Measurement of the Branching Fraction of Charmless Semileptonic B Meson Decays and the Quark-Mixing Matrix Element V_{ub} with the Belle II Experiment

Andrea Fodor

Department of Physics, Faculty of Science
McGill University
Montreal, Quebec, Canada

July 2024

A thesis submitted to McGill University in partial
fulfillment of the requirements of the degree of
Doctor of Philosophy

Abstract

This thesis presents the first inclusive untagged analysis of $B \to X_u \ell \nu$ decays using data from the Belle II experiment. Within the Standard Model theory of particle physics, these are weak decays of the B meson to a hadron containing an up quark, X_u, a charged lepton (either an electron or a muon), ℓ, and the corresponding lepton neutrino, ν. The decay rate is governed by the Cabbibo-Kobayashi-Maskawa quark mixing matrix element V_{ub}. The value of the V_{ub} matrix element is one of the fundamental parameters of nature, a quantity that cannot be theoretically predicted and therefore must be measured. The V_{ub} value can be extracted based on the measured $B \to X_u \ell \nu$ partial branching fraction. In this thesis, the measurement of the $B \to X_u \ell \nu$ branching fraction is done inclusively, whereby the outgoing hadron is not reconstructed. Only the charged lepton is selected. To avoid the dominant $B \to X_c \ell \nu$ background, in which X_c is a hadron containing a charm quark, the signal yield is measured in the lepton momentum endpoint region, between 2.1 and 2.7 GeV. For the first time in the B-factory era of flavour physics studies, this measurement is uniquely performed using both the electron and muon final states, an approach not used before in

previous B-factory inclusive measurements of these decay modes.

The analysis makes use of $(198.0 \pm 3.0) \times 10^6$ B meson pairs recorded by the Belle II experiment. The Belle II experiment is a B-factory experiment located at the SuperKEKB electron-positron collider in Japan. It started collecting physics-quality collision data in 2019.

Using the recorded data and a Monte Carlo simulated data sample, the signal selection is developed. The signal yield is measured using a fit to the lepton momentum, and the systematic uncertainties are evaluated. The analysis procedure is validated by measuring the hadron-flavour agnostic $B \to X\ell\nu$ branching fraction. In this thesis, the sensitivities for the measurement of the $B \to X_u\ell\nu$ partial branching fraction and the subsequently extracted V_{ub} value are presented. The expected experimental uncertainty for the partial branching fraction in the electron channel, $\Delta\mathcal{B}(B \to X_u e\nu)$, is $0.14\%_{\text{stat}} \pm 15.5\%_{\text{syst}}$, and in the muon channel, $\Delta\mathcal{B}(B \to X_u \mu\nu)$, is $0.16\%_{\text{stat}} \pm 12.0\%_{\text{syst}}$. The sensitivity in this early data-taking period is limited by systematic effects due to dominant particle identification uncertainties. The overall systematic uncertainty is expected to improve as the Belle II experiment collects more data and deeper understanding of the detector performance, motivated in part by this first study, is established.

Abrégé

Cette thèse présente la première analyse inclusive non étiquetée des désintégrations $B \rightarrow X_u \ell \nu$ en utilisant les données de l'expérience Belle II. Dans le cadre de la théorie du Modèle Standard de la physique des particules, il s'agit de désintégrations faibles du méson B vers un hadron contenant un quark up, X_u, un lepton chargé (soit un électron, soit un muon), ℓ, et du neutrino correspondant, ν. Le taux de désintégration est régi par l'élément V_{ub} de la matrice Cabbibo-Kobayashi-Maskawa qui mélange les différentes saveurs de quarks. La valeur de l'élément de matrice V_{ub} est l'un des paramètres fondamentaux de la nature, une quantité qui ne peut être prédite théoriquement et doit donc être mesurée. La valeur de V_{ub} peut être extraite de la fraction de désintégration partielle mesurée $B \rightarrow X_u \ell \nu$. Dans cette thèse, la mesure de la fraction de désintégration $B \rightarrow X_u \ell \nu$ est effectuée de manière inclusive, où le hadron sortant n'est pas reconstruit. Seul le lepton chargé est sélectionné. Pour éviter le fond dominant $B \rightarrow X_c \ell \nu$, dans lequel X_c est un hadron contenant un quark charm, le rendement du signal est mesuré dans la région du moment du lepton entre 2,1 et 2,7 GeV. Pour la première fois dans l'ère des usines à B

des études de physique des saveurs, cette mesure est effectuée de manière unique en utilisant à la fois les états finaux électronique et muonique, une approche jamais utilisée auparavant dans les mesures inclusives de ces modes de désintégration dans les usines à B. L'analyse utilise $(198.0 \pm 3.0) \times 10^6$ paires de mésons B enregistrées par l'expérience Belle II. L'expérience Belle II est une expérience d'usine à B située au collisionneur électron-positron SuperKEKB en Japon. Elle a commencé à collecter des données de collision de qualité assez élevée pour l'analyse en 2019. La sélection du signal est développée en utilisant les données enregistrées et un échantillon de données simulées Monte Carlo. Le rendement du signal est mesuré en ajustant une courbe au moment du lepton, et les incertitudes systématiques sont évaluées. La procédure d'analyse est validée en mesurant la fraction de désintégration $B \to X\ell\nu$ agnostique à la saveur des hadrons. Dans cette thèse, les sensibilités aux mesures de la fraction de désintégration partielle $B \to X_u\ell\nu$ et de la valeur de V_{ub} extraite par la suite sont présentées. L'incertitude expérimentale attendue pour la fraction de désintégration partielle dans le canal électronique, $\Delta\mathcal{B}(B \to X_u e\nu)$, est de $0.14\%_{\text{stat}} \pm 15.5\%_{\text{syst}}$, tandis que dans le canal muonique, $\Delta\mathcal{B}(B \to X_u \mu\nu)$, elle est de $0.16\%_{\text{stat}} \pm 12.0\%_{\text{syst}}$. La sensibilité au cours de cette période de prise de données précoce est limitée par les effets systématiques liés aux incertitudes sur l'identification des particules qui dominent. L'incertitude systématique globale devrait s'améliorer à mesure que l'expérience Belle II collecte plus de données et qu'une compréhension plus approfondie des performances du détecteur, motivée en partie par cette première étude, est établie.

Contributions of the Author

The Belle II experiment is an international collaboration of more than 120 universities and institutes, consisting of more than 1000 members. The hardware and software used for the running of the experiment are products of collaborative work spanning over more than a decade. The data and the Belle II specific software tools used for the analysis presented in this thesis have been produced and developed by the Belle II collaboration.

The analysis of the semileptonic inclusive decay $B \to X_u \ell \nu$, using the data recorded by the Belle II detector and the simulation produced by the Belle II collaboration, is the work of the author. This type of analysis has not been done at the Belle II experiment before and is work of the author providing distinct contributions to knowledge. The author is the only analyst that worked on this analysis. Advice and guidance have been given by Prof. Andreas Warburton and Dr. Raynette van Tonder, postdoctoral researcher, based on their previous experience with particle physics analyses. The author also received valuable guidance and advice from the Belle II Semileptonic Working Group members, namely the conveners of the group, Dr. William Sutcliffe, Dr. Lu Cao, Dr. Christoph Schwanda, Dr. Racha Cheaib,

and Prof. Florian Bernlochner.

A separate earlier analysis of the $B \rightarrow X_u e \nu$ electron momentum endpoint region was performed solely by the author and represented the 'rediscovery' of charmless B decays using the Belle II experiment. This result was presented at the International Conference for High Energy Physics in 2020 (ICHEP2020) [1, 2], but was not officially published. This preliminary result was reviewed by a committee of Belle II members: Dr. Markus Prim, Prof. Kyung Kwang Joo, and Prof. Romulus Godang. Several of the foundational techniques used in the analysis presented in this thesis draw from this initial $b \rightarrow u$ rediscovery exercise.

The summary of contributions of the author for each thesis chapter is given below:

- Chapter 1: The author gives an introduction and motivation for the analysis presented in the thesis.

- Chapter 2: The author presents an introduction to the theoretical framework of particle physics that is relevant to this analysis.

- Chapter 3: The author provides a summary of the previous measurements of semileptonic B meson decays performed at other particle physics experiment.

- Chapter 4: The author gives an overview of the Belle II experiment and the detector used for collecting the data used in this thesis.

  The author contributed to the data taking at the Belle II experiment by performing control room shifts, both on site and remote. The author also partook in the data reconstruction, by taking online quality monitoring shifts.

  The author significantly contributed to the estimation of the beam background levels in the electromagnetic calorimeter of the Belle II detector. The author analyzed

beam background simulations and some of the first data coming from the Belle II experiment for this purpose. This work will not be presented in this thesis.

- Chapter 5: The author lays out the analysis method and provides an overview of the Belle II data, simulation and the Belle II proprietary software used for the analysis. The simulation and the software are the result of work of numerous members of the Belle II collaboration. The analysis method was developed by the author, by examining previous similar analyses done at other experiments and with advice and guidance from previously mentioned collaborators. The analysis results are presented blinded for this thesis, where only the relative uncertainties are shown. The author explains the blinding process. For the application of the hybrid model to the signal simulation, the author received guidance from Dr. Markus Prim and Prof. Florian Bernlochner. The software package for the application of the hybrid modelling weights was developed by Dr. Markus Prim.

- Chapter 6: The author explains the signal selection procedure. The signal selection was developed, tested and applied to data and simulation by the author. The data and simulation correction factors have been derived by the Belle II collaboration and applied by the author. The software module used for applying the particle identification corrections was developed by the collaboration. The author trained and tested multivariate algorithms for continuum background suppression.

- Chapter 7: The author explains how the signal yield is extracted using a fit to data. The fitting algorithm was designed and tested by the author. Guidance while developing the fitting algorithm was provided by Dr. Raynette van Tonder. The fitting software used was developed by the Belle II collaboration.

- Chapter 8: The author explains the systematic uncertainty sources for the signal yield. The systematic uncertainties were calculated by the author.

- Chapter 9: The author demonstrates the calculation of the partial branching fraction of $B \rightarrow X_u \ell \nu$ decay and the $|V_{ub}|$ value. The author calculated the experimental uncertainties on these quantities. The author also performed and presents a cross-check of the analysis methodology by measuring the flavour-agnostic $B \rightarrow X \ell \nu$ inclusive branching fraction. The author compares the obtained results to the ones of previous experiments on this subject. The author outlines the next steps in the analysis unblinding process.

- Chapter 10: The author provides a summary of the analysis performed for this thesis and comments on future work.

Acknowledgements

I want to convey my deepest gratitude to my supervisor, Prof. Andreas Warburton, for his constant support and patient guidance in performing this analysis. I am grateful to the Belle II collaboration members for enabling me to work on this analysis. Special thanks to Raynette van Tonder for her valuable advice and support in performing the analysis. Thank you to Prof. Steven Robertson who led the McGill group over the years and provided valuable insight during our group meetings. Thank you to the other members of the McGill Belle II group for their help: Hannah Wakeling, Trevor Shillington, Kai-Ning Chu, Robert Seddon, Racha Cheaib, and others that have come and gone over the years. Thanks to the multitude of my McGill peers, especially to Shreya Saha and Bryce Cyr, for being there from the start and making the graduate studies experience more wholesome. My friends on both side of the ocean: Sanja, Aleksandra, Nikolina, Ana, Manca and others.

To my parents Ede and Gordana, my sister Tamara, for unconditional support and understanding. To my husband, Mihajlo, for being there through the ups and downs throughout this experience, and for the constant encouragement. And to our son Teodor, who was with me throughout the writing and has waited patiently for me to finish before his big arrival.

Contents

List of Figures

xv

List of Tables

List of Acronyms

- ARICH: Aerogel Ring Imaging Cherenkov

- AUC: Area Under the Curve

- BDT: Boosted Decision Tree

- BLNP: Bosch Lange Neubert Paz model

- CDC: Central Drift Chamber

- CKM matrix: Cabbibo-Kobayashi-Maskawa matrix

- CMS: Center-of-Mass System

- CP: Charge-Parity

- DAQ: Data Acquisition

- DEPFET: Depleted Field Effect Transistor

- DFN: De Fazio Neubert model

- DSSD: Double-Sided Silicon Strip Detector

- DT: Decision Tree

- ECL: Electromagnetic Calorimeter

- FOM: Figure of Merit

- HAPD: Hybrid Avalanche Photon Detector

- HER: High Energy Ring

- HFLAV: Heavy Flavour Averaging Group

- HLT: High Level Trigger

- HQET: Heavy Quark Effective Theory

- IP: Interaction Point

- IR: Interaction Region

- KLM: Long Lived Kaon and Muon detector

- LER: Low Energy Ring

- MC: Monte Carlo

- MVA: Multi-Variate Analysis

- NP: New Physics

- PDG: Particle Data Group

- PID: Particle Identification

- PMT: Photomultiplier Tube

- PXD: Pixel Vertex Detector

- QCD: Quantum Chromodynamics

- ROC: Receiver Operating Characteristic

- RPC: Resistive Plate Chamber

- SF: Shape Function

- SM: Standard Model

- SVD: Silicone Vertex Detector

- TOP: Time Of Propagation

- VXD: Vertex Detector

1

Introduction

The Standard Model (SM)(some of the seminal works that established the SM theory are [3–6]) of particle physics is currently the most successful theory describing the elementary particles and their interactions. The SM predictions have been experimentally verified at a wide range of energy scales. The most recent experimental observation of the Higgs boson [7] completes the particle content of the SM. Still, there are various questions that the SM cannot answer. Based on cosmological observations, it is concluded that the SM only describes about 5% of the energy density of the universe [8], while the remainder is attributed to 'dark matter' and 'dark energy'. The matter-antimatter imbalance in the universe is also not understood. Another issue is that the SM has at least 26 free parameters that cannot be theoretically predicted, but need to be experimentally determined.

Weak decays played a significant role in the development of the SM. The weak processes enable the transitions between quarks of different flavours. The Cabbibo-Kobayashi-Maskawa (CKM) matrix parametrizes the probability of these quark transitions. The CKM matrix is the source of the charge-parity symmetry violation in the SM, albeit this violation

1

is not sufficient to explain the dominance of matter over antimatter in the universe. Measurements of the CKM matrix elements are the cornerstone of flavour physics experiments. The values of the CKM matrix elements cannot be predicted by theory and need to be measured. They are necessary for a better understanding of the SM and are a good avenue for New Physics (NP) searches.

The V_{ub} element of the CKM matrix can be experimentally determined in $B \rightarrow X_u \ell \nu$ decays, where the B meson is a particle composed of a b quark and a u or d anti-quark[1], and X_u is a hadron made up of a u quark and a u or d anti-quark, ℓ is a charged lepton and ν is the corresponding lepton neutrino. In this type of B meson decays, the b quark transforms to a u quark via a weak interaction. This transition is described by the V_{ub} CKM matrix element. These so-called charmless decays are rare compared to the decays involving the final state hadron with a charm quark, $B \rightarrow X_c \ell \nu$, which constitute the main background in $B \rightarrow X_u \ell \nu$ searches. The $B \rightarrow X_c \ell \nu$ decay rate is described by the V_{cb} CKM matrix element, which is an order of magnitude larger than the V_{ub} element, making this decay more probable.

The first detection of charmless semileptonic B meson decays came in the early '90s from the CLEO [9] and ARGUS [10] experiments. Since then, the Belle and BaBar experiments made precise measurements of the $B \rightarrow X_u \ell \nu$ decay rate and the $|V_{ub}|$ value. Tensions are present in the $|V_{ub}|$ results between the inclusive and exclusive analysis approaches [11]. The analyses taking an exclusive approach reconstruct the specific final state hadron, while the inclusive analyses do not require reconstruction of a specific hadron in the final state. This thesis presents the measurement of the inclusive $B \rightarrow X_u \ell \nu$ decay rate using data from the novel Belle II experiment. The Belle II experiment is an electron-positron collider

---

[1]Charge conjugation is implied throughout this document.

experiment located in the KEK laboratory in Tsukuba, Japan. The electron and positron beams are collided at at a centre-of-mass energy corresponding to the $\Upsilon(4S)$ mass resonance, where B meson pairs are copiously produced, making Belle II a 'B-factory'. The amount of recorded data at the Belle II experiment, during its planned running, will be higher than at any previous B-factory experiment. Specifically, this analysis uses data consisting of 197 million B meson pairs from the early data taking period of the Belle II experiment, recorded between 2019 and 2021.

This thesis presents an inclusive $B \to X_u \ell \nu$ analysis, with the goal of measuring the partial decay rate in the lepton momentum endpoint region, between[2] 2.1 and 2.7 GeV. Only the final state charged lepton is selected. No requirements are made on the final state hadron, while the final state neutrino escapes detection. The reconstructed final-state charged lepton is either an electron or a muon, $\ell = e, \mu$. The analysis is done using an untagged approach, where the second B meson created in the event is not reconstructed. This approach was chosen to increase the selection efficiency of the signal mode, thereby enhancing the statistics. The dominance of the $B \to X_c \ell \nu$ background is avoided by extracting the signal yield in the momentum endpoint of the charged lepton. Because the u quark is much lighter than the c quark, due to energy conservation the leptons from B decays with a b to u conversion have more energy available than those from B decays with a b to c conversion. By examining the lepton momentum endpoint, I measure the partial branching fraction of the charmless semileptonic B meson decays. Once the branching fraction is obtained, the value of the V_{ub} matrix element is calculated. The partial branching fraction and the $|V_{ub}|$ calculation are done separately for two theoretical models for the $B \to X_u \ell \nu$ decay.

The analysis is performed blinded, whereby the signal yield in the region of interest –

---

[2]Natural units are used, $\hbar = c = 1$.

the lepton momentum endpoint – is not known to the analyst. This is done to avoid experimental bias. The final results in this thesis are not shown; only the estimated relative experimental uncertainties of the measurements are presented. This is done to avoid any unofficial results becoming public before the official approval from the Belle II collaboration.

This analysis is inspired by previous untagged inclusive $B \to X_u \ell \nu$ analyses performed at the BaBar experiment, namely [12] and [13]. The BaBar results only looked at the electron final states, while here the muon final states are included as well. This is the first measurement of the $B \to X_u \ell \nu$ partial branching fraction in the charged lepton momentum endpoint performed at the Belle II experiment. This analysis also serves as a benchmark for examining the performance of the Belle II detector.

The outline of this thesis is as follows: Chapter 2 gives the reader an overview of the Standard Model of particle physics, as well as an overview of semileptonic B meson decays and how they can be used for measuring the values of the CKM matrix elements V_{ub} and V_{cb}. Chapter 3 goes over the previous measurements of the $B \to X_u \ell \nu$ decays, both the inclusive and exclusive. The Belle II experiment, with the SuperKEKB collider and the Belle II detector, is introduced in Chapter 4. Chapter 5 outlines the analysis strategy. It describes the Belle II recorded dataset and the Monte Carlo simulated data, as well as the Belle II software framework used for the analysis. The signal selection requirements are described in Chapter 6, along with the corrections applied to data and Monte Carlo simulation. Next, in Chapter 7, the fit that is performed on data using Monte Carlo simulation as fit templates in order to extract the signal yield is described. The estimation of systematic uncertainties on the signal yield in the lepton momentum range from 2.1 to 2.7 GeV is outlined in Chapter 8. The procedures for the partial branching fraction and the $|V_{ub}|$ calculations are

explained in Chapter 9. Since the analysis is presented blinded, the expected relative experimental uncertainties are presented, without the final central values. The analysis procedure is validated by measuring the total branching fraction of the $B \to X\ell\nu$ decay, where X can be any flavour of hadron(s). Finally, in Chapter 10, the analysis is summarized and the results are discussed in comparison to previous $B \to X_u\ell\nu$ measurements. Plans for the next steps in this analysis are also outlined.

2

Theoretical introduction

This chapter will provide a general overview of the Standard Model (SM) that is necessary to better understand the motivation and the results of the work presented in this thesis. The reader is referred to, e.g. [8, 14, 15] for an overview of the SM theory. A more specific overview of the semileptonic B meson decays will be provided, explaining how these decays are suitable for the measurement of the Cabbibo-Kobayashi-Maskawa (CKM) matrix elements V_{ub} and V_{cb}.

The Standard Model is, at present, the most successful theory that describes the matter in our universe and its interactions via the electromagnetic, weak, and strong forces. It is a gauge theory based on the symmetry group $SU(3) \times SU(2) \times U(1)$. The $SU(3)$ group represents the strong force, while $SU(2) \times U(1)$ represents the electroweak force. In the Standard Model, the matter particles interact by exchanging gauge bosons, spin 1 particles that are the force carriers of the fundamental forces. The strong force carriers are gluons, the $W^{\pm}$ and Z^0 bosons are the carriers of the weak force, and photons are the carriers of the electromagnetic force. There are twelve spin $1/2$ fermions, six quarks and six leptons,

that make up visible matter. These fermions all have their anti-matter partners, so there are six anti-quarks and six anti-leptons, which have the opposite quantum numbers from their matter counterparts. The most recently detected particle of the Standard Model is the Higgs boson, which interacts with the gauge bosons and fermions and gives them mass through the Higgs mechanism [16].

While the Standard Model predictions describe the experimental observations well, it has its limitations and it is not a complete theory. First, the gravitational force is not described in the Standard Model. The Standard Model has at least 26 parameters (masses, mixing angles, couplings) which are not predicted by theory and can only be measured. Also, astrophysical observations show that the visible matter that is described by the Standard Model only makes up about 5% of the mass of the universe. Another 25% of the universe is made up of a type of matter that we call Dark Matter [17], since it has been detected only through the gravitational interaction. The additional 70% of the universe is made up of Dark Energy [18, 19]. The Standard Model does not explain the presence of Dark Matter nor Dark Energy. The Standard Model also cannot completely account for the dominance of matter over anti-matter in the universe.

2.1   Leptons and quarks

The visible matter is made up of two types of elementary spin-1/2 particles: leptons and quarks. The SM describes 24 elementary fermions in total: 6 leptons, 6 quarks, as well as their respective anti-particles that have the same mass, lifetime and spin, but opposite sign of electric and other charges.

In the Standard Model, there are six types of leptons: electron (e^-), muon (μ^-) and the tau lepton (τ^-), with their associated neutrinos, the electron neutrino (ν_e), the muon neu-

2.1 Leptons and quarks

trino (ν_μ) and the tau neutrino (ν_τ). The e^-, μ^- and τ^- have an electrical charge $q = -1$, while the neutrinos are electrically neutral, $q = 0$. The leptons do not carry the strong force charge ('colour'), so they cannot interact via the strong force. The charged leptons can interact via the electromagnetic force, by photon exchange, and the weak force, by $W^\pm$ or Z^0 exchange, while the neutrinos can only partake in the weak force interactions.

On the other hand, there are also six types of quarks: up (u), down (d), charm (c), strange (s), top (t) and bottom (or beauty) (b). They have the electric charge of either $q = +2/3$ or $q = -1/3$. The quarks also carry colour charge. Thus, they can interact via the electromagnetic force, the weak force, as well as the strong force. The quarks cannot be isolated due to the colour confinement phenomenon of quantum chromodynamics. They can only be found in colour-singlet bound states. Baryons are bound states of three quarks (qqq), like the proton ($p = uud$) or neutron ($n = udd$). Mesons are bound states of a quark and an anti-quark ($q\bar{q}$). Some examples of mesons are pions ($\pi^+ = u\bar{d}$, $\pi^0 = u\bar{u}$ or $d\bar{d}$), kaons ($K^+ = u\bar{s}$, $K^0 = d\bar{s}$), and B mesons ($B^+ = u\bar{b}$, $B^0 = d\bar{b}$). More exotic bound states have also been detected, such as tetraquarks ($qq\bar{q}\bar{q}$) and pentaquarks ($qqqq\bar{q}$) [8].

Leptons and quarks are divided into three generations that can be written in the form of $SU(2)_Y$ doublets:

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix} \tag{2.1}$$

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}. \tag{2.2}$$

Some of the properties of quarks and leptons, along with the gauge bosons, are summarized in Fig. 2.1.

Standard Model of Elementary Particles
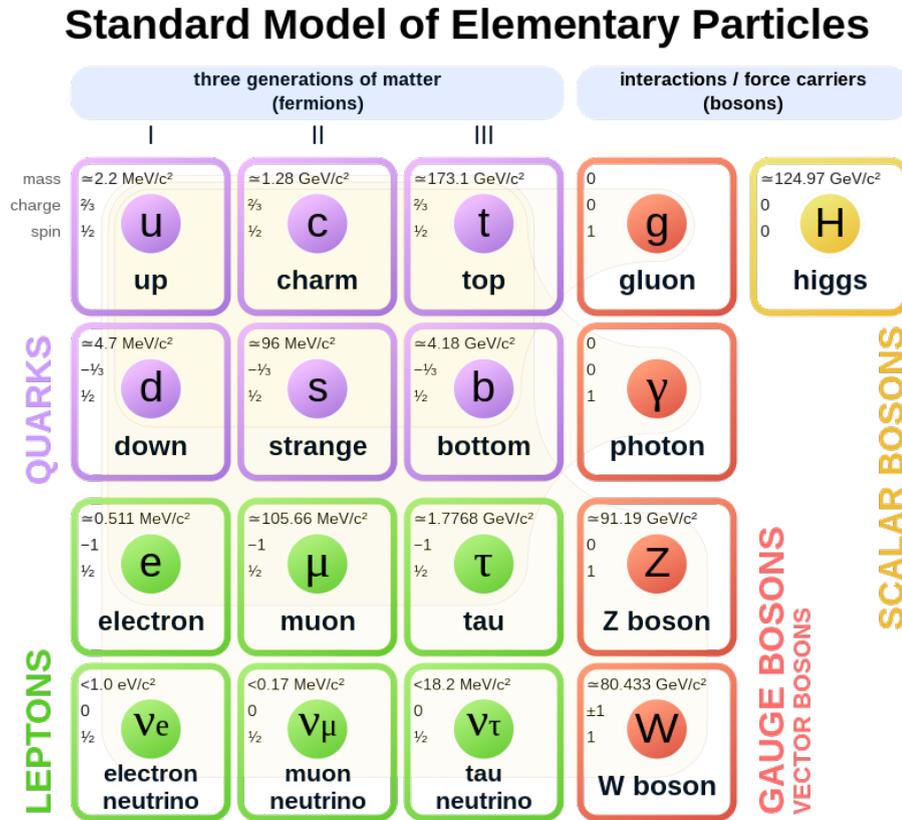
Figure 2.1: The elementary particles of the Standard Model and information on their mass, charge and spin. The quarks are shown in purple; the leptons are shown in green. The force carriers, gauge bosons, are shown in red. The Higgs boson is shown in yellow. Figure from [20].

2.2 Force-carrying bosons

Spin-1 gauge bosons mediate interactions in the Standard Model. Their number is equal to the number of generators of the relevant symmetry group. The strong force has 8 mediators that correspond to 8 generators of the $SU(3)_C$ symmetry. These mediators are called gluons, $g_\alpha, \alpha = 1, ..., 8$. The electromagnetic force has $U(1)$ symmetry, so it has one gauge

2.2 Force-carrying bosons

boson mediator, the photon (γ). The weak force carriers are three gauge bosons equivalent to the generators of the $SU(2)$ group, $W^\pm$ and Z^0.

The range of the force is inversely proportional to the mass of its mediating gauge boson. This means that the electromagnetic force, with a massless and electrically neutral photon as the carrier, has an infinite range. Gluons are also massless particles, but they carry the colour charge. Hence, the strong force does not have an infinite range. Unlike photons, gluons can interact amongst themselves because they carry colour charge. This makes the effective range of the strong force about 10^{-15} m, roughly the size of an atomic nucleus. The weak force carriers, $W^\pm$ and Z^0, are massive particles with mass of 80.4 and 91.2 GeV, respectively [8]. Because of this, the effective range of the weak force is very short, at 10^{-17} m.

The Higgs boson is the scalar boson in the Standard Model, whose existence was confirmed experimentally only a decade ago at the LHC [7]. It is a spin-0 particle with a mass of 125 GeV. The non-zero mass of the $W^\pm$ and Z^0 bosons indicated the existence of a field, the Higgs field, with a non-zero vacuum expectation value which would cause spontaneous electroweak symmetry breaking. At low energies, the electroweak $SU(2)_L \times U(1)_Y$ symmetry gets broken to the electromagnetic subgroup $U(1)_Q$. Four Goldstone bosons are required for the spontaneous symmetry breaking of the gauge group. Three of the four Goldstone bosons mix with the $W^\pm$ and Z^0, giving them mass, while the photon remains massless. The last component of the Goldstone boson is the spin-0 Higgs boson. Quarks and leptons acquire mass via Yukawa couplings to the Higgs field. The interactions of the fermions and the Higgs field cause the fermions to flip chirality (e.g. flip between right-handed and left-handed). So far, only left-handed neutrinos have been detected, which means they can't acquire mass through the Yukawa interactions with the Higgs field. From

detecting neutrino oscillations [21, 22], we know that the neutrinos need to be massive, but the value and the exact origin of their mass is still unknown.

2.3   Weak interaction and the CKM matrix

All fermions of the Standard Model experience the weak interaction, carried by the $W^{\pm}$ and Z^0 bosons. All the charged leptons are equivalent under the weak interaction, a property known as lepton universality. However, the same is not true for quarks. The quarks can change flavour mediated by the weak interaction. This phenomenon is known as quark mixing. It occurs because the mass eigenstates of quarks are not the same as their weak eigenstates. The Cabbibo-Kobayashi-Maskawa (CKM) matrix [23] relates the mass eigenstates to the weak eigenstates:

$$
\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}
\tag{2.3}
$$

The probability of a flavour transition is proportional to the magnitude of the corresponding CKM matrix element. The on-diagonal elements represent the flavour changing processes within the same generation of quarks and they are close to 1. The off-diagonal elements represent the flavour changes between generations and they are much smaller, meaning that such transitions are more rare. The current world-average measured values of the CKM

11

matrix elements are as follows [8]:

$$|V_{CKM}| = \begin{pmatrix} 0.97370 \pm 0.00014 & 0.2245 \pm 0.0008 & (3.82 \pm 0.24) \times 10^{-3} \\ 0.221 \pm 0.004 & 0.987 \pm 0.011 & (41.0 \pm 1.4) \times 10^{-3} \\ (8.0 \pm 0.3) \times 10^{-3} & (38.8 \pm 1.1) \times 10^{-3} & 1.013 \pm 0.030 \end{pmatrix}.$$

(2.4)

At tree level, the quark mixing occurs via the exchange of a single $W^{\pm}$ boson. In this exchange, a down-type quark with charge $q = -1/3$ changes into an up-type quark with charge $q = +2/3$ by emitting a W^- boson. Thus, these types of processes are called flavour changing charged currents. Flavour changing neutral current processes, where an e.g. up-type quark changes into another up-type quark of a different generation, are not allowed in the Standard Model at the tree level. They can only happen at the loop level, which means that multiple gauge bosons are exchanged in multiple interactions, and are thus highly suppressed.

The CKM matrix is a unitary ($VV^{\dagger} = I$) matrix due to the probability conservation. The unitarity condition can be rewritten as:

$$\sum_i V_{ij} V_{ik}^* = \delta_{jk}$$

(2.5)

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

(2.6)

The parity (P) operation changes a left-handed particle into a right-handed particle. The weak interactions can only occur between left-handed particles or right-handed anti-particles, which means that the weak interaction violates the P symmetry. Instead, it was thought that the weak interaction respects the charge-parity (CP) symmetry. The CP operation transforms a left-handed particle into its corresponding anti-particle. The CP symmetry viola-

tion was first discovered in $K^{\pm}$ decays [24]. This means that if a particle is interchanged with its anti-particle with opposite parity, their interactions under the weak force will not be the same.

Based on the unitarity requirement, the CKM matrix can be reduced to three real parameters and one complex phase. The complex phase accounts for the CP violation in the quark sector. The CKM matrix can be written as[25]:

$$
V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}\exp^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}\exp^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}\exp^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}\exp^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}\exp^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (2.7)
$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ for $i < j = 1,2,3$. The angles $\theta_{12}, \theta_{13}, \theta_{23}$ are the mixing angles between the quark generations and the phase δ describes the CP violation. Taking into account the hierarchy between the mixing angles, the CKM matrix can be rewritten in terms of the Wolfstein parametrization [26] in terms of the four parameters λ, A, ρ, η, such that $s_{12} \equiv \lambda$, $s_{23} \equiv A\lambda^2$ and $s_{13}\exp^{-i\delta} \equiv A\lambda^3(\rho - i\eta)$. By making an expansion in λ, the CKM matrix can be written as:

$$
V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}\left(\lambda^4\right). \quad (2.8)
$$

The unitarity conditions can be graphically represented in the form of unitarity triangles. The most common representation of the unitarity triangle, originating from Eq. 2.6 and the Wolfstein parametrization, is shown in Fig. 2.2. Dividing Eq. 2.6 by $V_{cd}V_{cb}^*$, a triangle is obtained, with a unit base and vertices at $(0,0)$, $(0,1)$, and (ρ, η). The sides and angles of

the unitarity triangle can be experimentally determined through measurements of various processes. By measuring the quantities that make up the sides and the angles of the unitarity triangle, it is possible to test the unitarity of the CKM matrix. For example, $\sin 2\alpha$ can be constrained by measuring CP violation in $B^0 \rightarrow \rho^+ \rho^-$ decays; $\sin 2\beta$ can be constrained by measuring CP violation in $B \rightarrow J/\psi K_S$ decays; $|V_{ub}|$ and $|V_{cb}|$ can be measured in semileptonic B meson decays [8]. The area of the unitarity triangle is a measure of the size of the CP violation.

From Eq. 2.7 and Eq. 2.8, it can be seen that the V_{ub} matrix element, which this thesis aims
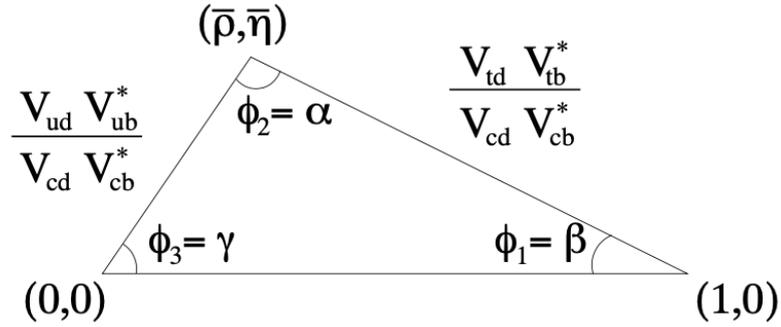
Figure 2.2: The standard unitarity triangle. Figure from [27].

to measure, depends on the CP violating phase. This makes it possible to probe the CP violation in the CKM matrix by measuring the V_{ub} value.

The CKM matrix is not the only source of the CP violation in the SM. The neutrino mixing showed that the neutrinos are massive particles, which makes the CP violation possible in the lepton sector as well. Another possible source of CP violation is in the strong interaction processes. However, there have been no strong interaction CP violating processes observed in any experiments. The CP asymmetry arising from the SM alone is not enough to account for the matter-anti-matter imbalance in the universe. Nevertheless, most NP scenarios that could explain the CP violation have observable effects in the quark-

flavour experiments [28].

2.4   Semileptonic B meson decays

The semileptonic decays of B mesons of the type $B \to X_c \ell \nu$ and $B \to X_u \ell \nu$ are the main avenue for measuring the values of the CKM matrix elements V_{cb} and V_{ub}, respectively. These decays proceed via leading order weak interactions, and no significant contributions from non-Standard-Model processes are typically expected. This makes them ideal candidates for the determination of the V_{cb} and V_{ub} values.

These semileptonic decays proceed via weak interaction, where the b quark converts to a c or a u by emitting a W boson. The W subsequently decays to a charged lepton and a lepton neutrino. The Feynman diagram of these processes is shown in Fig. 2.3.
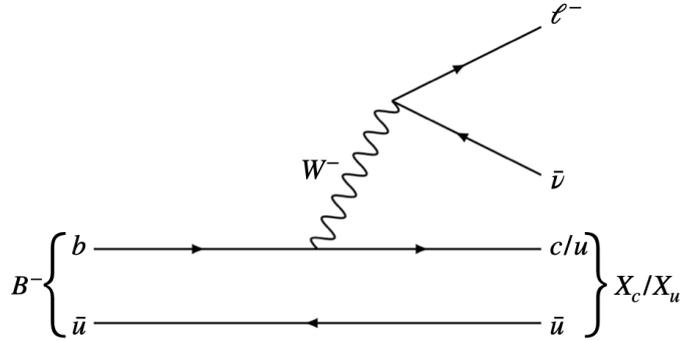
Figure 2.3: Feynman diagram of the semileptonic B meson decay ($B \to X \ell \nu$).

Within the SM, the charged-current semileptonic decays can be explained as an exchange of the $W^\pm$ boson between a quark and a lepton current. The quark current is given by [29]:

$$J_\mu = (\bar{\mathcal{U}}_L \gamma_\mu V_{CKM} \mathcal{D}_L), \tag{2.9}$$

2.4 Semileptonic B meson decays

where γ_μ are the Dirac gamma matrices, and $\mathcal{U}_L$ and $\mathcal{D}_L$ represent the left-handed up and down type quarks:

$$
\mathcal{U}_L = \begin{bmatrix} u_L \\ c_L \\ t_L \end{bmatrix}, \quad \mathcal{D}_L = \begin{bmatrix} d_L \\ s_L \\ b_L \end{bmatrix}.
\tag{2.10}
$$

Since the top quark does not form hadrons as it is too unstable, $\mathcal{U}_L$ can be written as:

$$
\mathcal{U}_L = \begin{bmatrix} u_L \\ c_L \\ 0 \end{bmatrix}.
\tag{2.11}
$$

The lepton current is written as:

$$
j_\mu = (\bar{e}_L \gamma_\mu \bar{\nu}_{e,L} + \bar{\mu}_L \gamma_\mu \bar{\nu}_{\mu,L} + \bar{\tau}_L \gamma_\mu \bar{\nu}_{\tau,L}).
\tag{2.12}
$$

The $W^\pm$ can be considered to be infinitely heavy compared to the masses of the quarks participating in the weak hadron decays. This leads to the effective Hamiltonian of the interaction:

$$
H_{\text{eff}} = \frac{4G_F}{\sqrt{2}}(\bar{\mathcal{U}}_L \gamma_\mu V_{CKM} \mathcal{D}_L) \times (\bar{e}_L \gamma_m u \bar{\nu}_{e,L} + \bar{\mu}_L \gamma_m u \bar{\nu}_{\mu,L} + \bar{\tau}_L \gamma_m u \bar{\nu}_{\tau,L}) + \text{h.c.},
\tag{2.13}
$$

where h.c. indicates the hermitian conjugate of the expression and G_F is the Fermi coupling constant:

$$
G_F = \frac{g^2}{4\sqrt{2}M_W^2},
\tag{2.14}
$$

with g being the coupling constant of the weak interaction, and M_W the mass of the W

gauge boson. Looking at the effective Hamiltonian, it is clear that the semileptonic decays can be used to access the values of all the CKM matrix elements, except the ones involving the top quark (V_{tq}). This way, semileptonic B meson decays of the type $B \rightarrow X_c \ell \nu$ and $B \rightarrow X_u \ell \nu$ can be used to access the V_{cb} and V_{ub} CKM matrix elements.

2.4.1 Decay kinematics

For the semileptonic B meson decays of the type $B \rightarrow X \ell \nu$, denoting the momentum of the B meson as p_B, the momentum of the charged lepton as p_ℓ and the momentum of the neutrino as p_ν, the momentum conservation can be written using 4-vectors as:

$$p_B = p_X + p_\ell + p_\nu,$$
$$p_B^2 = m_B^2, \quad p_X^2 = m_X^2, \quad p_\ell^2 = m_\ell^2, \quad p_\nu^2 = 0, \tag{2.15}$$

where m_B is the mass of the B meson, m_X is the mass of the outgoing hadron and m_ℓ is the mass of the charged lepton. For these purposes, it is sufficient to regard the neutrino mass as negligible.

For the fixed mass m_X, useful kinematic quantities are the momentum transfer squared, q^2, and the energy of the charged lepton, E_ℓ:

$$q^2 = (p_\ell + p_\nu)^2 = (p_B - p_X)^2, \quad m_\ell \leqslant q^2 \leqslant (m_B - m_X)^2,$$
$$E_\ell = \frac{p_B p_\ell}{m_B}, \quad m_\ell \leqslant E_\ell \leqslant \frac{1}{2m_B}(m_B^2 - m_X^2 + m_\ell^2). \tag{2.16}$$

Depending on the final state hadron, the leptons from semileptonic B meson decays have spectra with different endpoints. The final states with heavier charmed hadrons have lower lepton momentum endpoints compared to lighter charmless hadrons. Because the u quark

is lighter than the c quark, the charged leptons from $B \rightarrow X_u \ell \nu$ decays can have higher momenta due to energy conservation. This property can be experimentally exploited when measuring the $B \rightarrow X_u \ell \nu$, to avoid the dominant $B \rightarrow X_c \ell \nu$ background in the lower lepton momentum region. The lepton momentum distributions for $B \rightarrow X_u \ell \nu$ and $B \rightarrow X_c \ell \nu$ are shown in Fig. 2.4.
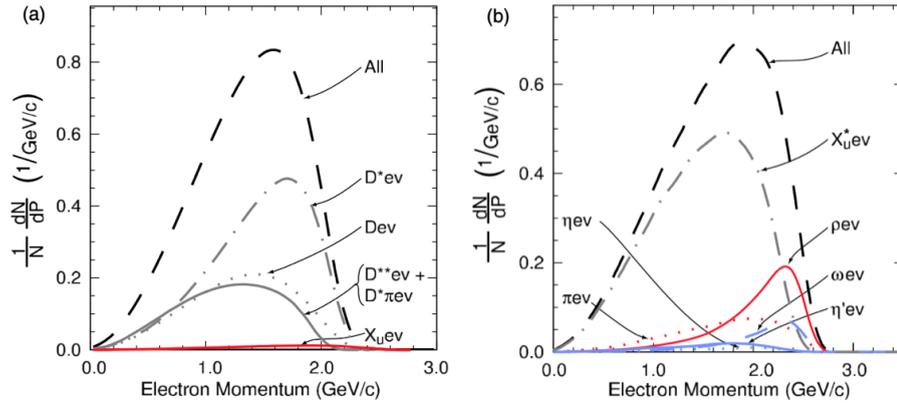
Figure 2.4: Monte Carlo generated electron momentum spectra for (a) $B \rightarrow X_c e \nu$ and (b) $B \rightarrow X_u e \nu$ decays. The $B \rightarrow X_u e \nu$ spectrum is shown in (a) with the red line to enable comparison with $B \rightarrow X_c e \nu$. In (b), different resonant decays of $B \rightarrow X_u e \nu$ are shown as labeled, with the full inclusive spectrum shown with the dashed black line. Figure taken from [29].

In the context of heavy-to-light quark transitions, where a heavy quark (c, b, t) transitions to a light quark (u, d, s), it is useful to define the light-cone components of the momenta:

$$P_- = E_X + |\vec{p}_X|, \quad P_+ = E_X - |\vec{p}_X|, \tag{2.17}$$

where p_X and E_X are the momentum and the energy of the outgoing hadron in the B meson

rest frame. The phase space for these quantities is given by:

$$\frac{m_X^2}{P_-} \leqslant P_+ \leqslant P_- \leqslant m_B. \tag{2.18}$$

From this point, the focus will be on the $B \to X_u \ell \nu$ semileptonic decays, which are the topic of this thesis. There are two approaches to the calculation of the $B \to X_u \ell \nu$ decay rate. One is the exclusive decay rate, where a specific final state hadron is considered. The other one is the inclusive decay rate, which includes all possible hadronic final states. These approaches are summarized in the next two sections.

2.4.2 Inclusive charmless semileptonic B meson decays

Both the $B \to X_c \ell \nu$ and $B \to X_u \ell \nu$ total inclusive decay rates can be can be calculated by using the Heavy Quark Expansion (HQE) method [30]. This method uses the fact that the mass of the b quark, m_b, is heavy compared to the QCD scale Λ_{QCD}. The QCD matrix elements are expanded in powers of Λ_{QCD}/m_b. This method can predict the total decay rate with uncertainties below 5% [31]. However, the total decay rate of the inclusive $B \to X_u \ell \nu$ process is difficult to measure due to the high backgrounds from the $B \to X_c \ell \nu$ process. The measurements can only be performed in the region of phase space where the $B \to X_c \ell \nu$ contributions are not dominant. For this reason, it is necessary to calculate the partial decay rates for the charmless inclusive decay in these limited phase space regions. Here, the HQE cannot be used, but it is necessary to introduce the "shape functions" (SF), non-perturbative distribution functions, which have an unknown form. The shape functions describe the motion of the b quark inside the B meson. They are important in the part of the phase space where the light-cone momentum component P_+ is not large compared to

Λ_{QCD}. Such is the case for the endpoint of the $B \to X_u \ell \nu$ lepton spectrum. At the leading order of $1/m_b$, a single shape function occurs. This shape function can be measured from $B \to X_s \gamma$ decays, as it is universal for all transitions of the b quark to a light quark. At the higher orders, multiple shape functions appear and they are not universal for different B meson decays. There are multiple theoretical approaches and calculations of the shape functions [32–34].

The triple differential decay rate for inclusive charmless semileptonic decays in the leading order can be written as [29]:

$$\frac{d^3\Gamma}{dP_+ dP_- dE_\ell} = \frac{G_F^2 |V_{ub}|^2}{192\pi^3} \int dk C\left(E_\ell, P_-, P_+, k\right) F(k) + \mathcal{O}\left(\frac{\Lambda_{QCD}}{m_b}\right), \qquad (2.19)$$

where $P_\pm$ are previously defined light-cone moments, E_ℓ is the energy of the lepton, m_b is the mass of the b quark. The coefficient $C(E_\ell, P_-, P_+, k)$ describes the decay $b \to u\ell\nu$ and can be calculated perturbatively. The function $F(k)$ is the non-perturbative shape function. A problem that arises when calculating the differential decay rate using Eq. 2.19 is that it does not predict any resonances in the hadronic invariant mass spectrum. For this reason, the phase space is usually analyzed by integrating out the resonant states. This problem is further explained in Chapter 5 in the context of $B \to X_u \ell \nu$ Monte Carlo simulation.

Several theoretical models of the inclusive $B \to X_u \ell \nu$ partial and total branching fractions are available and are used for the extraction of the $|V_{ub}|$ value. The two models used in this thesis are the ones developed by De Fazio and Neubert (DFN model) [32] and Bosch, Lange, Neubert and Paz (BLNP model) [35].

The DFN model provides next-to-leading order perturbative corrections to the triple differential decay rate from Eq. 2.19. It predicts the differential spectrum up to the order of
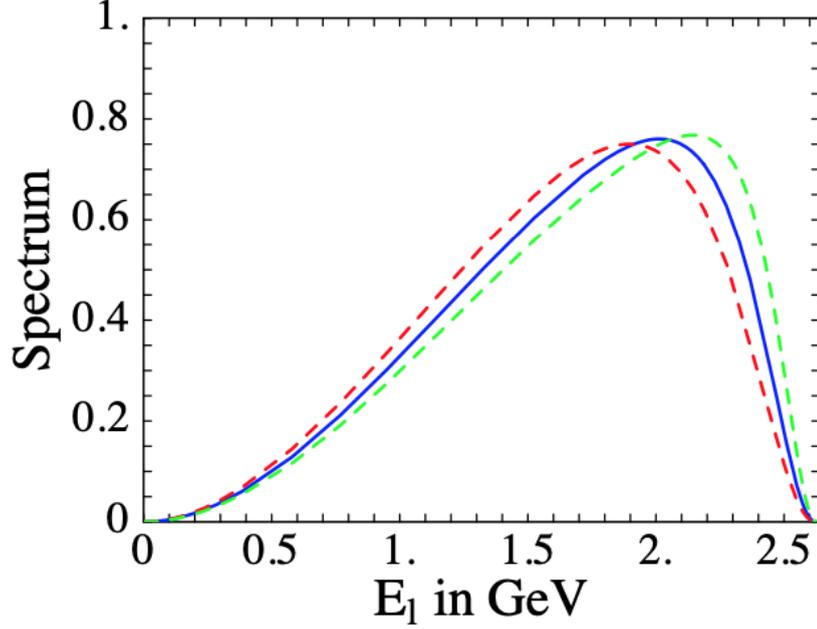
Figure 2.5: Energy spectrum of the charged lepton from $B \to X_u \ell \nu$ decays predicted by the DFN model. The different lines show the predictions for different m_b values: $m_b = 4.8$ GeV (solid blue), 4.65 GeV (dashed red) and 4.95 GeV (dashed green). Figure taken from [32].

α_s. This model parametrizes the leading-power non-perturbative shape function. The shape function is described with two parameters, $\bar{\Lambda}^{SF} = M_B - m_b$ and λ_1^{SF}. These parameters are determined from the measurements of the photon energy moments from $B \to X_s \gamma$ decays. In the phase-space endpoint region, where the non-perturbative corrections to the spectrum are large, the model predicts the decay distribution by a convolution of the parton model spectrum with the shape function. The charged lepton energy spectrum predicted by this model is shown in Fig. 2.5, where different lines show the predictions for different m_b values.

The BLNP model is a more recent model that includes all known perturbative and power corrections and interpolates between the HQE and shape function regions. The differential
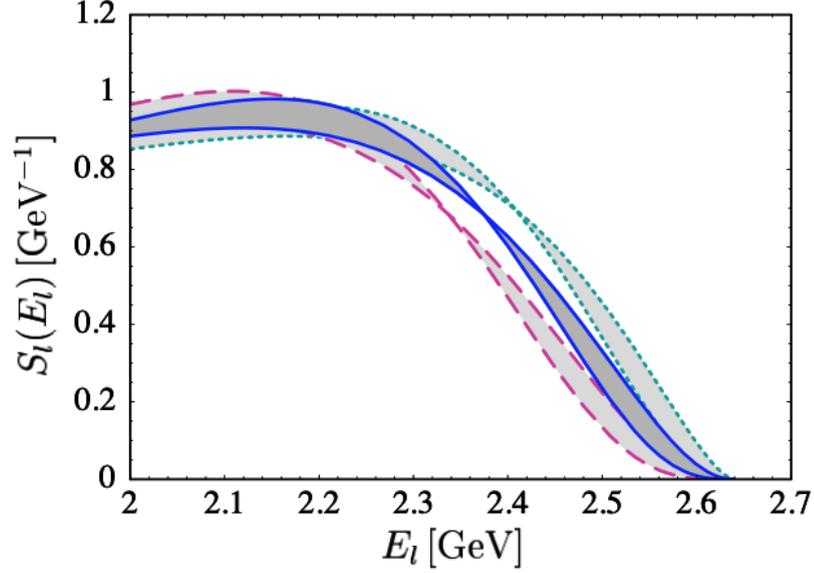
Figure 2.6: Energy spectrum of the charged lepton from $B \to X_u \ell \nu$ decays in the momentum endpoint, as predicted by the BLNP model. The different lines show the predictions for different $M_B - m_b$ values: 0.63 GeV (solid blue), 0.79 GeV (dashed red) and 0.56 GeV (dotted green). The width of the curves shows the sensitivity to the μ_π^2 value, taken between 0.2 and 0.34 GeV^2. Figure taken from [35].

decay rates are calculated at next-to-leading order in perturbative theory and at the leading power in HQE. The shape function is parametrized with two parameters, m_b and μ_π^2, where the μ_π^2 parameter corresponds to the kinetic energy of the heavy quark inside the heavy meson. The shape function parameters are determined by fits to moments of the hadron mass and lepton energy spectra from inclusive $B \to X_c \ell \nu$ decays and the photon energy moments in $B \to X_s \gamma$ decays. The predicted charged lepton energy spectrum in the endpoint region is shown in Fig. 2.6 for different values of $M_B - m_b$. Previous measurements of the inclusive $B \to X_u \ell \nu$ decays are summarized in Section 3.1.

2.4.3 Exclusive charmless semileptonic B meson decays

The exclusive semileptonic charmless decays, particularly $B \to \pi \ell \nu$, are another avenue for measuring the $|V_{ub}|$ value. Experimentally, examination of this type of decays is easier because the selection of the specific final hadron state eliminates the majority of the dominant $B \to X_c \ell \nu$ backgrounds. The downside is that the decay rates are only a fraction of the full inclusive decay rate, leading to low signal yields. Theoretically, the exclusive final state decay rates are possible to calculate, but different problems arise compared to the calculation of the inclusive decay rate.

The charmless final state hadrons can be divided into two categories: the pseudoscalar mesons (π^+, π^0, η and η') and the vector mesons (ρ^+, ρ^0, ω, ϕ). For the pseudoscalar meson final states, the vector current contributes, and for the vector final state, both the vector and the axial currents contribute. The matrix elements of these currents can be parametrized in terms of form factors. Form factors are scalar functions which describe the momentum transfer to leptons, q^2.

For the case of $B \to \pi \ell \nu$ decay, the total decay rate in the zero lepton mass limit ($m_\ell \to 0$) is [8]:

$$\frac{d\Gamma}{dq^2} = \frac{G_F^2 \left|V_{ub}\right|^2}{24\pi^3} \left|p_\pi\right|^3 \left|f_+\left(q^2\right)\right|^2, \tag{2.20}$$

where p_π is the momentum of the pion in the B rest frame and f_+ is the form factor. There are different theoretical approaches for the form factor estimation, many based on lattice-QCD (e.g. [36]).

Since this thesis focuses on the inclusive $B \to X_u \ell \nu$ decays, the exclusive decays will not be further explored here, but a more comprehensive summary is available in [8] and [29]. The next section explores the current experimental results of the semileptonic $B \to X_u \ell \nu$

2.4 Semileptonic B meson decays

decays and $|V_{ub}|$ measurements.

3

Previous measurements of $B \to X_u \ell \nu$

In this chapter, the previous measurements of the inclusive and exclusive $B \to X_u \ell \nu$ branching fractions and the $|V_{ub}|$ values are summarized. These measurements were performed at the previous $e^+ e^-$ collider experiments: CLEO, ARGUS, BaBar and Belle. A few selected measurements of V_{ub} from the LHCb experiment are also summarized.

The CLEO experiment [37] was operating for 30 years, from 1979 to 2008, making it the longest operating experiment in particle physics. It was located at Cornell University, at the CESR accelerator. It first discovered the $\Upsilon(4S)$ $b\bar{b}$ resonance [38], which is slightly above the threshold for $B\bar{B}$ meson pair production, making it the ideal state for the study of B mesons. CLEO operated at multiple energies, but a substantial part of the operations was dedicated to operation at the $\Upsilon(4S)$ resonance. The direct competitor of the CLEO experiment at the time was the ARGUS experiment [39], located at the DESY laboratory, at the DORIS II $e^+ e^-$ collider. It operated from 1982 until 1992. The CLEO and ARGUS experiments almost simultaneously reported the first observations of the charmless semileptonic B meson decays, confirming the Standard Model prediction of the non-zero $|V_{ub}|$ value.

Previous measurements of $B \to X_u \ell \nu$

The Belle [40] and BaBar [41] experiments are the next generation of experiments dedicated to flavour physics research. They both operated at the $\Upsilon(4S)$ resonance and due to their high luminosity were called B-factory experiments. They are also called asymmetric B factories due to the difference in energies of the electron and positron beams. The asymmetry in beam energies meant that the B meson pairs were created in motion, thus enabling the detection of individual B meson decay vertices. Using this approach, it was possible to measure the time dependence of B^0 and $\bar{B}^0$ meson decays and the CP violation. The Belle experiment was located at the KEKB e^+e^- collider at the KEK laboratory in Japan, with 3.5 GeV positron and 8.0 GeV electron beams. It operated from 1999 to 2010. The BaBar experiment was located at the PEP-II collider, at the SLAC laboratory in the USA, colliding 3.1 GeV positron and 9.0 GeV electron beams. It operated from 1999 to 2008. Even though these experiments are no longer collecting data, the analysis of the existing datasets is still ongoing. Belle and BaBar experiments first observed the CP symmetry violation in the B meson sector [42, 43]. This experimental observation confirmed the earlier postulated theory of Kobayashi and Maskawa [23], which led to them being awarded the Nobel prize for their work in 2008.

The e^+e^- collider experiments operating at the $\Upsilon(4S)$ resonance primarily produce one $B\bar{B}$ pair for each event. One of the B mesons in the event is reconstructed as signal. The untagged analyses do not do any reconstruction for the companion B meson. The tagged analyses reconstruct the companion B meson using pre-defined algorithms, where the detector signal is interpreted and the reconstruction is done in multiple semileptonic or hadronic B meson decay channels. While the tagged analysis approach provides better understanding of the event kinematics, the reconstruction of the companion B has a limited efficiency on the order of less than 1%, thus significantly reducing the signal selection

efficiency. On the other hand, the untagged analyses are dominated by the $B \to X_c \ell \nu$ backgrounds, and the signal extraction is limited to a small region of phase space. The untagged analyses rely on the lepton momentum endpoint region for signal extraction. However, because the HQE does not converge properly in this part of the phase space, there are large uncertainties on the theoretical predictions needed for $|V_{ub}|$ determination. The below summarized $B \to X_u \ell \nu$ analyses rely on both of these analysis strategies.

The LHCb experiment [44] at the LHC proton-proton collider, currently operating at the CERN laboratory, also provided some relevant measurements for $|V_{ub}|$. While the direct measurement of $|V_{ub}|$ is not accessible to LHCb, the experiment provided multiple measurements relevant for the ratio $|V_{ub}/V_{cb}|$. One such measurement is the first observation of the $B_s^0 \to K^- \mu^+ \nu_\mu$ [45], where the branching fraction was measured normalized to the $B_s^0 \to D_s^- \mu^+ \nu_\mu$ decay. The Λ_b^0 baryons (udb quark content) are also accessible to the LHCb experiment. The $|V_{ub}/V_{cb}|$ ratio was determined using the ratio of the measured branching fractions of $\Lambda_b^0 \to p \mu^- \bar{\nu}_\mu$ and $\Lambda_b^0 \to \Lambda_c^0 \mu^- \bar{\nu}_\mu$ [46].

3.1 Inclusive $B \to X_u \ell \nu$ measurements and $|V_{ub}|$ estimations

The first experimental evidence of the charmless semileptonic B decays came from the CLEO experiment in 1990 [9]. The observation was made using $244 \times 10^3 \, \Upsilon(4S)$ events. The momentum spectrum of the final state muons and electrons was analyzed. An excess of events was seen in the region between 2.4 and 2.6 GeV, where the background from $B \to X_c \ell \nu$ is negligible. An excess of $70.4 \pm 20.3_{\text{stat}} \pm 10.4_{\text{syst}}$ events was observed. Shortly thereafter, the ARGUS experiment reported the analysis of the endpoint lepton spectrum

3.1 Inclusive $B \to X_u \ell \nu$ measurements and $|V_{ub}|$ estimations

[10]. In the region between 2.3 and 2.6 GeV, they observed an excess of $41 \pm 10_{\text{exp}}$ events above known backgrounds.

An updated result came from CLEO in 1993 [47], with an increased dataset consisting of $955 \times 10^3 \ \Upsilon(4S)$ events. In the lepton momentum region between 2.3 and 2.6 GeV, there was an observed excess of $107 \pm 15_{\text{stat}} \pm 11_{\text{syst}}$ events, which were attributed to $B \to X_u \ell \nu$. The measured partial branching fraction in this momentum region was found to be $\Delta \mathcal{B}(B \to X_u \ell \nu) = (154 \pm 22_{\text{stat}} \pm 20_{\text{syst}}) \times 10^{-6}$. The ratio $|V_{ub}/V_{cb}|$ was also measured at 0.076 ± 0.008.

A decade later, the most precise endpoint analysis was performed at BaBar[12], using 88 million $B\bar{B}$ events. Only the electron final state was considered. Relative contributions of the signal, $B\bar{B}$ (mainly $B \to X_c \ell \nu$) background and continuum background were estimated in a fit to data. A simulation based on the DFN model of the inclusive $B \to X_u e \nu$ momentum was used to approximate the signal contributions. The signal momentum distribution is shown in Fig. 3.1 after the background subtraction. To reduce the effect of the signal shape modelling, the data in the endpoint was combined into a single bin for the fitting, as indicated in Fig. 3.1. The obtained partial branching fraction in the momentum region 2.0-2.6 GeV is $\Delta \mathcal{B}(B \to X_u e \nu) = (0.572 \pm 0.041_{stat} \pm 0.065_{syst}) \times 10^{-3}$. The $|V_{ub}|$ value was determined to be $|V_{ub}| = (4.44 \pm 0.25_{exp} \pm \left(^{+0.42}_{-0.38} \right)_{SF} \pm 0.22_{theory}) \times 10^{-3}$, where the uncertainties come from the experiment, the shape function (SF) and the theoretical modelling. The extrapolated total branching fraction is $\mathcal{B}(B \to X_u e \nu) = 2.27 \pm 0.26_{exp} \pm \left(^{+0.33}_{-0.28} \right)_{SF} \pm 0.17_{theory}$.

The latest improvement of this measurement, completed in 2017, includes the full BaBar dataset of 466.5 million $B\bar{B}$ pairs [13]. The large dataset enabled the estimation of the partial branching fraction starting at the electron momentum of 0.8 GeV up to the kine-

3.1 Inclusive $B \to X_u \ell\nu$ measurements and $|V_{ub}|$ estimations
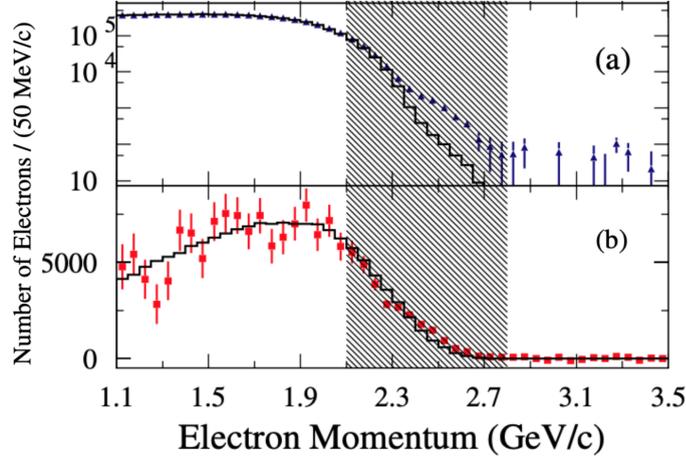
Figure 3.1: BaBar's results for electron-momentum spectra in the $\Upsilon(4S)$ rest frame: (a) on-resonance data after subtraction of the fitted non-BB background (triangles—blue) compared to simulated BB background that is adjusted by the combined fit to the on- and off-resonance data (histogram); (c) on-resonance data after subtraction of all backgrounds (linear vertical scale, data points—red), compared to the simulated $B \to X_u e\nu$ signal spectrum (histogram). The bins in the shaded area are combined into a single bin for the fit, in order to reduce the sensitivity to the shape of the signal spectrum in this region. Figure from [12].

matic endpoint. Again, the data between 2.1 and 2.7 GeV were combined into a single bin in order to reduce the fit sensitivity to the signal shape modelling. Four different models were used for the signal, DFN, BLNP, GGOU [34] and DGE [48]. The results show model dependence. The resulting partial branching fractions, the total branching fractions and the $|V_{ub}|$ values are listed in Table 3.1.

Another inclusive untagged analysis with a slightly different approach was performed at BaBar, as reported in [49]. Events with an electron with the momentum in the region $2.1 - 2.8$ GeV were selected. The missing 4-momentum in the event was calculated, defined as

$$p_{miss} = p_{e^+e^-} - p_{vis}, \qquad (3.1)$$

3.1 Inclusive $B \to X_u \ell \nu$ measurements and $|V_{ub}|$ estimations

$\Delta\mathcal{B}\,[10^{-3}]$	$\mathcal{B}\,[10^{-3}]$	$|V_{ub}|\,[10^{-3}]$
DFN		
$1.397 \pm 0.078_{\text{exp}}{}^{+0.214}_{-0.153\text{SF}}$	$1.494 \pm 0.084_{\text{exp}}{}^{+0.239}_{-0.167\text{SF}}{}^{+0.030}_{-0.003\,\text{theory}}$	$3.794 \pm 0.107_{\text{exp}}{}^{+0.292}_{-0.219\text{SF}}{}^{+0.078}_{-0.068\,\text{theory}}$
DGE		
$1.433 \pm 0.081_{\text{exp}}$	$1.537 \pm 0.086_{\text{exp}}{}^{+0.031}_{-0.003\,\text{theory}}$	$3.848 \pm 0.108_{\text{exp}}{}^{+0.084}_{-0.070\,\text{theory}}$
GGOU		
$1.554 \pm 0.082_{\text{exp}}{}^{+0.095}_{-0.086\text{SF}}$	$1.665 \pm 0.087_{\text{exp}}{}^{+0.103}_{-0.093\text{SF}}{}^{+0.002}_{-0.011\,\text{theory}}$	$3.959 \pm 0.104_{\text{exp}}{}^{+0.164}_{-0.154\text{SF}}{}^{+0.042}_{-0.079\,\text{theory}}$
BLNP		
$2.268 \pm 0.125_{\text{exp}}{}^{+0.191}_{-0.163\text{SF}}$	$2.418 \pm 0.134_{\text{exp}}{}^{+0.205}_{-0.176\text{SF}}{}^{+0.003}_{-0.003\,\text{theory}}$	$4.563 \pm 0.126_{\text{exp}}{}^{+0.230}_{-0.208\text{SF}}{}^{+0.162}_{-0.163\,\text{theory}}$

Table 3.1: Results from BaBar [13] for the partial branching fraction, the total branching fraction, $|V_{ub}|$ values based on the electron momentum range $\Delta p = 0.8 - 2.7$ GeV for different theoretical predictions, with experimental, SF, and theory uncertainties.

where $p_{e^+e^-}$ is the center of mass energy of the beams and p_{vis} is the total detected momentum in the detector. This was interpreted as the 4-momentum of the undetected neutrino, with certain corrections, $p_\nu\,(|\vec{p}_{miss}|, \vec{p}_{miss})$. Knowing the missing neutrino momentum enables the calculation of the momentum transfer, as $q^2 = (p_e + p_\nu)^2$. Additional selection requirements were chosen based on these variables to minimize the $B \to X_c \ell \nu$ background and the total theoretical and experimental uncertainties on $|V_{ub}|$. The measured partial branching fraction was determined to be $\Delta\mathcal{B}(2.0 < E_e < 3.5 \text{ GeV}) = (3.54 \pm 0.33_{stat} \pm 0.34_{syst}) \times 10^{-4}$. The obtained $|V_{ub}|$ is $|V_{ub}| = (3.95 \pm 0.26_{\text{exp}} \pm \left({}^{+0.58}_{-0.42}\right)_{\text{HQE}} \pm 0.25_{\text{theory}}) \times 10^{-3}$. This approach has lower signal efficiency and has additional systematic uncertainties from the estimation of the missing momentum.

The tagged analyses have access to the full set of kinematic variables (E_ℓ, m_X, q^2, etc.), but suffer from low signal yield, $\mathcal{O}(1000)$ events for the Belle and BaBar datasets. The tagged B is fully reconstructed, and an electron or a muon candidate is required, while the remaining detector signal is assigned to the X_u system. In this approach, up to 90% of the $B \to X_u \ell \nu$ decay rate is accessible. Additional requirements on the square of the missing mass in the event ($M^2_{miss} \approx 0$ if the neutrino is the only missing particle) are possible to

3.1 Inclusive $B \to X_u \ell \nu$ measurements and $|V_{ub}|$ estimations

implement in order to reduce the $B \to X_c \ell \nu$ background. However, the phase space regions where $B \to X_u \ell \nu$ signal is dominant are still the most sensitive, and the $B \to X_c \ell \nu$ subtraction still remains an important source of uncertainty.

The most recent tagged inclusive $B \to X_u \ell \nu$ partial branching fraction measurement came from Belle [50]. The full Belle dataset was used, 711 fb^{-1} of data recorded at the $\Upsilon(4S)$ resonance. The electron and muon final states were included. A machine learning algorithm was trained for $B \to X_c \ell \nu$ background suppression. Three variables, M_X, q^2, E_ℓ, were used for the partial branching fraction calculation, accessing from 31% up to 85% of the phase space. The most inclusive measurement, with the requirement $E_\ell > 1$ GeV, from a two-dimensional fit to the M_X and q^2 observables, gives the partial branching fraction $\Delta\mathcal{B}(B \to X_u \ell \nu) = (1.59 \pm 0.07_{\text{stat}} \pm 0.16_{\text{syst}}) \times 10^{-3}$. The calculated $|V_{ub}|$ value is $|V_{ub}| = (4.10 \pm 0.09_{\text{stat}} \pm 0.22_{\text{syst}} \pm 0.15_{\text{theory}}) \times 10^{-3}$.

Belle also reported the first measurement of the differential branching fractions[51], with the similar analysis setup as [50]. The partial branching fractions were reported as a function of the lepton energy, the 4-momentum transfer squared, the hadronic mass, the hadronic mass squared and the light-cone momenta. The results are shown in Fig. 3.2. These results can be used in the future for model independent determinations of $|V_{ub}|$.

The Particle Data Group (PDG) reports the averaged $|V_{ub}|$ value from inclusive $B \to X_u \ell \nu$ measurements, averaged over different signal models and assigned a modelling uncertainty [8]:

$$|V_{ub}| = (4.13 \pm 0.12_{\text{exp}}{}^{+0.13}_{-0.14}{}_{\text{theory}} \pm 0.18_{\Delta\text{model}}) \times 10^{-3} \quad \text{(inclusive)}. \qquad (3.2)$$

The Heavy Flavour Averaging Group (HFLAV) provides $|V_{ub}|$ values based on different
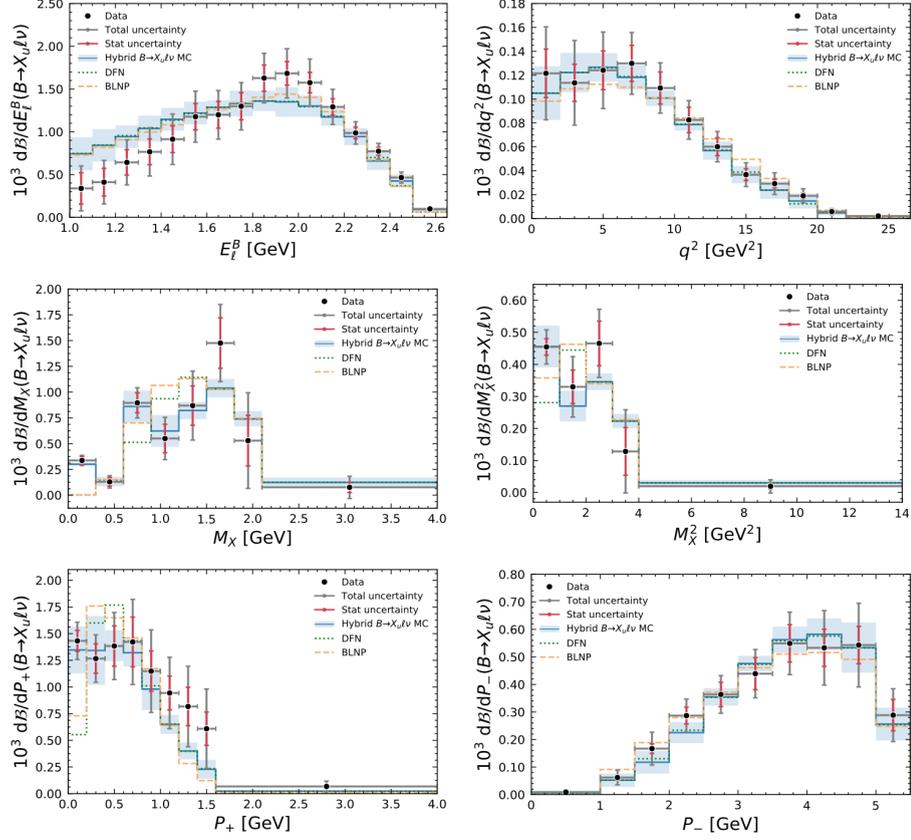
Figure 3.2: The results from Belle where the differential $B \to X_u \ell \nu$ branching fractions were measured: the lepton energy in the B rest frame (E_ℓ^B), the four-momentum transfer squared (q^2), the invariant hadronic mass and mass squared of the X_u system (M_X, M_X^2), and the light-cone momenta of the hadronic X_u system ($P_\pm$). The MC prediction and the DFN and BLNP inclusive calculations are also shown. Figure from [51].

$B \to X_u \ell \nu$ models. The relevant result for this thesis is the one based on the BLNP modelling. It is shown in Fig. 3.3. The results using the DFN model were not presented.

3.2 Exclusive $B \rightarrow X_u \ell \nu$ measurements and $|V_{ub}|$ estimations

The exclusive final state charmless semileptonic decay searches can be performed tagged and untagged. The decay $\bar{B} \rightarrow \pi \ell \bar{\nu}$ is the most reliable decay mode for $|V_{ub}|$ determination. The untagged measurements use the missing momentum of the event to estimate the kinematics of the undetected neutrino. While the signal to background ratio in these analyses is low (S/B<1), they have high statistical power. In the tagged measurements, the signal to background ratio is high (S/B~10) and the q^2 resolution is better, but the statistical power is limited.

CLEO [53, 54], BaBar [55] and Belle [56] performed untagged analyses of $\bar{B} \rightarrow \pi \ell \bar{\nu}$ and $\bar{B} \rightarrow \rho \ell \bar{\nu}$. The dominant systematic uncertainties in these analyses come from the neutrino reconstruction and the modelling of the $B \rightarrow X_u \ell \nu$ background.

The untagged measurements can be performed using a semileptonic tag, where the companion B meson is reconstructed in the $B \rightarrow \bar{D}^{(*)} \ell^+ \nu_\ell$ decay mode, or using a hadronic tag, where only the fully hadronic final states are considered. The semileptonic tag has higher selection efficiency than the hadronic tag, but it has additional missing neutrinos. Both the semileptonic and hadronic tag measurements were performed at Belle [57, 58] and BaBar [59, 60]. The dominant systematic uncertainties in these analyses come from the calibration of the tagging algorithm. The $|V_{ub}|$ value can be obtained from the $\bar{B} \rightarrow \pi \ell \bar{\nu}$ measurements using the measured branching fraction and the q^2 spectrum. The HFLAV group provides the average measured $|V_{ub}|$ value from these measurements [52]:

$$|V_{ub}| = (3.70 \pm 0.10_{\text{exp}} \pm 0.12_{\text{theory}}) \times 10^{-3} \quad \text{(exclusive)}. \tag{3.3}$$

3.2 Exclusive $B \to X_u \ell \nu$ measurements and $|V_{ub}|$ estimations

The inclusive and exclusive $|V_{ub}|$ measurements are marginally consistent. This tension between the inclusive and exclusive measurements has been a long-standing puzzle in B physics. The results from inclusive measurements are performed over different regions of phase space and using different theoretical models, but the results are consistent. The exclusive measurements mainly focus on the $\bar{B} \to \pi \ell \bar{\nu}$ decay mode, where the theoretical predictions are better known. To resolve the inclusive-exclusive tension, advancements are necessary, both in the theoretical modelling of the shape function at higher orders and in the experimental techniques.
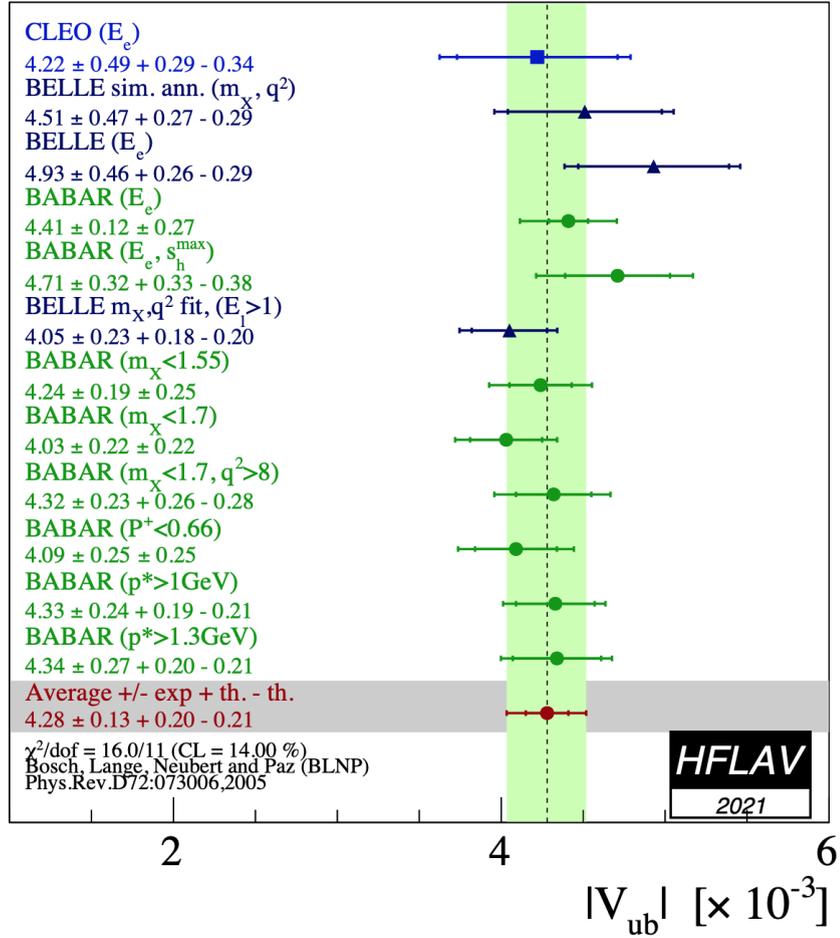
Figure 3.3: Measurements of $|V_{ub}|$ from inclusive semileptonic decays and their average based on the BLNP prescription. The labels indicate the variables and selections used to define the signal regions in the different analyses. The quoted uncertainties are statistical and systematic experimental uncertainties. The bars show the total uncertainty with the statistical uncertainty also indicated on the bar. Figure from [52].

4

The Belle II Experiment

The Belle II experiment is the successor of the Belle and BaBar B-factory experiments. It is located at the SuperKEKB e^+e^- collider, in the KEK laboratory in Tsukuba, Japan. The Belle II experiment aims to collect 50 ab^{-1} of e^+e^- collision data, which is 50 times more than Belle and BaBar collected each. The data will be used to search for physics beyond Standard Model, by looking for rare and forbidden processes, as well as to perform precision measurements of the Standard Model predictions. After two phases of collider and detector commissioning, the experiment started collecting data in March 2019.

4.1 The SuperKEKB accelerator

The SuperKEKB collider is an asymmetric electron-positron collider with beam energies of 7 GeV and 4 GeV, respectively. These beam energies give a total center of mass energy of 10.58 GeV, which is just above the mass of the $b\bar{b}$ quark resonance, $\Upsilon(4S)$. The $\Upsilon(4S)$ decays almost exclusively ($> 96\%$ [8]) to pairs of B and anti-B mesons, either $B^0\bar{B}^0$ or

B^+B^-. The entangled B meson pairs allow for a variety of measurements that test the flavour sector of the Standard Model, including the CKM quark mixing model and measurements of CP violation. The asymmetric beam energies provide a boost to the center of mass system of $\beta\gamma \approx 0.284$. By producing the boosted B meson pairs, it is possible to measure the location of the decay vertex of each B meson. This allows Belle II to perform measurements of time dependent CP violation in the B meson system.

A schematic rendering of the SuperKEKB accelerator is shown in Fig. 4.1. It uses the existing 3 km long circular tunnel structure for beam pipes from the KEKB accelerator. Electrons and positrons are first accelerated in a linear accelerator before being injected into the SuperKEKB accelerator storage rings: the Low Energy Ring (LER) for 4 GeV positrons, and the High Energy Ring (HER) for 7 GeV electrons. The frequency of e^+e^- collisions is 509 MHz. The Belle II detector is positioned at the intersection point of the two storage rings. The accelerator has been upgraded to achieve a record high instantaneous luminosity of 6×10^{35} cm^{-2}s^{-1}, which is 40 times higher than the KEKB luminosity. The instantaneous luminosity is defined as:

$$\mathcal{L} = \frac{1}{\sigma}\frac{dN}{dt}, \tag{4.1}$$

where σ is the cross section of the given interaction and dN is the number of events detected in a certain period of time (dt). This relates to the total integrated luminosity, describing the total amount of the collected data, as:

$$\mathcal{L}_{int} = \int \mathcal{L}dt. \tag{4.2}$$

4.1 The SuperKEKB accelerator

The increase in luminosity is achieved by using higher beam currents and by implementing a 'nano-beam' scheme. The higher beam currents mean that the number of electrons and positrons in the beams is larger than at KEKB and other previous electron-positron colliders. The beam size is also compressed to achieve the so-called nano-beams, which increases the rate of collisions. The downside of this luminosity increase is an increase in the beam-induced backgrounds. Consequently, the Belle II detector is exposed to much higher radiation during its operation. The Belle II detector design needed to achieve high performance in this high-rate environment.
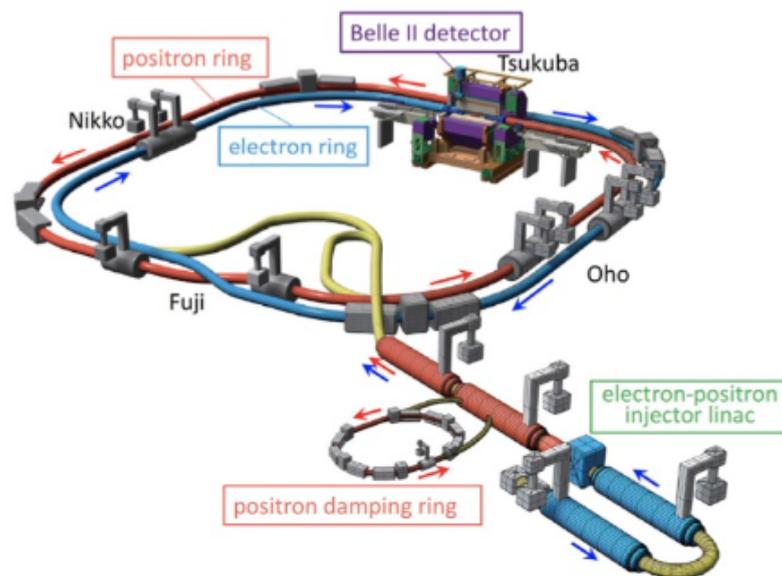
Figure 4.1: Artistic rendering of the SuperKEKB accelerator, with the Belle II detector located at the interaction point. The electron and positron storage rings are indicated, as well as the injection linac. Picture was taken from [61].

4.2   The Belle II detector

The Belle II detector is located at the interaction point of the SuperKEKB collider. It is used to identify and reconstruct decays initiated in e^+e^- collisions. The detector is located in a 1.5 T magnetic field provided by a solenoid. The detector comprises multiple sub-detectors, each with a specific purpose, such as particle identification, energy measurement, and charged particle tracking. The function, operating principles and layout of each sub-detector system are described in the following subsections. These are based on references [62] and [63], unless otherwise indicated. A schematic of the Belle II detector is shown in Fig. 4.2.
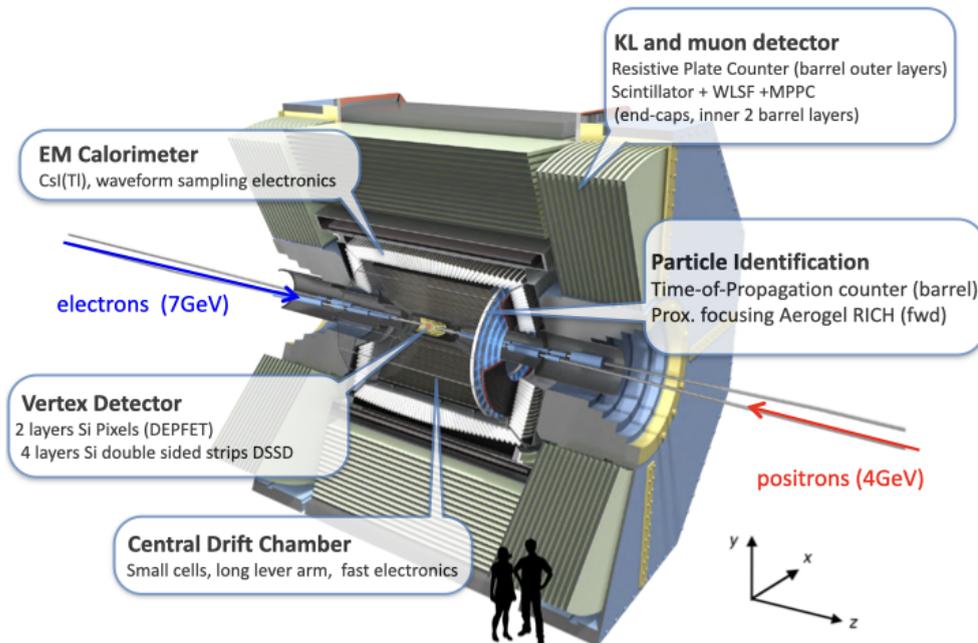
Figure 4.2: Artist rendering of the Belle II detector. The subdetectors are labeled, as well as the coordinate system used for the experiment. Two people are shown for size comparison. Picture modified from [64].

4.2.1   Pixel vertex detector (PXD)

The Pixel Vertex Detector (PXD) together with the Silicon Vertex Detector (SVD) form the Vertex Detector (VXD) of the Belle II detector. The VXD is used to reconstruct the decay vertices of fast-decaying particles, primarily the B mesons. This is based on the measurements of the position of charged particles that pass through the PXD and the SVD. The PXD is located closest to the interaction point, only 4 mm away from the beam pipe and 14 mm away from the interaction point.

The PXD uses silicon based DEPFET (DEpleated P-channel Field-Effect Transistor) technology. Electron-hole pairs are created in the depleted silicon by a charged particle passing through. Due to the applied voltage bias, a current is created by the electrons drifting towards the internal gate. This current activates the transistor switch, signalling that a particle has passed through and the readout commences. After the readout, the accumulated charge is cleared by a clear contact placed on the periphery of each pixel.

The DEPFET technology was chosen due to its thinness (75 μm in Belle II) and the internal amplification, which was needed to comply with the low material budget required. This was necessary in order to avoid multiple scatterings of charged particles, which would degrade the vertex resolution. Due to the increased luminosity of the SuperKEKB collider, there was also a need for a radiation-hard device that can withstand the high level of background, which DEPFET does.

The PXD consists of two cylindrical layers, the first at $r_1 = 14$ mm, and the second at $r_2 = 22$ mm. The small inner radius of the PXD was necessary to compensate for the smaller beam energy asymmetry compared to the previous B-factory experiments. This

results in smaller separation of B and $\bar{B}$ vertices. The inner (outer) PXD layer consists of 8 (12) ladders, each with a width of 15 mm and a sensitive length of 90 (123) mm. The lengths of the ladders differ between the two layers to offer the coverage over the full acceptance region, between $17°$ and $150°$ polar angle. The sensors are pixelated to avoid occupancy (fraction of detector channels triggered in an event) issues due to the high luminosity and high background environment. The PXD consists of 10^7 readout channels. It is operated in the continuous readout mode with readout time of 20 μs per frame. The readout electronics are situated outside the acceptance region, thus significantly improving (reducing) the material budget. The detector configuration is shown in Fig. 4.3.

The expected vertex resolution of the full PXD is 10 μm. Initially, the inner layer was completely installed, but only two ladders in the outer layer were installed due to production delays. The full outer layer was installed during the long shutdown of the Belle II detector in the summer of 2023. The resolution of the partially installed portion of the PXD was measured at 14 μm [66] and is in agreement with the expected performance based on simulation.

4.2.2 Silicon vertex detector (SVD)

The Silicon Vertex Detector (SVD), together with the PXD, forms the Belle II VXD. Figure 4.4 is an image taken during the assembly of the VXD detectors. Its purpose is to reconstruct the decay vertices of the two B mesons in each event, as well as the decay vertices of D mesons and τ leptons. SVD, together with PXD, also measures the tracks of low-momentum charged particles (order of 10 MeV) that do not reach the Central Drift Chamber. SVD also measures the decay vertex of K_S^0 mesons that decay outside the PXD.
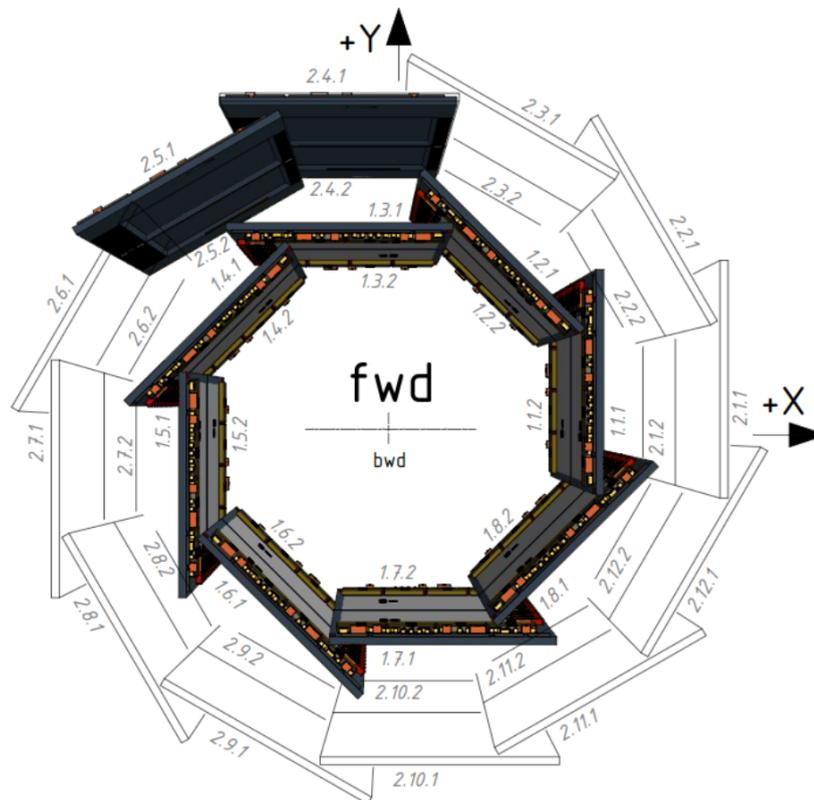
Figure 4.3: Schematic of the PXD modules' cross-section in beam direction. The data used in this thesis was recorded using only the coloured modules. The remaining modules were installed in summer 2023. Image taken from [65].

It also provides data necessary to extrapolate the charged-particle tracks from the Central Drift chamber to the PXD.

The SVD consists of four layers at radii $r_1 = 38$ mm, $r_2 = 80$ mm, $r_3 = 115$ mm, and $r_4 = 140$ mm. The layers are cylindrical, around the PXD and the beam pipe, with slanted sections in the forward portion of the three outer layers. Each layer consists of slightly overlapping ladders with double-sided sensors. A schematic of the SVD is shown in Fig. 4.5.
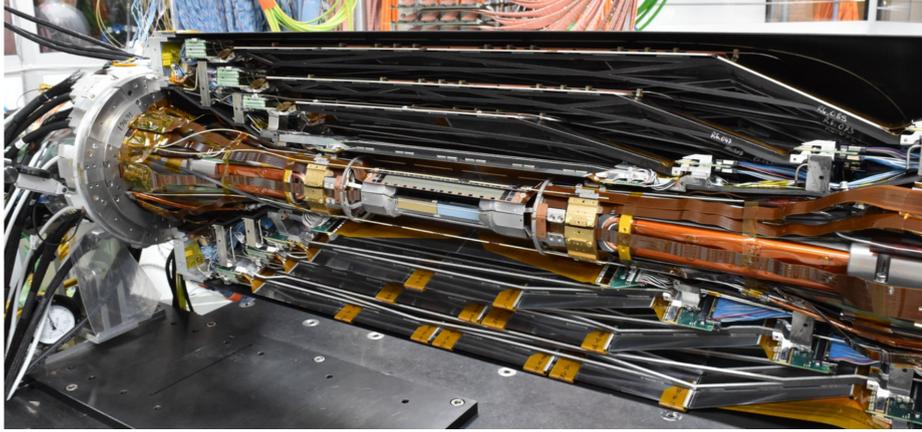
Figure 4.4: Image taken during the assembly of the VXD detectors. The PXD detector is visible in the middle. The slanted SVD sensors in the forward region (right) are visible. Image taken from [66].

The SVD uses double-sided silicon strip detectors (DSSD). Strips are used instead of pixels to reduce the number of readout channels. There is no occupancy issue such as with the PXD because the the background hit rate decreases with $1/r^2$. The strips are arranged perpendicularly on the two sides of the sensor. On the inner side, closer to the interaction region, the p-side of the strips are along the z-axis. On the outer side, the n-side of the strips are along the $r - \phi$ direction. As a charged particle passes through a sensor, it liberates a shower of electrons and holes along its path. The electrons drift towards the n-side of the strips, while the holes drift towards the p-side. By reading out the signal from each side, we can infer the 2D position of the charged particle. This is illustrated in Fig. 4.6.

The SVD covers the polar angle region between $17°$ and $150°$, the same as the PXD. The material budget is at 0.568% X_0 per layer, where X_0 is the radiation length. The expected spatial resolution based on Monte Carlo simulation of 0.5 GeV single muons is $2 - 5$ μm in $r - \phi$ and $7 - 27$ μm in z direction, depending on the layer and the incident angle.
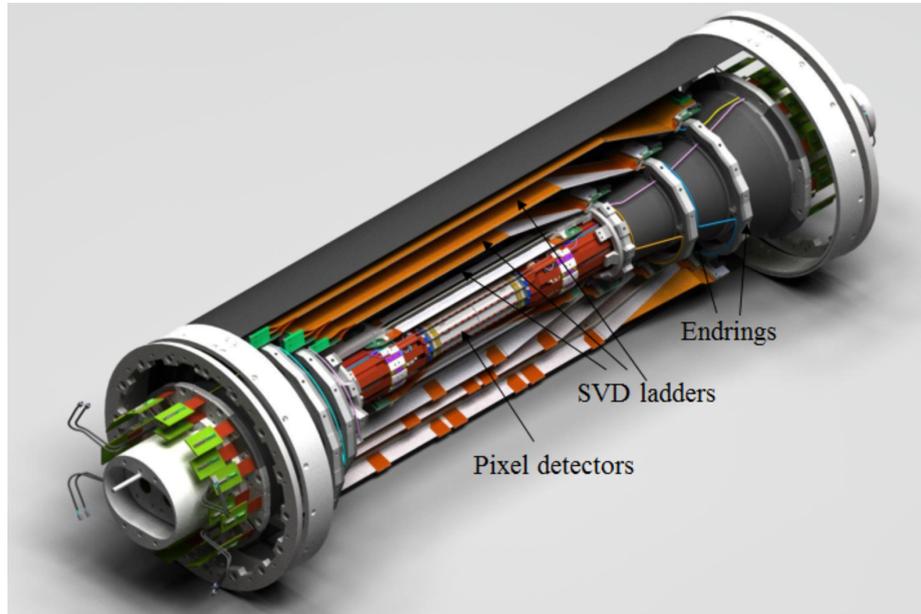
Figure 4.5: 3D model of the Belle II VXD [67]. The SVD ladders are indicated.

The measured resolution in data collected in 2019 is not as good, an observation attributed to the simulation being too optimistic, and the reconstruction algorithms being tuned on MC [68]. The agreement between data and MC is expected to improve by improving the simulation and the reconstruction resolution on data.

4.2.3   Central Drift Chamber (CDC)

The Central Drift Chamber (CDC) is the central tracking device of the Belle II detector. The CDC reconstructs the trajectories of charged particles and measures their momenta precisely. It is also used for particle identification, by measuring the energy loss rate, dE/dx, which is especially useful for low momentum particles that do not leave the CDC volume. The CDC is also used to provide trigger signals for charged particles.
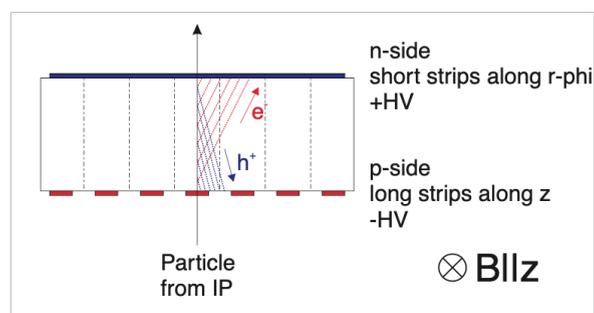
Figure 4.6: Illustration [62] of the passage of a charged particle through a DSSD sensor. The direction of the magnetic field is indicated (parallel to the z-axis).

The CDC is a large volume drift chamber with small drift cells. It has a cylindrical shape, with the inner radius of 160 mm and the outer radius of 1130 mm. The length of CDC is 2.3 m. There are 14336 sense wires, with either axial or stereo orientations. Axial orientation sense wires are parallel to the solenoidal magnetic field. Stereo orientation is skewed with respect to the axial orientation by between 45 and 74 mrad. Sense wires are made out of gold-plated tungsten and are operated at a high voltage of approximately 2.3 kV. There are 42240 aluminium field wires which are grounded. Figure 4.7 shows how the wires are organized in layers and superlayers. The innermost superlayer has smaller cells to reduce the occupancy due to high beam backgrounds.

The CDC volume is filled with a 50:50 gas mixture of helium and ethane. This gas is chosen since it performed well in the Belle detector, with low radiation length, good position and energy loss resolutions, and little radiation damage.

As a charged particle traverses the drift chamber, it ionizes the gas molecules. Because of the applied electric field, the electrons drift towards the closest sense wire, ionizing more gas molecules along the way, thus creating an avalanche. The measurements of the position and timing of the signal created by the avalanche in the sense wire enables us to measure
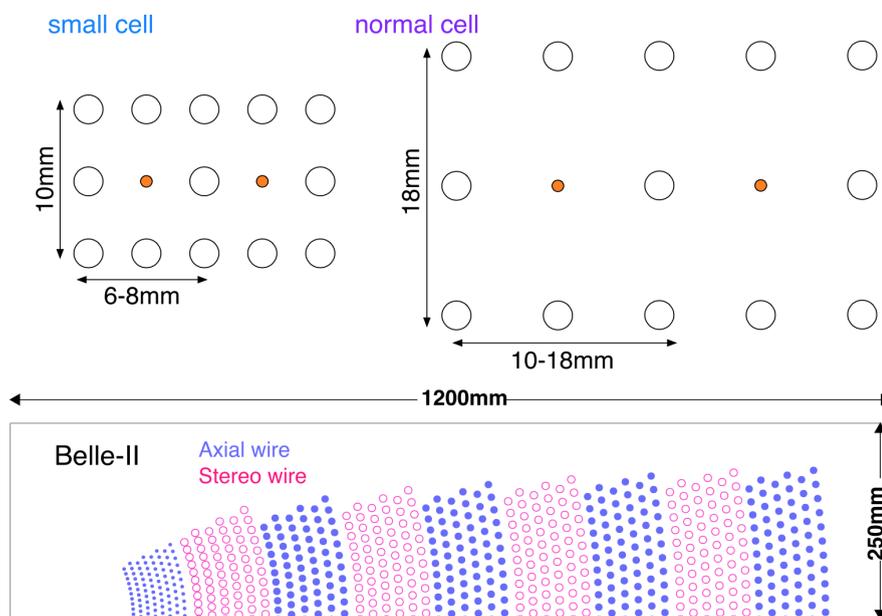
Figure 4.7: The upper portion of the figure shows the wire configuration in a small cell (inner superlayer) and in a normal cell. The sense wires are shown with orange dots, while the white circles show the grounded field wires. The lower portion of the figure shows the layer configuration of the Belle II CDC. Image taken from [69].

the position of the charged particle. The size of the avalanche is proportional to the energy lost by the particle in the initial ionization event. By combining the signal from all the ionization events a particle produces, we can obtain the dE/dx. Since the energy loss of a particle at a given momentum is dependent on its mass, based on the Bethe formula, the particle type can be identified. The dE/dx distribution for different particles is shown in Fig. 4.8.

The CDC has the same polar angle coverage as the VXD, between $17°$ and $150°$. The position resolution is about $100\ \mu$m. The dE/dx resolution is dependent on the incident angle and it is between 8% and 12%.
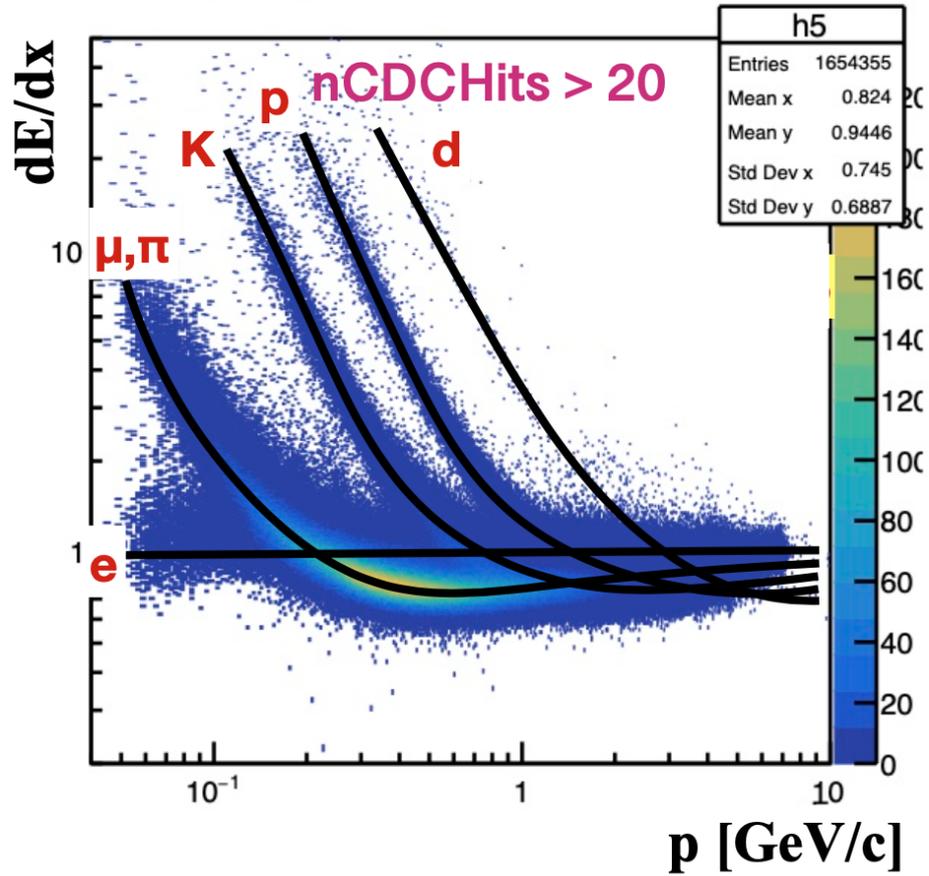
Figure 4.8: Energy loss rates (dE/dx) of charged particles in the Belle II CDC. Solid lines show predictions from the Bethe formula, dots show Belle II data recorded in 2019. Figure from [70].

4.2.4 Time-of-propagation counter (TOP)

Good K/π separation is necessary for Belle II to perform competitive B meson reconstruction. The time of propagation counter (TOP) is the sub-detector that provides the particle identification in the barrel region of the Belle II detector. The position of the TOP counter, in front of the calorimeter, puts additional constraints on its design by requiring the material

budget to be low. The TOP counter is the only operational Cherenkov time of propagation counter of its kind. It distinguishes between different particles based on the Cherenkov radiation they emit when traversing the TOP subdetector medium.

The TOP counter consists of 16 modules arranged in a barrel shape, around the CDC sub-
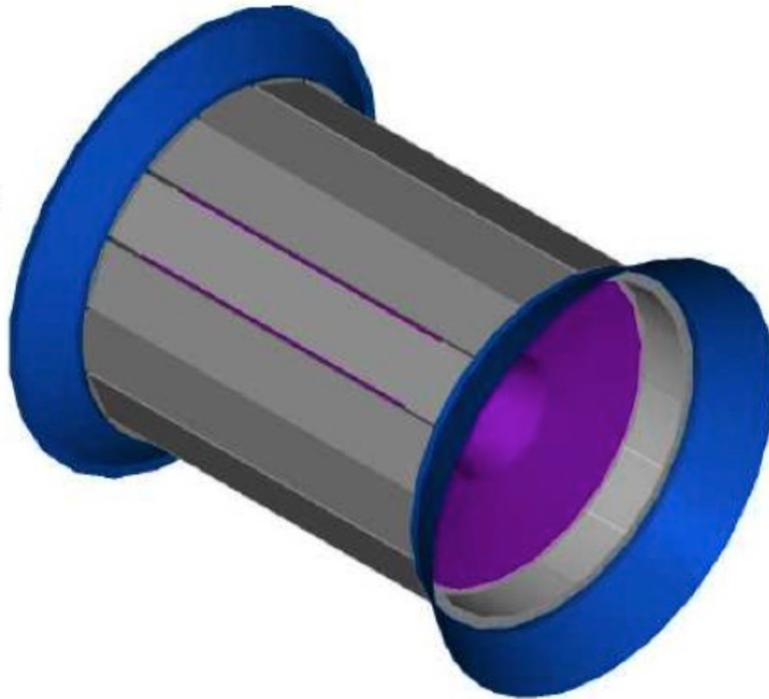
Figure 4.9: 3D conceptual rendering of the TOP detector (grey) integrated together with the CDC (purple). Figure from [62].

detector, as shown in Fig. 4.9. Each detector module consists of a 250 cm long, 45 cm wide and 2 cm thick quartz radiator bar with a mirror on one end. On the other end, the radiator is glued to a 10 cm long prism that couples to a readout array of micro-channel-plate photomultiplier tubes (MCP PMTs). There are 32 MCP PMTs connected to each quartz bar, each with 16 readout channels. The total number of readout channels for the TOP detector

is 8192. The azimuthal coverage of the TOP counter is 93%. The schematic of one TOP module is shown in Fig. 4.10.

 As a charged particle traverses the quartz radiator, it emits Cherenkov radiation. The an-
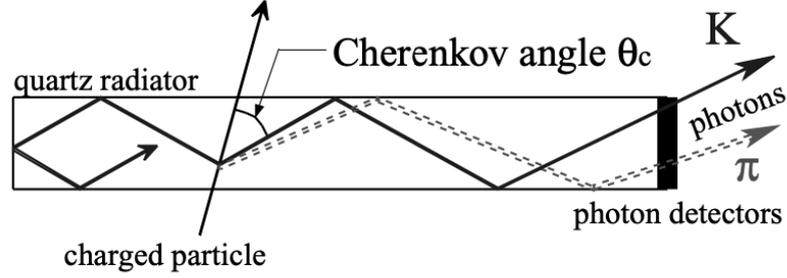
Figure 4.10: Side view of the inner reflection of the Cherenkov light inside the TOP counter. Figure from [71].

gle of the Cherenkov photons with respect to the particle trajectory, θ_C, depends on the particle's velocity:

$$\cos\theta_c = \frac{1}{\beta n}, \tag{4.3}$$

where n is the refractive index of the medium and β is the particle's velocity as a fraction of the speed of light [8]. The photons travel inside the quartz bar and are internally reflected until they reach an MCP PMT. The photon time of arrival is a sum of the time of flight of the charged particle to the quartz radiator and the time of propagation of the Cherenkov photon in the radiator. Based on the time of arrival and the $x - y$ detection position, it is possible to distinguish between different particles. The position and timing information is compared to the expected Probability Density Functions (PDFs) for six particle hypotheses (e, μ, π, K, p, d). The identification probabilities are assigned based on the six corresponding likelihood values.

The performance is tested using the MC simulations. The results show that a kaon identifi-

cation efficiency in the momentum region between 0.5 and 2 GeV is 90% with a pion fake rate below 5%.

4.2.5   Aerogel ring-imaging Cherenkov detector (ARICH)

To complement the particle identification of the TOP counter in the barrel region, an aerogel ring-imaging Cherenkov (ARICH) detector is used in the forward endcap. The ARICH subdetector provides discrimination between pions and kaons from the momentum of 0.4 GeV up to 4 GeV. It also provides pion-electron separation up to 1 GeV. ARICH consists of two 2 cm thick aerogel layers with different refractive indices ($n = 1.045$ upstream and $n = 1.055$ downstream). The detector screen of 540 hybrid avalanche photon detectors (HAPDs) is used for the photon readout and it is separated from the aerogel radiators by a 20 cm expansion gap. To obtain the needed Cherenkov angle resolution, enough photons need to be produced in the radiator and detected. This can be achieved with a thicker radiator, which in turn increases the uncertainty of the point of emission. The double layer of aerogel radiator with carefully chosen refractive indices provides focusing of the Cherenkov photons and reduces the emission point uncertainty. Aerogel material is used for this purpose because its refractive index can be chosen during production. The principle of operation of the double radiator is illustrated in Fig. 4.11. The HAPD sensor provides single photon sensitivity with high granularity. The Cherenkov angle resolution of ARICH is 3.1 mrad.

4.2.6   Electromagnetic calorimeter (ECL)

The Belle II electromagnetic calorimeter (ECL) is a scintillation calorimeter that is primarily used for detecting the energy and position of photons from neutral particle decays (e.g.
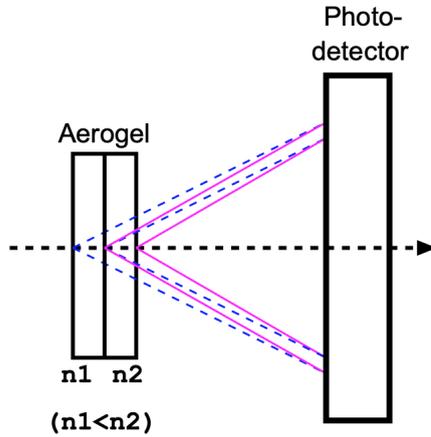
Figure 4.11: Principle of ARICH detection: Cherenkov photons (dashed blue and solid purple lines), emitted at different angles from the two radiators, are detected by the photo-detector. Figure from [72].

$\pi^0 \to \gamma\gamma$). Almost a third of decay products of a B meson are π^0's, so it is important to properly reconstruct them. For that reason, the ECL needs to provide high efficiency photon detection in a wide range of energies, between 20 MeV and 4 GeV. The ECL is also used for electron and, together with the KLM (Section 4.2.7), muon and K_L^0 identification. The calorimeter is also part of the trigger system and it is used for measuring the on-line and off-line luminosities.

When a high-energy photon traverses the calorimeter medium, it goes through pair production, $\gamma \to e^+e^-$. High-energy electrons, produced in pair production or in other processes, lose energy by Bremsstrahlung radiation, $e^- \to e^-\gamma$. These processes continue in a chain as long as the decay products have sufficient energy. This chain process is called an electromagnetic shower. Once the produced electrons and photons no longer have enough energy for Bremsstrahlung or pair production, they continue losing energy by ionisation and excitation of the calorimeter medium. Ideally, all the energy of the original photon or electron is absorbed in the calorimeter. As the energy is absorbed by the scintillating material of the

calorimeter, it re-emits the energy in the form of light. The light yield from the scintillator is proportional to the absorbed energy. Thus, by measuring the light output, the energy of the original photon/electron can be calculated.

The ECL is located outside the TOP and ARICH detectors. The position of the ECL is illustrated in Fig. 4.12. It consists of 8736 CsI(Tl) scintillating crystals arranged in a cylindrical shape with a total weight of 43 tons. The forward and backward endcaps are located at $z = 1.96$ m and at $z = -1.02$ m, and contain 2112 crystals in total. The 3 m long barrel region has an inner radius of 1.25 m and it contains 6624 crystals. The crystals are 30 cm long (16.1 radiation lengths) truncated pyramids with an average cross section of 6×6 cm^2. The same crystals and their support structure were used in the Belle experiment. Thallium doped caesium-iodide crystals are chosen because of their good energy resolution of less than 5%, high light output of 54000 photons per MeV of deposited energy and short radiation length of 1.86 cm.

There are two photodiodes with amplifiers that are used for readout on the back end of each crystal. The readout system was upgraded compared to Belle in order to combat the problem of pile-up present in the high-luminosity environment of Belle II. Faster readout is needed to avoid pile-up, where signals from one event overlap with the signal from the subsequent event because the time between the events is much shorter compared to the Belle experiment. This issue is lessened by the use of faster readout electronics that incorporate waveform sampling. Waveform sampling, where the shape of potential signals is fitted using timing information, is used to ensure only signals that are in time with e^+e^- collision are kept.

The ECL covers 90% solid angle in the center of mass frame and provides polar coverage between $12.4°$ and $155.1°$. For photons of energies 100 MeV and 8 GeV, expected energy

resolution is 4% and 1.6% and angular resolution is 13 mrad and 3 mrad [62], respectively.
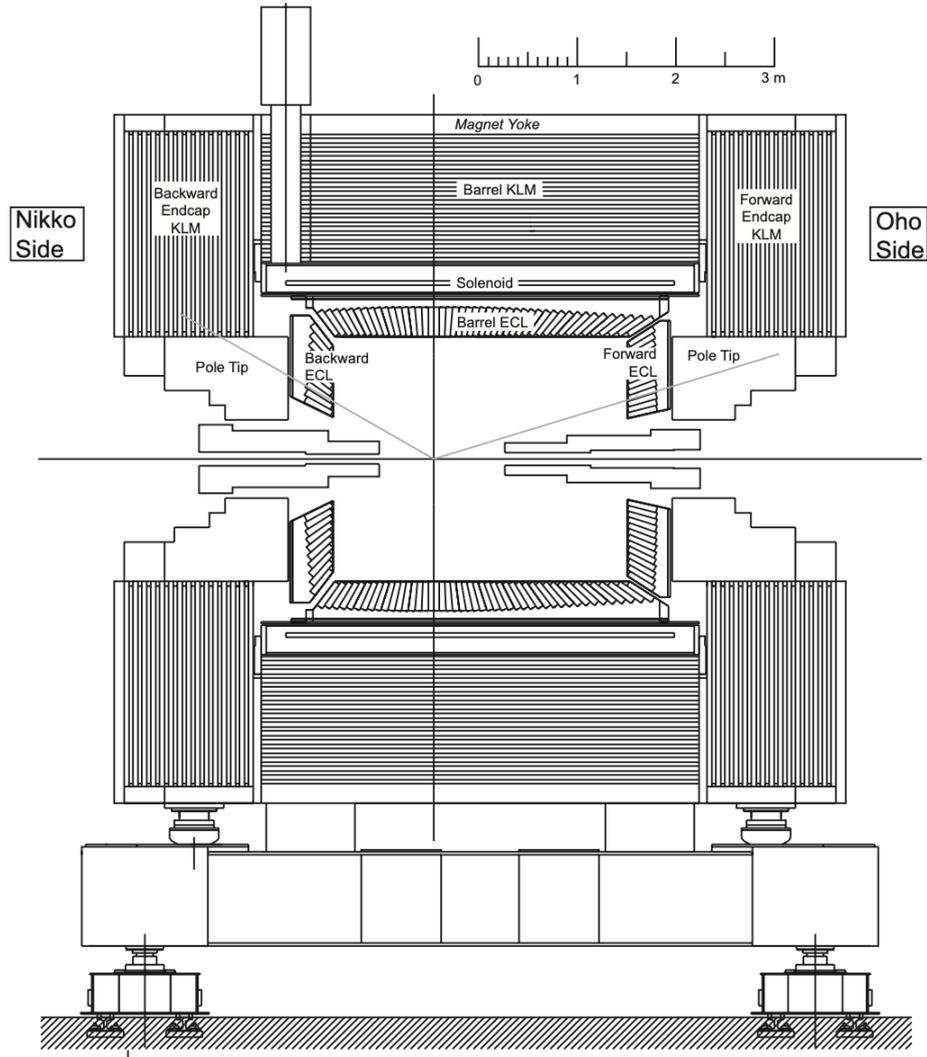
Figure 4.12: Schematic of the side view of the Belle II detector. The ECL crystals are indicated. The KLM is visible, located outside the ECL and solenoid. The gray lines mark the nominal polar angular acceptance of Belle II. Figure from [62].

4.2.7 K-Long and muon detector (KLM)

The K-Long and muon detector (KLM) is used to provide K_L^0 and muon reconstruction. It consists of alternating 4.7 cm thick iron plates and active detector layers, located outside the Belle II solenoid. The iron plates in the KLM also act as the magnetic flux return for the solenoid. The iron plates provide 3.9 interaction lengths of material, beyond the 0.8 interaction lengths of the ECL. The schematic of the KLM is shown in Fig. 4.12.

The KLM consists of the octagonal barrel region and two endcap sections. The barrel region has 14 detection layers. All but the two innermost layers use glass-electrode resistive plate chambers (RPC). These RPCs are inherited from the KLM of the Belle detector where they showed good performance. Each detector layer contains two RPC layers that are orthogonal to each other in order to provide z and ϕ positional information. An RPC is a large planar capacitor with the gap between the electrodes filled with gas. As a charged particle travels through the gas, it causes ionisation. The ionisation is further amplified by the electric field generated by the electrodes, turning into an 'avalanche'. When the avalanche comes into contact with the electrodes, signal is created. Neutral hadrons, such as K_L^0, produce hadronic showers as they interact with iron nuclei. Secondary charged particles are created in these showers, which then can be detected.

An RPC works on the principle that it discharges when a charged particle passes through and it must recharge before the next detection. This dead time during recharge poses a problem when used in a high luminosity experiment. To mitigate this problem, the endcaps and the two inner layers in the barrel use scintillation detectors. Here, in each gap between the iron flux return plates, strips of scintillating polystyrene are placed. In each detection layer the strips are arranged in two orthogonal planes to provide z and ϕ positional infor-

mation. The scintillating strips are covered in reflective coating and they contain an optical fibre running though the middle. The optical fibre is mirrored on one end and on the other it is attached to a photodiode used for readout. The fibre also shifts the wavelength of the blue scintillation light to a green wavelength in which photodiodes have greater efficiency. A diagram of the principle of operation of the scintillation strips is shown in Fig. 4.13.
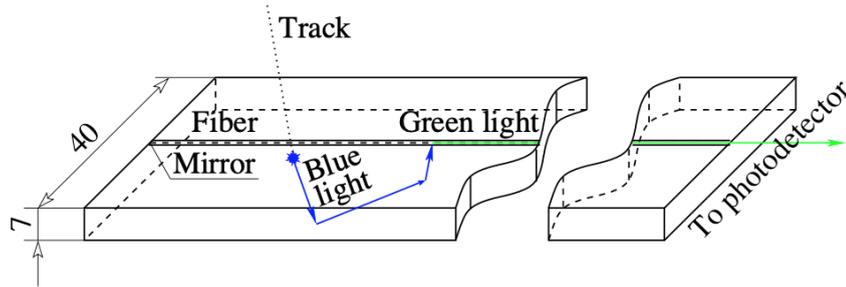
Figure 4.13: Detection of scintillation light in a scintillating strip of the KLM detector. Figure from [62].

The KLM provides polar angle coverage of $25° - 155°$ and close to 49000 readout channels. The muon detection efficiency is 89% above 1 GeV, with a hadron fake rate of 1.3%. The K_L^0 detection efficiency rises linearly with momentum from zero at 0 GeV to a plateau of 80% at 3 GeV [62].

4.2.8 Triggering and data acquisition

Since operating at a high instantaneous luminosity, there is a need for event selection during the data taking in Belle II. A trigger is used to select the events of interest and filter out the unwanted background. Information from the CDC, ECL, TOP and ARICH is combined and forwarded to a final-decision logic. Total cross sections and trigger rates at the design

4.2 The Belle II detector

Physics process	Cross section (nb)	Rate (Hz)
$\Upsilon(4\,\mathrm{S}) \to B\bar{B}$	1.2	960
Hadron production from continuum	2.8	2200
$\mu^+\mu^-$	0.8	640
$\tau^+\tau^-$	0.8	640
Bhabha ($\theta_{\mathrm{lab}} \geqslant 17°$)	44	$350^{(a)}$
$\gamma\gamma$ ($\theta_{\mathrm{lab}} \geqslant 17°$)	2.4	$19^{(a)}$
2γ processes ($\theta_{\mathrm{lab}} \geqslant 17°, p_t \geqslant 0.1\,\mathrm{GeV}/c$)	~ 80	~ 15000
Total	~ 130	~ 20000

$^{(a)}$ rate is pre-scaled by a factor of $1/100$

Table 4.1: Total cross section and trigger rates [62] with $L = 8 \times 10^{35}$ cm^{-2} s^{-1} from various physics processes at $\sqrt{s} = 10.58$ GeV.

luminosity[1] of 8×10^{35} cm^{-2} s^{-1} for the processes of interest are given in Table 4.1. Here, the hadron production from continuum refers to the events where an $\Upsilon(4S)$ meson is not produced, but a quark pair is produced. At the Belle II center-of-mass energy, these are the $e^+e^- \to u\bar{u}, d\bar{d}, s\bar{s}, c\bar{c}$ processes. The Bhabha scattering refers to $e^+e^- \to e^+e^-$ processes. While the Bhabha and $\gamma\gamma$ events are easy to distinguish based on their detector signature, a fraction of them is saved to be used for measuring the luminosity and calibrating the detector responses.

The trigger consists of two levels, the online Level 1 trigger (L1) and the offline High Level Trigger (HLT). The trigger is a part of the Data Acquisition system (DAQ). The aim of the Belle II DAQ is to record the events of interest and to perform multi-step data reduction. This is done using a large number of CPUs, ranging from the on-board detector readout to the HLT computing farm. The L1 trigger performs online event selection based on the track trigger and the energy trigger. The track trigger consists of 2D and 3D tracking

---

[1]The initial aim of the SuperKEKB collider and the Belle II experiment was to achieve the instantaneous luminosity of 8×10^{35} cm^{-2} s^{-1}, but this goal has since been revised to 6×10^{35} cm^{-2} s^{-1}. For that reason, some of the estimations in the detector and DAQ design were performed using the 8×10^{35} cm^{-2} s^{-1} instantaneous luminosity.

algorithms, while the energy trigger computes the total event energy and looks for isolated clusters, all done with a precise trigger timing [73]. The expected data rate coming from the L1 trigger at the goal luminosity is up to 30 kHz, with 1.1 MB per event, or up to 30 GB/s. Such a vast amount of data needs to be further reduced to a manageable level, which is done offline, using the HLT. Full scale offline reconstruction is performed and only the physics level events are selected. This process further reduces the data rate down to a few kHz. The HLT classifies the selected events into different categories: hadronic, Bhabha, $\mu\bar{\mu}$, $\tau\bar{\tau}$, $\gamma\gamma$, cosmic, other scaled calibration events including random triggered events, and other background events [74]. The analysis described in this thesis uses events from the hadronic category. This trigger is highly efficient in selecting the $B\bar{B}$ events, with efficiency of higher than 99.5%, estimated using Monte Carlo simulation.

5

Analysis strategy and tools

5.1 Analysis strategy

The goal of this analysis is to measure the partial branching fraction of the decay $B \rightarrow X_u \ell \nu$ in the lepton momentum endpoint region, between 2.1 and 2.7 GeV in the center-of-mass frame[1] of the $\Upsilon(4S)$. The analysis is done inclusively, where the X_u hadron is not reconstructed, but only the outgoing lepton is selected, either an electron or a muon. The analysis is performed on the data from the Belle II experiment. This analysis builds on the analyses performed at the previous B-factory experiments, as described in Chapter 3, especially the BaBar measurements [12, 13]. Uniquely, unlike in the previous inclusive measurements performed at BaBar and Belle, this analysis is performed simultaneously on the electron and muon final states.

The analysis tools and dataset used are described in this chapter. The following chapter describes the signal selection. A binned template fit is performed on the selected data to

---

[1]Further on, the use of the $\Upsilon(4S)$ center-of-mass reference frame will be implied, unless otherwise noted.

extract the signal yield in the lepton momentum endpoint. The Monte Carlo simulated data samples are used as templates for signal and backgrounds and are fitted to the measured Belle II data. The fit is performed on the lepton momentum, between 1.0 and 3.0 GeV, separated into bins of 50 MeV. The endpoint region, between 2.1 and 2.7 GeV is combined into a single wide bin, as was done by BaBar [13]. The signal yield in the endpoint is obtained from the fit result in the wide bin. The fit strategy and results are described in more detail in Chapter 7. The systematic uncertainties of the extracted endpoint signal yield is estimated for various sources, as described in Chapter 8. The partial branching fraction in the lepton endpoint is calculated based on the measured signal yield, as described in Chapter 9. Further on, the value of the V_{ub} CKM matrix element is calculated, based on the measured partial branching fraction. The results presented in this thesis are blinded, as described in the next section. No central values are shown; only the relative statistical and systematic uncertainties of the measured quantities are presented.

5.1.1   Blinded analysis strategy

The analysis is performed blinded, such that the value of the endpoint signal yield is hidden from the analyst. This is done to avoid any experimental bias. The signal yield in the full fitted momentum range, between 1.0 and 3.0 GeV, is available to the analyst, but the signal yield in the endpoint, between 2.1 and 2.7 GeV, is hidden. The hidden value is further propagated through the analysis, and it is used to estimate the relative systematic uncertainties, but the absolute systematic uncertainties are also blinded.

Before the final results are ready to be unblinded, various checks are performed to validate the analysis methodology. A cut flow analysis of the signal selection requirements was done to verify the selection efficiency for each of the requirements, as described in Sec-

tion 6.5. Following the signal selection, the distribution of the lepton momentum in data is compared to the Monte Carlo simulation prior to fitting, as shown in Section 6.6. This step is necessary in order to verify that the Monte Carlo simulation provides a good approximation of the data. In the next step, the template fit was performed. To demonstrate that the fit algorithm performs correctly, Asimov data was used. A fit to the Asimov data is a fit where the established fitting algorithm is performed, but in place of the recorded data, a Monte Carlo dataset is used. If the fitting algorithm works correctly, such a fit is expected to produce the results identical to the Monte Carlo simulation. The fitting algorithm checks with Asimov data are described in Section 7.2. Finally, the whole analysis procedure is validated by measuring the $B \to X\ell\nu$ branching fraction, as described in Chapter 9. The branching fraction of the flavour-agnostic $B \to X\ell\nu$ process is well known and the agreement of the measured branching fraction with the previous results demonstrates the robustness of the analysis algorithm designed by the author.

In this thesis, the results are presented blinded - only the estimated measurement uncertainty is shown. The unblinding will be performed at a later date, leading to the journal publication. This is done to prevent any results that are not approved by the Belle II collaboration entering public circulation. The final decision on the unblinding readiness is given by the Belle II collaboration, following a multi-stage review process. After an analysis is complete, for it to receive the unblinding permission, the Working Group and the assigned Review Committee need to give their approval. The path to unblinding for this analysis is further discussed in Chapter 9. Now, the dataset and the tools used for this analysis are described in the following sections.

5.2   Belle II dataset

The Belle II experiment has collected 424 fb^{-1} of e^+e^- collision data up through July 2022. The experiment has recently ended the Long Shutdown One period, during which the pixel detectors and the collider magnets were upgraded. Operations resumed in February 2024. This was the first part of the preparation for achieving the full targeted instantaneous luminosity of 6×10^{35} cm^{-2}s^{-1}. The maximal instantaneous luminosity achieved before the shutdown was 4.7×10^{34} cm^{-2}s^{-1}, which is higher than any other e^+e^- collider has achieved before. The accumulation of the total integrated luminosity for the run period between 2019 and present is shown in Fig. 5.1.

 Due to the need for central processing of the recorded data and the time consuming process of data verification and systematic error estimation, this thesis only uses the data recorded until December 2021. This amounts to a total integrated luminosity of 189 fb^{-1} recorded at the center of mass energy corresponding to the $\Upsilon(4S)$ resonance, at $\sqrt{s} = 10.58$ GeV, and another 18 fb^{-1} recorded 60 MeV below the resonance. The data recorded below the $\Upsilon(4S)$ resonance is called "off-resonance" data and it is used to study non-resonant contributions from the e^+e^- collisions, where the $\Upsilon(4S)$ resonance is not produced, but various backgrounds of type $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$, e^+e^-, $\mu^+\mu^-$ or $\tau^+\tau^-$ are. These backgrounds are called "continuum" backgrounds. The dataset used in this thesis is summarised in Table 5.1. The data used is skimmed using the HLT hadron skim to reduce the input data size, as described in Chapter 4.
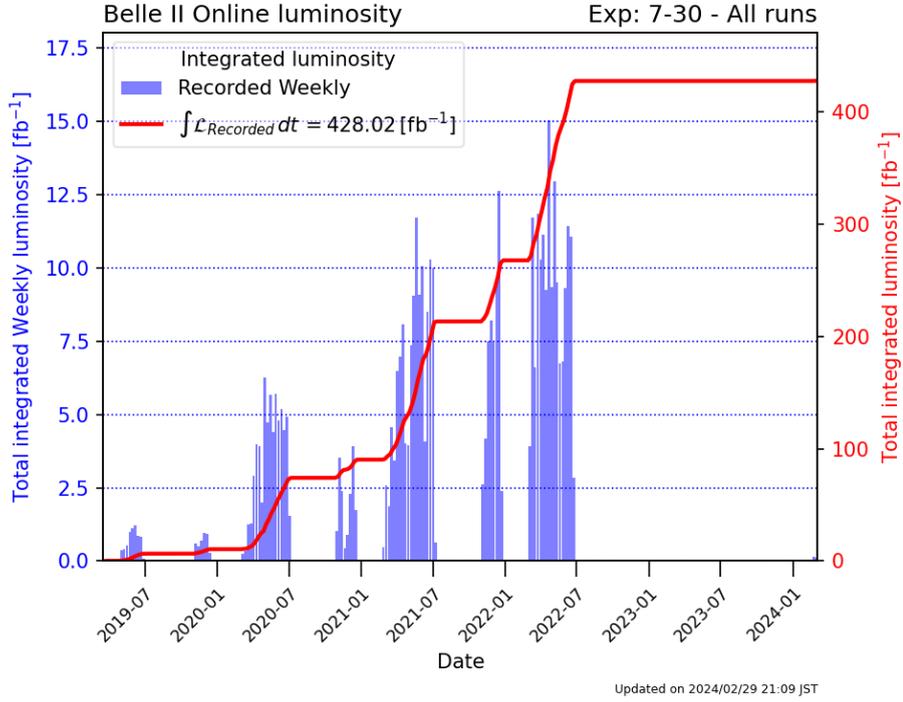
Figure 5.1: The total weekly integrated luminosity and the total cumulative integrated luminosity of the Belle II experiment since the start of its operations in 2019 [75].

5.3 Monte Carlo Simulation

Simulated Monte Carlo datasets (MC) are used to compare the recorded data to the theoretical predictions, as well as to optimize the signal selection and estimate the signal selection efficiency and background rate. Using the MC generators, a simulated counterpart of the Belle II data is created. Event generator packages are used to produce possible e^+e^- interactions based on theoretical branching fraction predictions using MC generators. The EvtGen [76] package is used to produce $B\bar{B}$ samples, while the continuum events are produced using the KKMC [77] generator for the $\mu^+\mu^-$, $\tau^+\tau^-$ or $q\bar{q}$ processes; BBREM [78], BHWIDE [79] and BABAYAGA.NLO [80] are used for e^+e^- processes. The final-state

$\sqrt{s}$ [GeV]	Experiment number	Luminosity [fb^{-1}]
10.58	7	0.510 ± 0.002
	8	4.459 ± 0.003
	10	3.635 ± 0.001
	12	54.388 ± 0.004
	14	16.385 ± 0.005
	16	10.321 ± 0.004
	17	10.714 ± 0.004
	18	89.176 ± 0.010
		total: 189.880 $\pm$ 0.013
10.52	8	0.813 ± 0.001
	12	8.716 ± 0.002
	18	8.424 ± 0.003
		total: 17.953 $\pm$ 0.004

Table 5.1: Summary of the collected Belle II data that is used in this analysis. The experiment number corresponds to a period of data collection with the same accelerator parameters. The presented uncertainties are statistical uncertainties only and are provided by the Belle II Data Production group.

radiation is simulated using the PHOTOS [81] package. The event list with decay products and the information on their kinematics is then forwarded to the Geant4 [82] simulation of the Belle II detector, which simulates the detector response. Because the information of the processes produced in an event, the truth information, is known in MC, it enables us to optimize selection criteria for the analysis of a desired signal mode, as well as to study the background processes that mimic the signal.

This analysis uses 100 fb^{-1} of MC simulated data. This includes B^+B^-, $B^0\bar{B}^0$, $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$ and $\tau^+\tau^-$ samples. Together, these MC samples are called 'generic' MC. The MC samples were provided by the Belle II Collaboration and are from the 14th production campaign of MC simulation (MC14). The continuum background samples were used initially for the optimization of the signal selection, but later in the analysis the off-resonance samples were used instead, as motivated in Section 6.2. A dedicated MC sample of the signal

Process $(e^+e^- \rightarrow)$	Cross-section [nb]	Number of events per 100 fb^{-1} ($\times 10^6$)
$\Upsilon(4S) \rightarrow B^+B^-$(charged)	0.540	54
$\Upsilon(4S) \rightarrow B^0\bar{B}^0$(neutral)	0.510	51
$u\bar{u}$	1.605	160.5
$d\bar{d}$	0.401	40.1
$s\bar{s}$	0.383	38.3
$c\bar{c}$	1.329	132.9
$\tau^+\tau^-$	0.919	91.9
Total	**6.835**	**683.5**

Table 5.2: Summary of the branching fractions and the total number of events generated for the MC used in this analysis, for $\sqrt{s} = 10.58$ GeV.

$B \rightarrow X_u \ell \nu$ events is used, with an improved modelling, as described in section 5.3.1. The MC samples used are summarized in Table 5.2.

5.3.1 Signal $B \rightarrow X_u \ell \nu$ MC with hybrid modelling

The previous B-factory experiments used the hybrid model, originating at CLEO [83], to describe the $B \rightarrow X_u \ell \nu$ decays. The inclusive $B \rightarrow X_u \ell \nu$ decay reconstruction includes both resonant and non-resonant final states. The triple-differential decay rate for inclusive $B \rightarrow X_u \ell \nu$ decays, $\frac{d^3\Gamma}{dq^2 dE_\ell dm_X}$, does not account for the resonances in the mass spectrum of the hadron. In MC, this can be ameliorated by integrating out the resonant contributions from the inclusive decay simulation and combining with the separate resonant decay simulations. To do this, the inclusive distribution is split into bins in hadron invariant mass, m_X,

the momentum transfer, q^2, and the lepton energy in the B rest frame, E_ℓ^B:

$$m_X = [0.00, 1.40, 1.60, 1.80, 2.00, 2.50, 3.00, 3.50] \text{ GeV}, \qquad (5.1)$$

$$q^2 = [0.00, 2.50, 5.00, 7.50, 10.0, 12.5, 15.0, 20.0, 25.0] \text{ GeV}^2, \qquad (5.2)$$

$$E_\ell^B = [0.00, 0.50, 1.00, 1.25, 1.50, 1.75, 2.00, 2.25, 3.00] \text{ GeV}. \qquad (5.3)$$

The resonant contributions are subtracted from the inclusive predictions in each bin, giving the rate of only the non-resonant final states. The resonant contribution is then added based on the exclusive predictions. Thus, the total number of events that the hybrid model predicts is:

$$H_i = R_i + \omega_i I_i, \qquad (5.4)$$

where H_i is the total number of events in bin i, R_i is the number of resonant contributions, I_i is the number of events from the inclusive distribution, and the weight $\omega_i = \frac{I_i - R_i}{I_i}$ is assigned such that in each bin $H_i = I_i$.

The hybrid model was implemented using a module available in the eFFORT package [84]. We used signal $B \to X_u \ell \nu$ MC from the MC14 campaign. The branching fractions were updated to the most recent PDG values. The results for the partial branching fraction and $|V_{ub}|$ will be presented for both BLNP and DFN signal models.

The samples are described in Table 5.3. The distribution of m_X before and after the hybrid model implementation is shown in Fig. 5.2. As can be seen in this figure, the inclusive MC, shown by the blue dotted line, does not include the resonant peaks in m_X. These resonances are simulated separately, shown by the green dash-dotted line. When the hybrid reweighting is performed, the rate of the resonant contributions is subtracted from the inclusive contribution. This brings down the inclusive contributions, as shown with the orange dash-

dotted line. The total signal contribution, consisting of the resonant and newly reweighted inclusive contributions, is shown with the solid black line. The resulting discontinuity in the m_X spectrum is an expected side effect of the hybrid model implementation. Since the m_X spectrum was divided into bins to calculate the new hybrid weights, the discontinuity lies on the edge between the first and the second bin. This region is most severely affected because all the resonances lie in the first m_X bin. While there is some concern that this new m_X spectrum is unphysical, it is not a worry in this analysis, since the m_X quantity is not directly used.

This signal MC with hybrid weights assigned is used to replace the signal events from the $B\bar{B}$ generic MC. The difference between the lepton momentum distribution between these two signal samples is shown in Fig. 5.3. It is evident that the new signal MC has higher number of events, purely because the branching fractions were updated when the hybrid MC was produced.

	B^+	B^0	Model
$\pi\ell\nu$	$(0.78 \pm 0.027) \cdot 10^{-4}$	$(1.50 \pm 0.06) \cdot 10^{-4}$	BCL
$\rho\ell\nu$	$(1.58 \pm 0.11) \cdot 10^{-4}$	$(2.94 \pm 0.21) \cdot 10^{-4}$	BCL
$\eta\ell\nu$	$(0.39 \pm 0.05) \cdot 10^{-4}$	-	ISGW2
$\eta'\ell\nu$	$(0.23 \pm 0.08) \cdot 10^{-4}$	-	ISGW2
$\omega\ell\nu$	$(1.19 \pm 0.09) \cdot 10^{-4}$	-	BCL
$X_u\ell\nu$ incl.	$(2.21 \pm 0.32) \cdot 10^{-3}$	$(2.05 \pm 0.21) \cdot 10^{-3}$	BLNP / DFN

Table 5.3: Branching fractions and models used to simulate resonant and inclusive $B \to X_u\ell\nu$ MC samples. The BCL[85] and ISGW2[86] models are used to describe the exclusive charmless decays. The previously discussed DFN [32] and BLNP [35] models are used to describe the inclusive charmless decays.
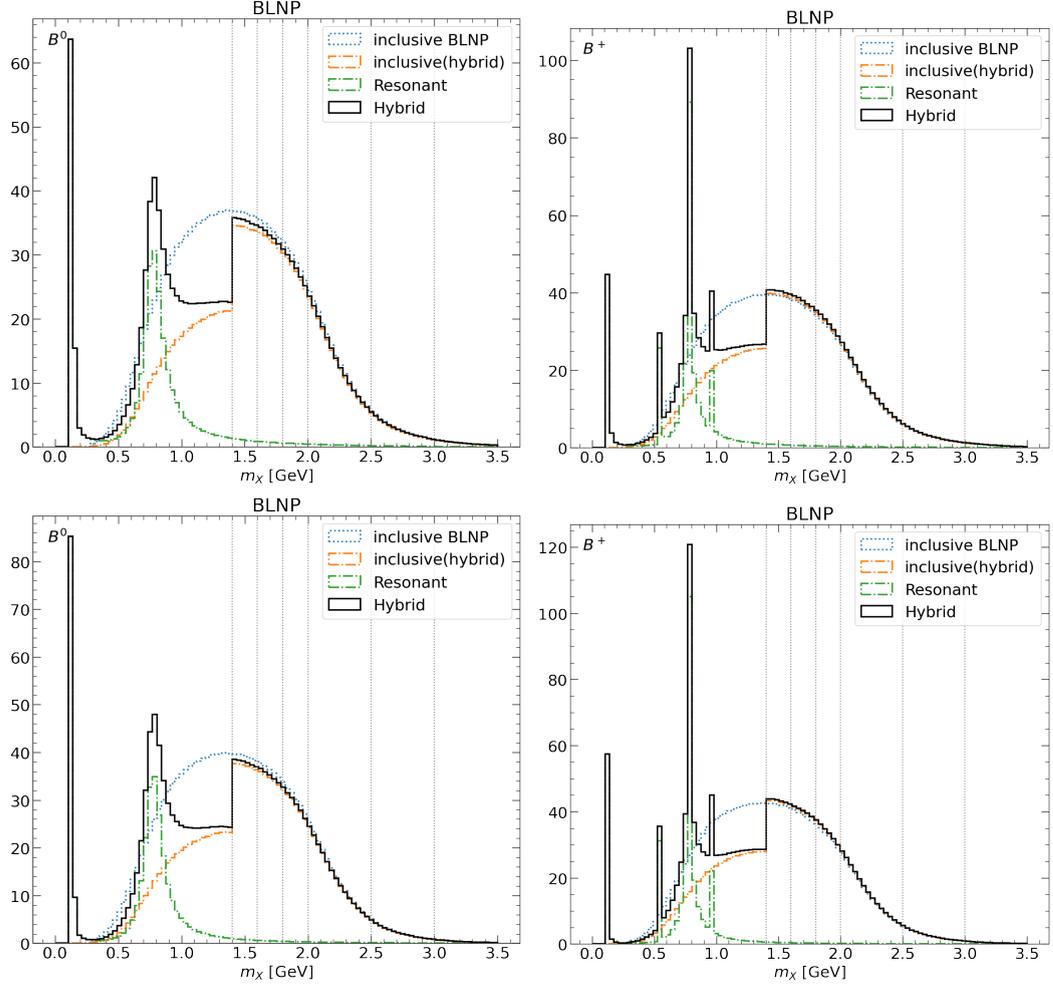
Figure 5.2: Distribution of the hadron invariant mass for inclusive MC (blue dotted line), resonant MC (green dash-dotted line), the inclusive contribution after the hybrid model applied (orange dash-dotted line), and the hybrid resonant and non-resonant contributions combined (black solid line), for B^0 decays (left) and for B^+ decays (right), for electrons (top) and muons (bottom), scaled to luminosity of 1 fb^{-1}. The vertical grey dotted lines show the bins in m_X used for hybrid reweighting.
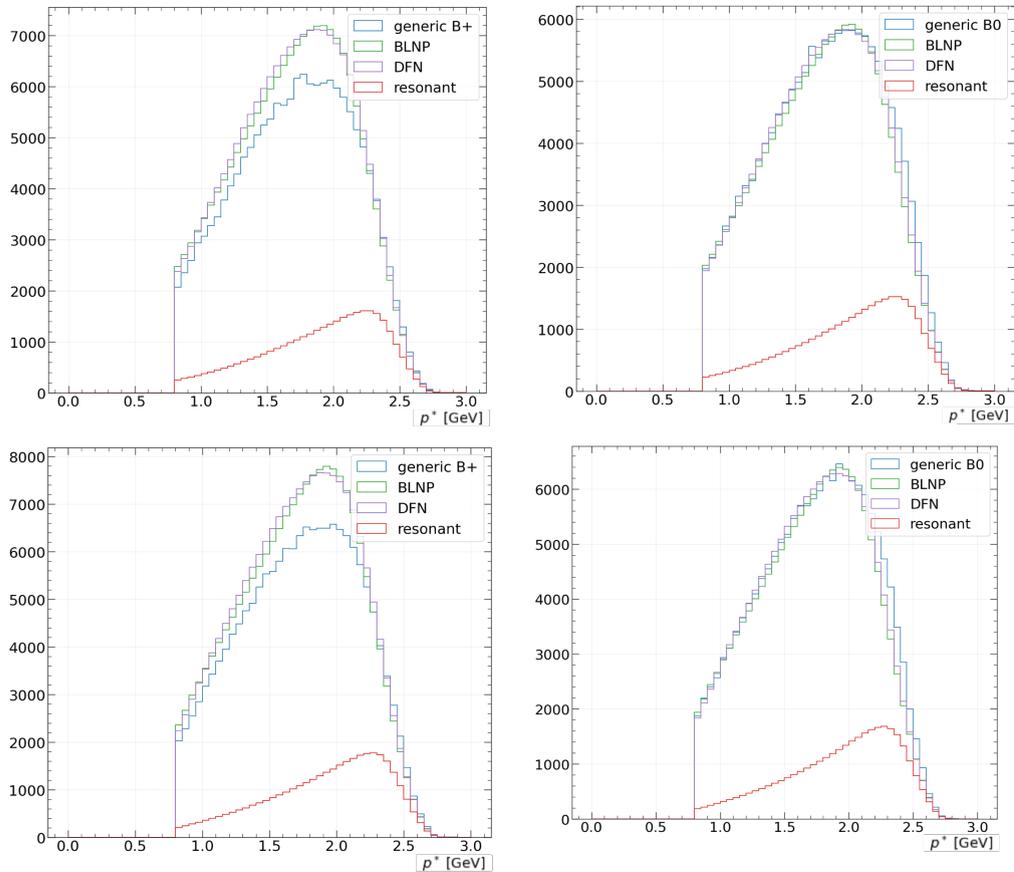
Figure 5.3: Signal electron (top) and muon (bottom) spectrum in the center of mass frame in the generic $B\bar{B}$ MC (blue) and in the signal MC with hybrid weighting (resonant modes (red) and two different models for non-resonant decays, BLNP (green) and DFN (purple)), for B^+ decays (left) and for B^0 decays (right).

5.4   Belle II Analysis Software Framework

The Belle II experiment uses its dedicated Belle II Analysis Software Framework (`basf2`) [87] for online and offline data processing, detector simulation and generation of MC samples. The framework consists of independent processing blocks, called modules, which are specialized to perform small tasks. The modules are combined and executed linearly within a path defined in use-specific steering files. The modules exchange information through a common object store, called DataStore. This is illustrated in Fig. 5.4. The input/output is handled using CERN's ROOT library [88].

Once the data is recorded, it is centrally processed and the detector calibration is per-
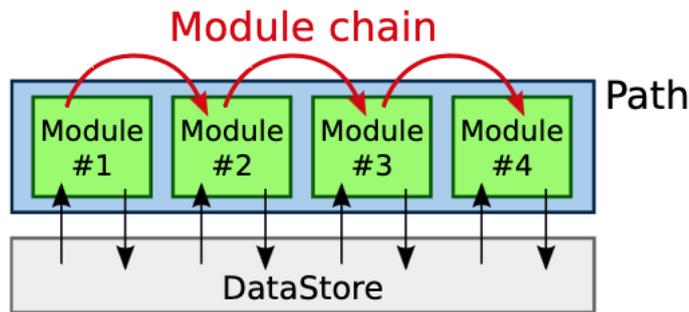
Figure 5.4: A simple data processing chain. One path containing 4 modules that exchange data with each other using the common DataStore. Image taken from [89].

formed. Similarly for MC, the simulation is centrally run using the generators described previously. Both of these operations are done using `basf2`, and ROOT files containing data objects are made available to analysts. These files contain detector objects, such as `Tracks`, `ECLClusters` and `KLMCLusters`. For a specific analysis, the analyst further processes these ROOT files using `basf2`. At this stage, the detector objects are interpreted as particles and are gathered in `ParticleLists`. The analyst applies selection criteria

and has the option to further combine particles to reconstruct decays. The results are again output as ROOT files containing particle candidates organized into ROOT's `TTree` objects. Further analysis and data manipulation is performed outside `basf2` ('offline'), most commonly using Jupyter Notebooks [90].

The data and MC samples used in this analysis are processed using the release 5 of `basf2` software. The analysis selection is performed using `basf2` release `light-2205-abys`. The next chapter describes the signal reconstruction done using `basf2`, as well as the further signal selection and the reweighting of the data and MC done offline.

6

Signal Selection

6.1 Signal selection requirements

The signal for the inclusive $B \rightarrow X_u \ell \nu$ analysis consists solely of the outgoing charged lepton. In this analysis approach, the final-state hadron is not reconstructed, while the lepton neutrino escapes the detection.

Because the B mesons are always produced in pairs at the Belle II experiment, there are different analysis approaches based on the treatment of the companion B meson. One of the B mesons is, at least partially, reconstructed as the signal, B_{sig}. In this analysis, B_{sig} is partially reconstructed with the outgoing charged lepton. The other B meson in the event is referred to as B_{tag}. If the B_{tag} is not reconstructed, the analysis is labelled as untagged. Otherwise, the B_{tag} can be reconstructed using hadronic or semileptonic tagging algorithms. While the tagged approach provides more information on the companion B and offers better understanding of the event kinematics, the tradeoff is a loss of signal selection efficiency. Due to the suppressed nature of the $B \rightarrow X_u \ell \nu$ signal decay, this analysis uses

6.1 Signal selection requirements

the untagged approach, to maximize the signal selection efficiency.

To optimize the ratio of signal to background events, a set of requirements is imposed on the data. These selection requirements are colloquially referred to as "cuts". Since the signal consists of either an electron or a muon, these particles need to be correctly identified in data. These charged particles leave tracks as they traverse the CDC, thus a requirement for a track is the first step in their identification. Since any charged particle can leave a track in the CDC, additional information from other subdetectors needs to be combined to assign a likelihood for an electron or a muon. Information from other particle identification detectors (TOP, ARICH, ECL, KLM) is used to calculate likelihoods for the track signal to be caused by different particle types. A global PID likelihood for a given particle type can be calculated using the combined likelihood information from the subdetectors:

$$\frac{\mathcal{L}_{\text{particle}}}{\mathcal{L}_e + \mathcal{L}_\mu + \mathcal{L}_\pi + \mathcal{L}_K + \mathcal{L}_d}. \tag{6.1}$$

The selection cuts for electron and muon signal candidates used in this analysis are as follows:

- at least one charged track is required, with the impact parameters dr <1.0 cm and $|dz|$ <3.0 cm, to ensure that the track is originating from the interaction region; here dr is the distance in the $r - \phi$ plane and dz is the distance on the z-axis from the interaction point.

- the track is required to have the transverse momentum $p_t > 0.05$ GeV and to be in the CDC angular acceptance; this is to ensure that the track is properly measured by the CDC subdetector.

6.1 Signal selection requirements

- the track is required to have high global PID likelihood to be either an electron or a muon; `electronID` > 0.9 or `muonID` > 0.9 is required.

- only the tracks with the center of mass momentum, p^*, between 1.0 GeV and 3.2 GeV are saved for further analysis; the tracks with $p^* < 1.0$ GeV are not used because the particle identification does not have a good performance in that momentum region, resulting in high fake rate, where other particles are misidentified as electrons or muons; the tracks with $p^* > 3.2$ GeV come exclusively from continuum background events.

Figure 6.1 shows the distributions of these variables in signal and in background MC.

In addition to the track selection requirements, a set of requirements is imposed on the event level. The goal is to minimize contributions from the machine and continuum backgrounds. A requirement on the total number of tracks is imposed, such that at least 5 tracks are required in an event, `nTracks` > 4. This reduces the low-multiplicity backgrounds without significantly affecting the signal efficiency. The ECL clusters that pass the requirements $E > 0.1$ GeV and $|$`clusterTiming`$| < 200$ ns and that are in the CDC angular acceptance, are defined as `good clusters`. An event needs to have more than 2 `good clusters` to pass the selection. Similarly, a requirement on the number of `good tracks` in an event is selected to be greater than 3. A `good track` is defined as a track with $p_t > 0.05$ GeV, $dr < 1.0$ cm, $|dz| < 3.0$ cm and with an angle in the CDC angular acceptance. The continuum backgrounds are further minimized by training a Multivariate Analysis (MVA) algorithm, as described in Section 6.4.

Electron and muon pairs coming from the J/ψ ($m = 3.096$ GeV) meson decays are a
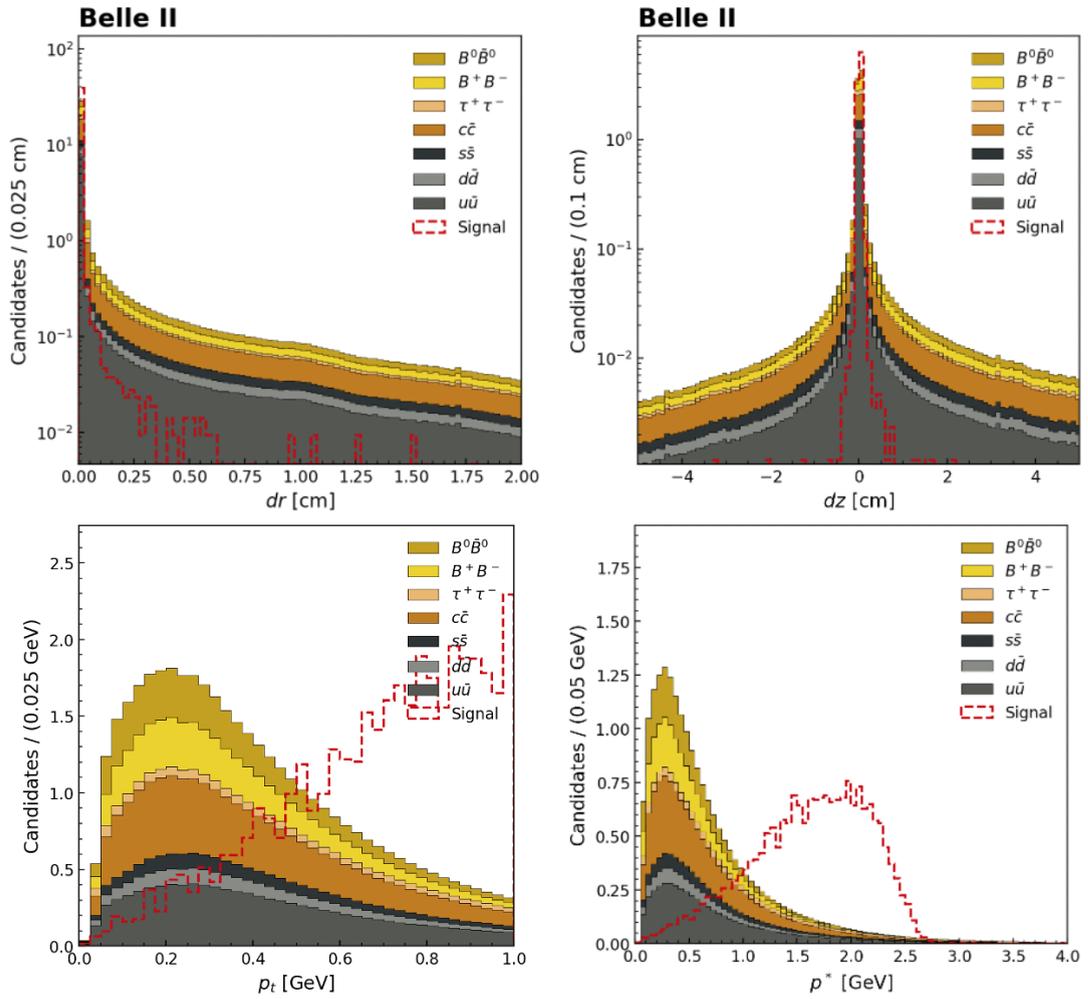
Figure 6.1: Distributions of variables used for track selection in MC. The continuum background and $B\bar{B}$ MC are shown in solid colours, while the signal distribution is shown with a dashed red line. Arbitrary scaling is used. The scaling for the signal distribution is higher than the background scaling, in order for the signal to be visible.

6.1 Signal selection requirements

significant source of background that can be easily suppressed. Charmonium vetoes were applied to electron and muon candidates as follows:

- Electron-pair invariant-mass exclusion: $2.9 < M_{e^+e^-} < 3.2$ GeV,

- Muon-pair invariant-mass exclusion: $2.9 < M_{\mu^+\mu^-} < 3.2$ GeV.

For the purpose of better understanding the signal distribution and the background contributions, the $B\bar{B}$ MC samples are further subdivided for the purpose of this analysis. The division is made according to the truth information of the selected signal lepton, as follows:

- $b \to u$ sample - the candidate lepton is from a $B \to X_u \ell\nu$ event

- $b \to c$ sample - the candidate lepton is from a $B \to D\ell\nu$, $B \to D^*\ell\nu$, $B \to D^{**}\ell\nu$ or other $B \to X_c\ell\nu$ decay

- J/ψ sample - the candidate lepton is a daughter of a J/ψ meson

- $secondaries$ sample - the candidate lepton is a daughter of neither a B meson nor a J/ψ

- $fake$ sample - the candidate lepton is not a true lepton

- $other$ sample - the candidate leptons that don't belong to any of the previous categories.

The momentum distributions of these different MC components is shown in Fig. 6.2 and Fig. 6.3 for electrons and muons, respectively. The $b \to u$ sample contains the signal events. The $b \to c$ sample is the most dominant background in the analysis. These events are very similar to signal events as they are semileptonic B decays as well, but they contain a charmed meson (D, D^*, D^{**}) in the final state in place of the charmless meson in the sig-

nal sample. The J/ψ sample contains the events where the selected signal sample is a true daughter of a J/ψ meson. This type of background is heavily reduced by the previously described charmonium veto. The *secondaries* sample represents the events where the selected lepton is not a daughter of a B meson, but comes from a secondary decay of a B meson daughter. In this case, the lepton most commonly comes from a semileptonic decay of a charmed meson (e.g. $D^- \to K^0 \ell^- \nu$). Electrons or muons could also be produced in τ decays from $B \to X_u \tau \nu_\tau$ decay, but this contribution is not significant. The *fake* sample contains the events where the selected lepton candidate is not a true lepton. In this case, the particles that are falsely identified as leptons are $\pi^\pm$, $K^\pm$, p^+, $\Sigma^\pm$ and $\Xi^\pm$. For the pions and kaons that are falsely identified as leptons, there are corrections available to improve the agreement between the MC simulation and data, as described in Section 6.3.5. All the events that do not fall into any of the previously described categories are grouped into the *other* sample. These are mainly the semileptonic decays of the B mesons that involve a strange quark, such as $B \to K\ell\ell$ and $B \to K\pi\ell\ell$.

6.2   Continuum Monte Carlo and off-resonance data

A portion of the Belle II data was taken at the center-of-mass energy 60 MeV below the $\Upsilon(4S)$ resonance, as described in Section 5.2. This is customarily done at B factory experiments in order to have better understanding of the continuum backgrounds. This type of data is often used in place of the continuum MC simulation because it more accurately represents the continuum backgrounds in the experiment. The amount and the kinematic properties of continuum backgrounds are highly dependent on the accelerator settings during data taking, mainly the center-of-mass energy of the beams. The MC simulation that

Figure 6.2: Momentum distributions for different categories of $B\bar{B}$ MC, for events with an electron candidate. The top left plot shows the momentum distribution for the signal $b \to u$ sample. Both the resonant contribution and the non-resonant contributions from BLNP and DFN models are indicated.The statistical uncertainty is shown.

Figure 6.3: Momentum distributions for different categories of $B\bar{B}$ MC, for events with a muon candidate. The top left plot shows the momentum distribution for the signal $b \to u$ sample. Both the resonant contribution and the non-resonant contributions from BLNP and DFN models are indicated. The statistical uncertainty is shown.

is used in this analysis is generated using a singular set of accelerator parameters, which does not accurately depict the data-taking conditions, which can vary over time. The off-resonance data is taken at different times of data-taking, to better represent the accelerator conditions of different data-taking periods.

By initially comparing the data and MC agreement it was evident that the continuum MC was not accurately representing the continuum processes seen in data. The samples compared have the selection requirements from Section 6.1 applied. The continuum MC was reweighted to match the off-resonance sample luminosity. The disagreement is assumed to be due to variations in operating conditions of the accelerator between different data-taking periods, mainly the small but significant variations in the center-of-mass energy of the collisions. Because of this issue, the data taken off-resonance is used to approximate the continuum contributions in the on-resonance data. The comparison of off-resonance data to continuum MC is given in Fig. 6.4.

The disagreement is most evident when looking at the R_2 variable. The R_2 variable quantifies the event shape, as described in more detail in Section 6.4. The values close to 1 represent jet-like continuum events, while events with a more spherical distribution have R_2 values close to 0, which are mainly $B\bar{B}$ events. Looking at the R_2 distribution in Fig. 6.4, it is evident that the MC simulation fails to accurately depict the continuum events seen in the detector. The peak in the off-resonance sample that is seen close to the R_2 value of 1 can be explained by Bhabha events, which are not included in the MC simulation. The Bhabha events were mostly expected to be eliminated with the requirement on the number of tracks in the event to be greater than 4. But seeing this peak in the off-resonance data tells us that a significant number of such events passes this cut. Still, this effect is not problematic since a
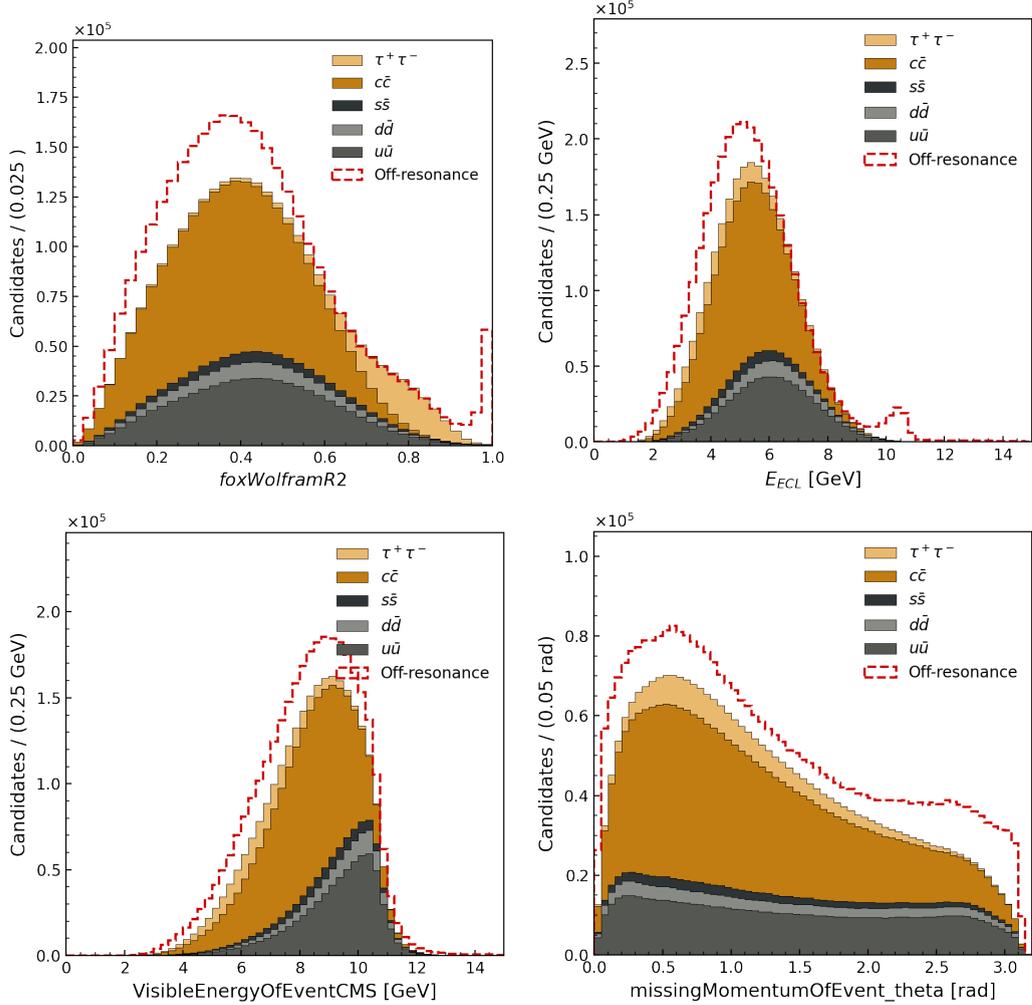
Figure 6.4: Comparison between continuum MC (solid colours) and off-resonance data (dashed red line). The MC samples were reweighted to match the off-resonance data luminosity. The R_2 variable, defined in the text, is shown top-left; the total energy detected in the ECL is shown top-right; the bottom-left figure shows the total visible energy in the event and the bottom-right figure shows the θ direction of the missing momentum of the event.

simple cut on the R_2 variable would eliminate these events. The disagreement in the lower R_2 region is problematic. This tells us that there is a significant contribution of 'spherical', $B\bar{B}$-like events, which the continuum MC does not accurately simulate.

Another variable used to compare off-resonance data and continuum MC, shown in Fig. 6.4 is the total energy seen by the calorimeter, E_{ECL}. The excess of events is seen around the $\Upsilon(4S)$ mass, 10.58 GeV. At lower energies, there is also an excess of events seen in off-resonance data.

To avoid these issues, it was decided to use solely the off-resonance data in place of continuum MC. The size of the off-resonance dataset is only about 10% of the dataset taken at $\Upsilon(4S)$ resonance, which introduces significant statistical uncertainty, the benefit is that the continuum events are more accurately represented with this dataset. The inadequate representation of continuum backgrounds in MC simulation was also seen by other Belle II analyses, mainly the untagged analyses with large selection efficiencies. A set of standard cuts was prescribed to mitigate this issue and those cuts were also applied in this analysis, from this point onwards. Those are the requirement on the total visible energy of the event in the center-of-mass system, $4 < \mathtt{E_{CMS}^{vis}} < 10$, and the requirement that the missing momentum of the event points within the detector acceptance, $0.2967 < \mathtt{p}_\theta^{\mathtt{missing}} < 2.705$.

6.3   Data and MC corrections

6.3.1   Luminosity correction

A normalization needed to be applied to the MC samples to account for the difference in luminosities between the generated MC samples and the recorded dataset. The MC samples

were normalized by the ratio of luminosities:

$$\frac{\mathcal{L}_{on-res\ data}}{\mathcal{L}_{MC}} = 1.893 \qquad (6.2)$$

where $\mathcal{L}_{on-res\ data}$ is the measured luminosity of the on-resonance data and the $\mathcal{L}_{MC}$ is the luminosity of the simulated MC sample. The MC samples were generated with a center-of-mass energy of $\sqrt{s_{MC}} = 10.580$ GeV.

6.3.2 Track momentum scaling in data

A momentum scaling of 0.99971 is applied to all tracks in data. This is done to correct for a discovered shift of invariant mass peak positions of $D^{0,+}$, J/ψ, K_S and Λ_c^+ states in the recorded dataset compared to the corresponding invariant mass value from the PDG used in MC simulation [91].

6.3.3 Photon energy bias correction in data

To correct for the photon energy bias in the region below 2 GeV, corrections are applied to the reconstructed photons. These corrections are obtained by analyzing the decays of π^0 and η into two photons [92].

6.3.4 Bremsstrahlung correction

The electron candidates (i.e. the candidates that pass the track selection cuts and the electronID > 0.9 requirement) are corrected for Bremsstrahlung emission during event selection. Bremsstrahlung radiation is a braking radiation that charged particles produce

while decelerating. Electrons are more susceptible to emitting Bremsstrahlung radiation than muons due to their lower mass.

In `basf2`, the `correctBremsBelle` module is used to handle the Bremsstrahlung correction for electron candidates. In order to account for the lost momentum due to the Bremsstrahlung radiation, photons that fall within a cone of $3.5°$ centred around the electron's flight path are identified as Bremsstrahlung photons. Photons are required to have an energy below 1.2 GeV. To reduce the probability of adding the low-energy photons coming from SuperKEKB machine backgrounds, specific minimum ECL cluster energy requirements are set. For the clusters in the forward ECL region, the minimum required cluster energy is 75 MeV. For the barrel region of the ECL, the minimum cluster energy is set to 50 MeV. For the backward region, the minimum required energy is 100 MeV. The required minimum cluster energy is the highest for the backward ECL region since that region is most affected by machine backgrounds. Subsequently, the four-momentum of the photon candidates that satisfy these requirements is added to the momentum of the corresponding electron candidate.

To verify that the Bremsstrahlung correction behaves as intended, some checks were performed on MC. For this analysis, it is important that the electron momentum is not over-corrected. An overcorrection would artificially push electron candidates into a higher momentum region than their true momentum. This would be problematic as it would wrongly inflate the electron yield in the endpoint region. To make sure this is not the case, the electron momentum after the Bremsstrahlung correction was compared to the true generated momentum in MC. The electron momentum before and after applying the Bremsstrahlung correction is shown in Fig. 6.5. Figure 6.6 shows the two-dimensional distribution of reconstructed momentum in the center-of-mass frame after the Bremsstrahlung correction is

applied, versus the ratio of the reconstructed and the generator level center-of-mass momentum. It can be seen that there are instances where the reconstructed momentum is over-corrected and the lepton is assigned a higher reconstructed momentum than what was generated. The root mean square of the difference between the reconstructed and generated momentum was calculated, both with and without the Bremsstrahlung correction. By comparing the root mean square values, it is evident that the Bremsstrahlung correction on average improves the difference between the generated and the reconstructed momenta. The comparison is shown in Table 6.1.

	$b \to u$	$b \to c$	J/ψ	$secondaries$	$fakes$	$other$
$\sqrt{(p^*_{uncorrected} - p^*_{gen})^2}$ [GeV]	0.027	0.013	0.030	0.005	0.001	0.018
$\sqrt{(p^*_{corrected} - p^*_{gen})^2}$ [GeV]	0.016	0.005	0.019	0.005	0.004	0.007

Table 6.1: Root mean square of the difference between the reconstructed and generated momentum. Here $p^*_{corrected}$ and $p^*_{uncorrected}$ indicate the center-of-mass momentum of the electron with and without the Bremsstrahlung correction, respectively, while p^*_{gen} indicates the generator level center-of-mass momentum.

6.3.5 Efficiency and fake rate corrections for particle identification

Because the simulation does not perfectly model the detector and its response, some calibration is needed to match the efficiency of particle identification (PID) in the simulation to the real detector response. Electron and muon identification efficiencies are studied using $e^+e^- \to \ell^+\ell^-(\gamma)$, $e^+e^- \to (e^+e^-)\ell^+\ell^-$ and $J/\psi \to \ell^+\ell^-$ events. These processes are well known, so they can be used to compare the performance of the PID in data and in simulation and to derive the needed calibrations to match these performances. The Belle II collaboration's performance group provides the calibration factors that are applied to the correctly reconstructed electrons and muons in MC. The corrections are given in bins of
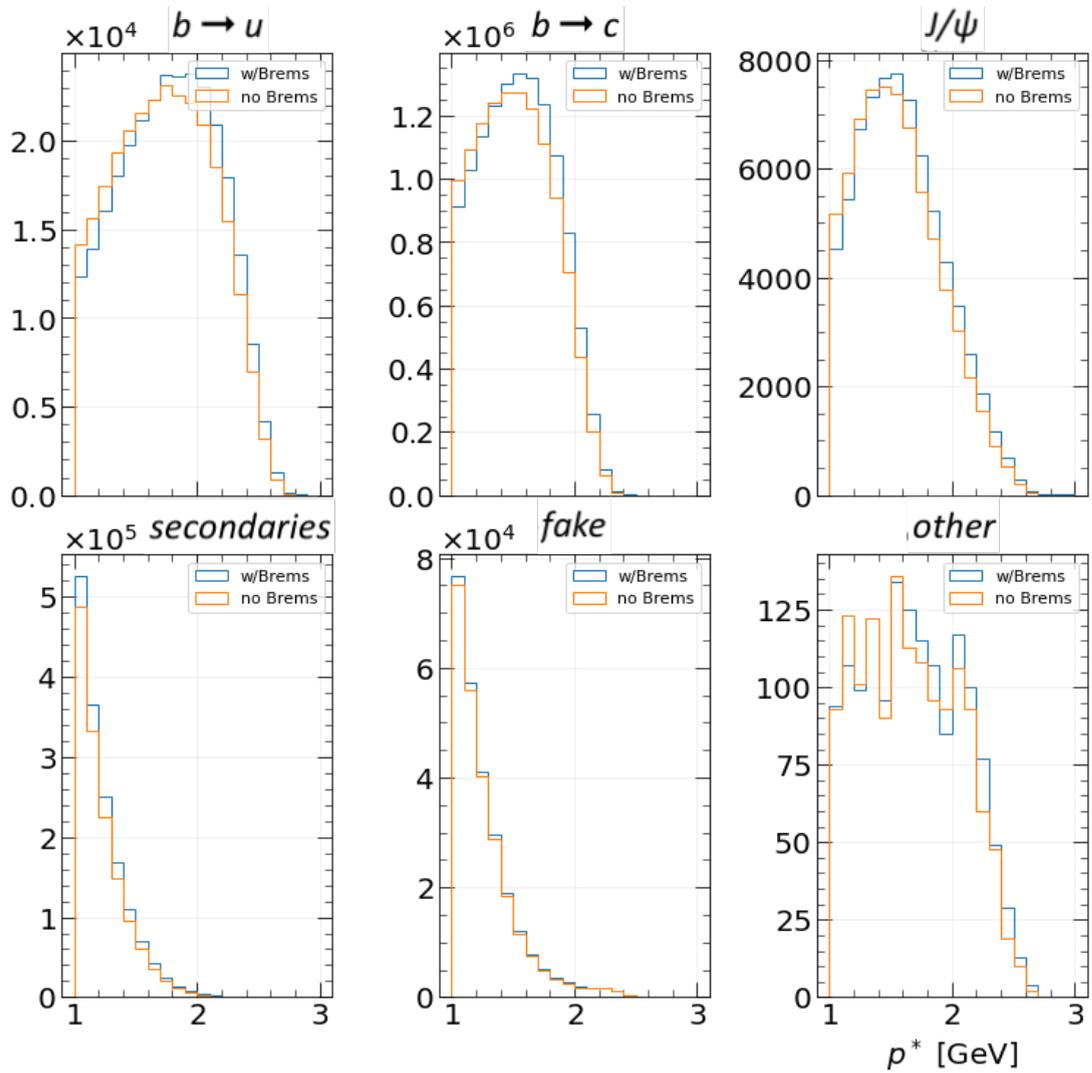
Figure 6.5: The distribution of the center-of-mass lepton momentum before the Bremsstrahlung correction, shown in orange; the corrected value is shown in blue.
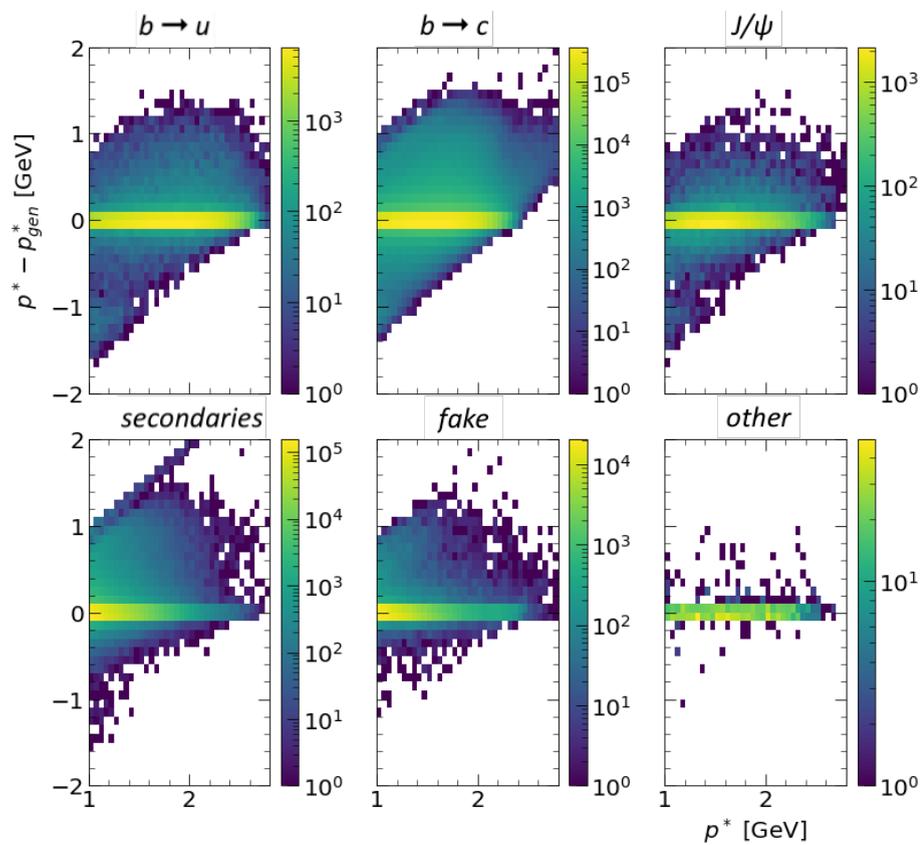
Figure 6.6: 2D distribution of the reconstructed center-of-mass electron momentum with Bremsstrahlung correction on the x-axis and the difference from the generator level value on the y-axis.

momentum, p, and azimuthal angle, θ. The bin coverage of the correctly identified leptons in MC is given in Fig. 6.7. All the electron candidates are covered by the provided corrections, while 1% of muon candidates is not covered by these corrections.

Similarly, the pion-lepton mis-identification probability (fake rate) is studied using the
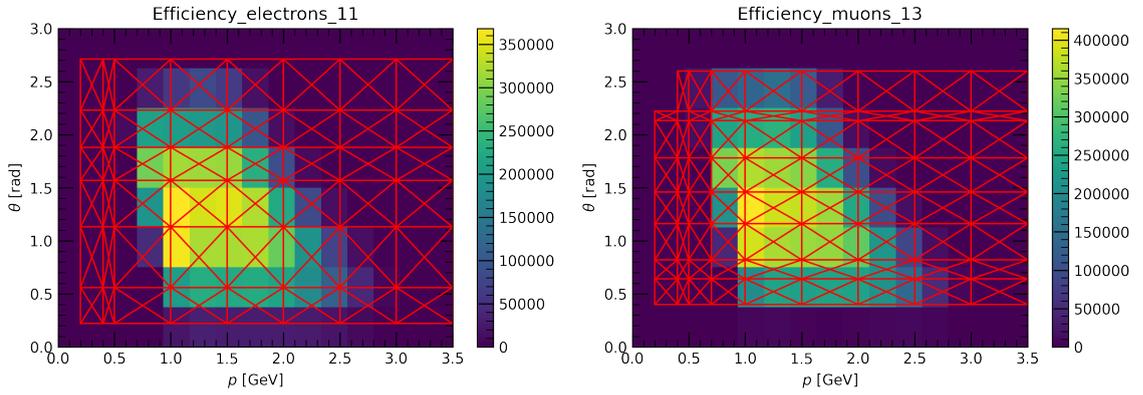
Figure 6.7: The 2D MC sample distributions of correctly identified electrons (left) and muons (right) in angle, θ [rad], vs. momentum, p [GeV], is shown, overlaid with the lepton identification efficiency correction bins that are indicated with red rectangles. The colour scale depicts the number of candidates.

$K_S^0 \rightarrow \pi^+\pi^-$ channel, and the kaon-lepton mis-identification probability is studied using the $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ channel. The performance group provides the corrections for cases when pions or kaons are mis-identified as electrons or muons. These corrections are provided in bins of p and θ, as well. Due to the limited statistics of the samples used for estimating the fake rate corrections, there are some gaps in coverage. The coverage for pions and kaons mis-identified as leptons is given in Fig. 6.8. In the case of pions mis-identified as electrons, 0.01% of candidates are not covered by the corrections. There is 0.7% of pions mis-identified as muons not covered by the corrections. The kaon corrections have more limited coverage, with 0.25% of electron candidates not covered by these corrections and 0.5% for muon candidates. In these cases, the candidates are not assigned any corrections.
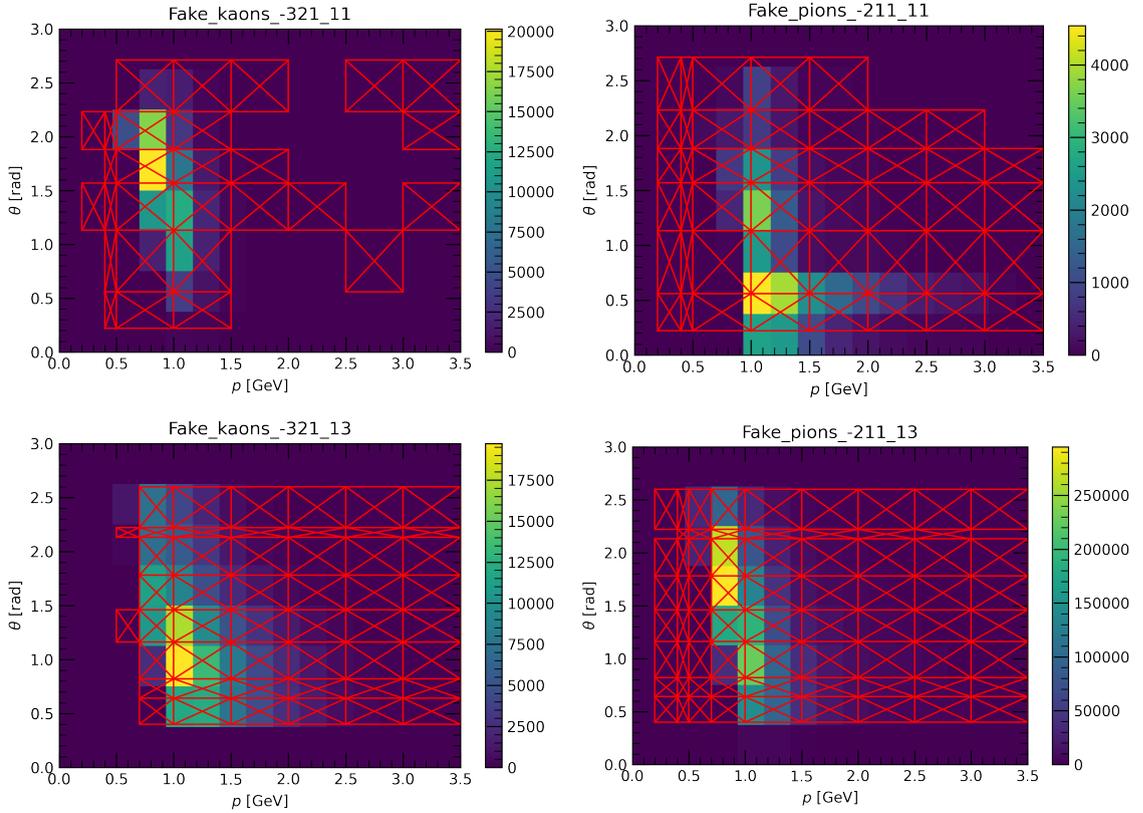
Figure 6.8: The 2D MC sample distribution of kaons (left) and pions (right) faking electrons (top) and muons (bottom) in angle, θ [rad], vs. momentum, p [GeV], is shown, overlaid with the lepton fake rate correction bins that are indicated with red rectangles. The colour scale depicts the number of candidates.

In this analysis, the PID efficiency corrections and the fake rate corrections are applied to the lepton candidates after the initial signal selection. The corrections are read from the provided tables using the `PIDvar` package developed by the Belle II collaboration members [93].

The statistical and systematic uncertainties for the efficiency and fake rate MC correction factors are provided by the performance group, and are later used to evaluate the systematic uncertainty due to the PID corrections. The average values and the uncertainties of

the PID efficiency corrections and fake rate corrections per momentum bin are shown in Fig. 6.9. The Belle II Performance Group continuously analyzes and provides updates to these corrections. With the data coming from the Belle II experiment, the PID performance is better understood and improved. This analysis uses PID correction factors released under the name `post-ICHEP-2022`; these are the most recent corrections made available for the dataset and MC campaign used in this analysis.

6.3.6 Continuum normalization

To replace the continuum MC with off-resonance data, the off-resonance data sample was scaled to account for the difference in luminosity and cross sections at lower center-of-mass energy by a factor r_L defined as:

$$r_L = \frac{s_{off}}{s_{on}} \frac{\int L_{on} dt}{\int L_{off} dt} = 10.429 \tag{6.3}$$

where s and $\int L dt$ refer to the center-of-mass energy squared and integrated luminosity of the on- and off-resonance samples. The off-resonance and on-resonance data samples from different data-taking periods were taken at slightly different center-of-mass energies. The mean center-of-mass energies were used in this reweighting, which were found to be:

$$\sqrt{s_{on}} = 10.576 \text{ GeV}, \tag{6.4}$$

$$\sqrt{s_{off}} = 10.518 \text{ GeV}. \tag{6.5}$$

For the off-resonance data, the momentum of the lepton candidates in the center-of-mass frame was scaled by the ratio $\frac{\sqrt{s_{on}}}{\sqrt{s_{off}}}$, to account for the different center-of-mass energies

(a) electronID efficiency corrections

(b) muonID efficiency corrections

(c) electronID pion fake rate corrections

(d) muonID pion fake rate corrections

(e) electronID kaon fake rate corrections

(f) muonID kaon fake rate corrections
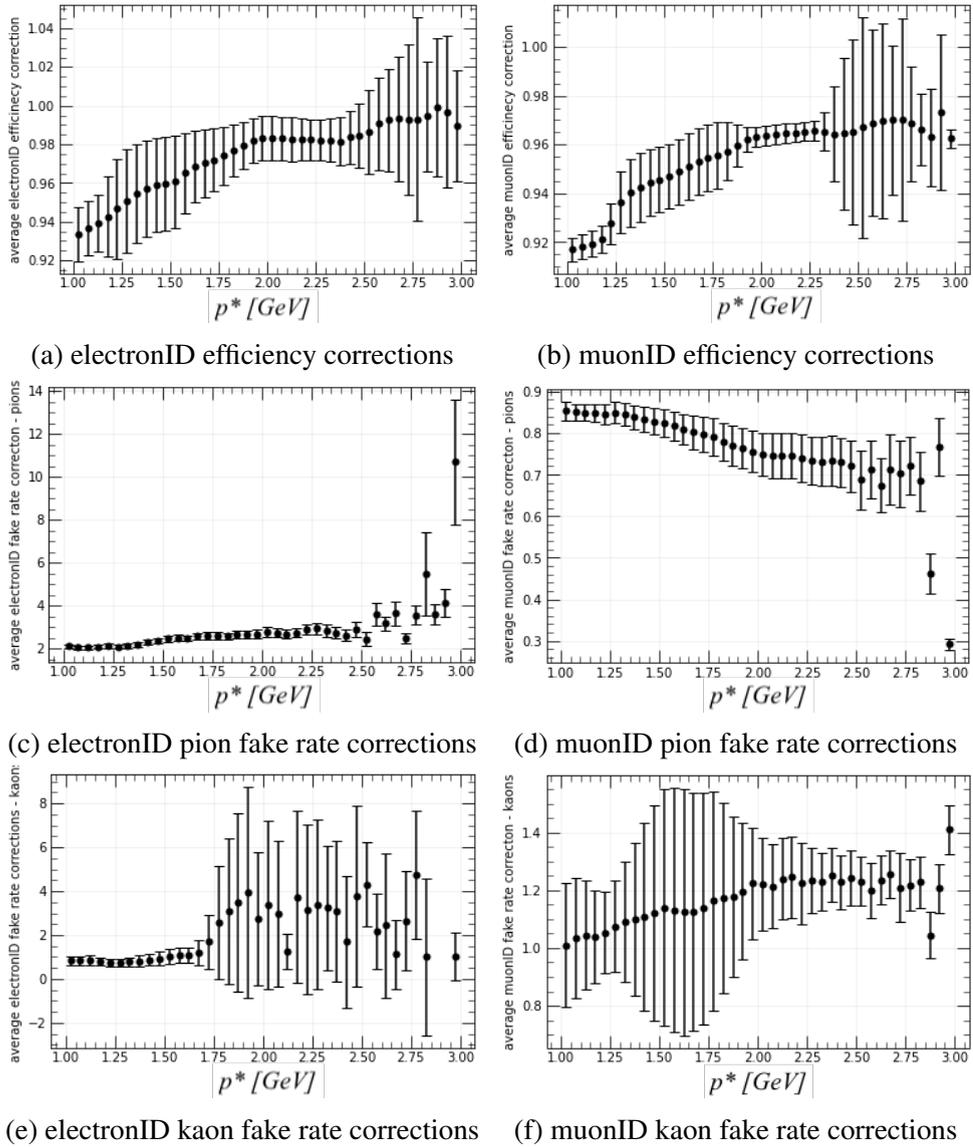
Figure 6.9: Average efficiency and fake rate corrections over the lepton momentum
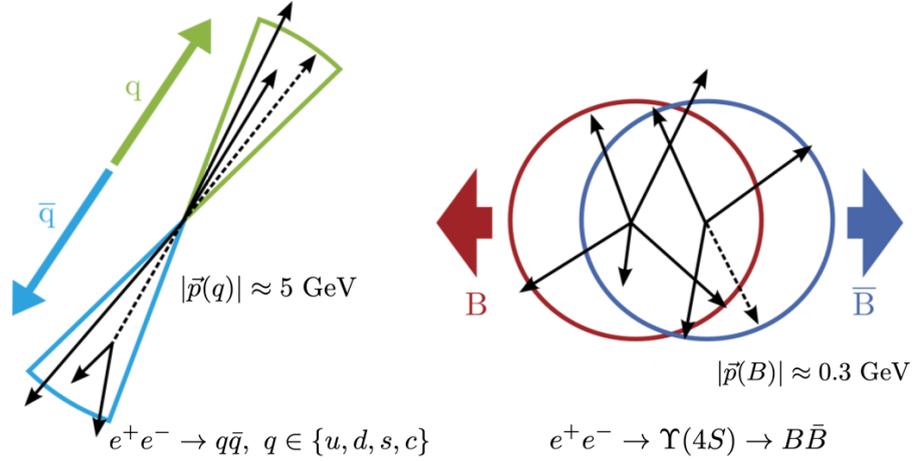
Figure 6.10: Comparison of the event shapes in a continuum event (left) and a $B\bar{B}$ event (right) in the center of mass reference frame.

when boosting to the CM frame from the lab frame.

6.4 Continuum suppression

As seen in Table 5.2, the branching fraction for the continuum backgrounds is high when operating at $\sqrt{s} = 10.58\,\text{GeV}$. The continuum background contribution needs to be reduced in order to increase the statistical significance of signal events. This can be done by taking advantage of the fact that the shapes of the continuum events differ from the $B\bar{B}$ events. Since the center-of-mass energy of e^+e^- collisions is exactly that of the $\Upsilon(4S)$ mass and the B meson pairs are produced almost at rest, their decay products are distributed evenly in the detector and the events are of spherical shape. In contrast, the collision products in the continuum events have a large boost and these events have a jet-like shape. This is illustrated in Fig. 6.10.

The event shapes can be quantified using different variables. The event shape variables are

calculated using the `good tracks` and `good clusters`, as described in Section 6.1. One type of such variables are the Fox-Wolfram moments [94]. The k-th Fox-Wolfram moment is defined as:

$$H_k = \sum_{i,j}^{N} |\vec{p_i}||\vec{p_j}|P_k(\cos\theta_{ij}), \tag{6.6}$$

where θ_{ij} is the angle between the momenta, $\vec{p}$, of particles i and j, and P_k is the k-th Legendre polynomial. The reduced Fox-Wolfram moments are defined as $R_k = H_k/H_0$. The most commonly used variable describing the event shape in B-factory experiments is the second reduced Fox-Wolfram moment, R_2. The R_2 has values close to 0 for spherical $B\bar{B}$ events and values close to 1 for jet-like continuum events.

A cut of $R_2 < 0.5$ is applied offline, after the event reconstruction and selection. The R_2 distribution in data and MC is shown in Fig. 6.11. The agreement between data and MC is not ideal and this is attributed to imperfect simulation of full event kinematics and to the small variations of the beam center-of-mass energies between different run periods and in MC simulation.

A multivariate analysis (MVA) based on boosted decision trees was used to further separate continuum and $b \rightarrow u$ signal events. A training set is used to 'teach' a classifier to distinguish between the signal and continuum background events. The MVA classifier is given a target variable, which is a binary variable that labels continuum events and signal events. A set of observable variables is also given, called the training variables, which have separation power between signal and background. Based on this information, the classifier assigns a signal probability to each event, such that the signal events have probability close to 1 and continuum events are assigned a probability close to 0. The performance of the classifier is evaluated using a test sample. For the test, the classifier is given the test data with its training variables values and the results are cross-checked with the target variable,
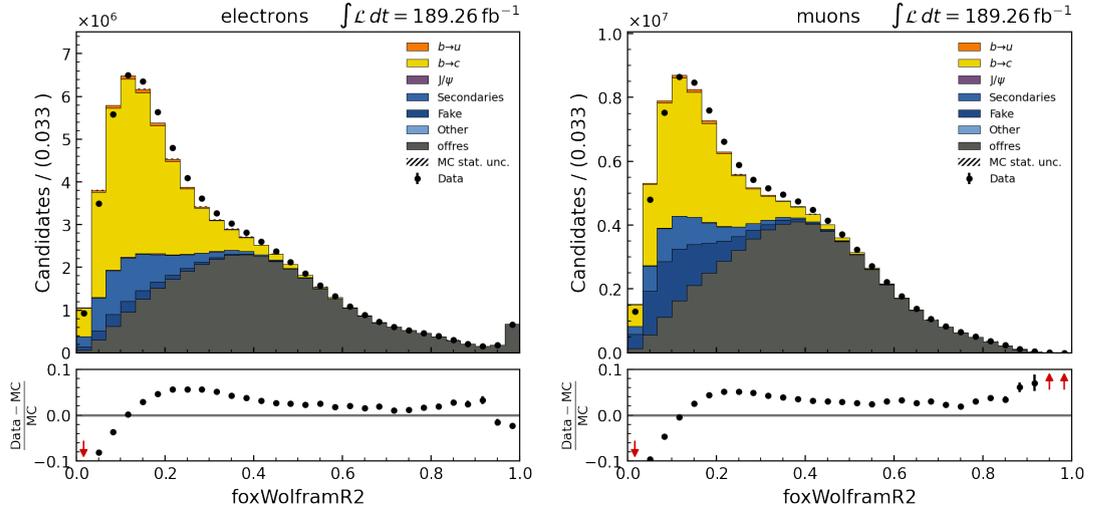
Figure 6.11: The 2nd order Fox-Wolfram moment, R_2, of the events; comparison between data (black dots) and MC (solid colours).

which in the test case is only available to the analyst. If the tested performance proves satisfactory, the classifier is then applied to the observed data and MC using the training variables. The classifier gives a classifier probability value for each event in the given data. A cut can then be applied on the new classifier probability variable to reduce the background contributions in the selected events.

For the MVA training, the FastBDT MVA model was used. FastBDT [95] is a stochastic gradient-boosted decision tree method which was optimized to provide good computing performance when used on the large Belle II data samples. A simple decision tree (DT) classifier uses a series of consecutive cuts to separate between signal and background. DT classifiers are sensitive to statistical fluctuations in the training sample and are prone to over-fitting. Boosted DT (BDT) classifiers use multiple weak-learner DTs that roughly separate signal and background and are thus more robust. Their performance can be further improved by using a gradient-descent in each boosting step to reweight the training sample.

6.4 Continuum suppression

A stochastic gradient BDT further uses a randomly drawn sub-sample of the training set in each boosting step. This approach is the most robust against over-fitting.

The classifier was trained on 80k events of $b \rightarrow u$ MC and off-resonance data each. It was tested using 20k events of each sample. The training was done separately for samples with electrons and muons. The training variables used and their definitions are listed in Table 6.2. The distribution of the 9 most important variables in the signal MC and the comparison between data and MC are shown in Fig. 6.12 for electrons, and in Fig. 6.13 for muons.

The performance of the MVA training is evaluated using the test samples. The comparison of the classifier output in the training sample and in the test sample is shown in Fig. 6.14. The agreement between the two distribution confirms that there is no over-training of the MVA. The receiver operating characteristic (ROC) curves are also used to estimate the MVA performance. They show the relation between the signal selection efficiency of the classifier versus the background rejection. These curves were compared between the training and the test samples, as well as the area under the curve (AUC), as shown in Fig. 6.15. The good agreement between the ROC curves and the AUC values is another indication that there was no over-training.

To decide on an optimal value for the cut on the classifier output, a figure of merit (FOM) is used. It is defined as:

$$FOM = \frac{N_{sig}}{\sqrt{N_{sig} + N_{bkg}}},$$ (6.7)

where N_{sig} is the number of signal events and N_{bkg} is the number of background events that pass the cut. By calculating the FOM value for cuts on various values of the classifier,

Variables	Description
`thrust`, T_Z, $\cos\theta(T)$	Thrust axis $\vec{T}$ is defined as the unit vector along which the total projection of all the momenta of the particles in the event is maximal [27]. The thrust is derived as: $T = \frac{\sum_{i=1}^{N} |\vec{T}\cdot\vec{p_i}|}{\sum_{i=1}^{n} |\vec{p_i}|}$. $cos\theta(T)$ is the cosine of the polar angle component of the thrust axis and T_z is the z component of the thrust axis.
$B_0 - B_4$	The harmonic moments [94] with respect to the thrust axis, from 0th to 4th order, where $B_\ell = \sum_{i=1}^{N} \frac{p_i}{\sqrt{s}} P_\ell(\cos\alpha_i)$
$p^*_{miss}, \theta(p^*_{miss}), p^*_{miss\ z}$	The total momentum of the detected particles in an event, p^*_{miss} and the θ and z projections, $\theta(p^*_{miss}), p^*_{miss\ z}$
R_1, R_3, R_4	Reduced Fox-Wolfram moments of order 1, 3 and 4
$M^2_{miss}, M^2_{miss}/E_{miss}$	The missing mass squared, M^2_{miss} and the ratio of the missing mass squared and missing energy in the event, M^2_{miss}/E_{miss}
$E_{FH}, p_{FH}, E_{BH}, p_{BH}$	The forward hemisphere (FH) energy, E_{FH}, and backward hemisphere (BH) energy, E_{BH}, are defined as the total energy of the particles flying in the same and in the opposite direction to the thrust axis, respectively. Similarly, the forward and backward hemisphere momenta are defined as the total momenta of the particles flying in the same and in the opposite direction to the thrust axis.

Table 6.2: The list of the variables used in continuum suppression MVA and their descriptions.

(a) The thrust of the events. (b) The z-axis projection of (c) The θ direction of the the thrust of the events. missing momentum in the event.

(d) The missing momentum (e) The forward energy of the (f) The 0th order harmonic of the event in the CMS event. moment w.r.t. the thrust axis. frame.

(g) The 2nd order harmonic (h) The 4th order harmonic (i) The 1st order Fox-moment w.r.t. the thrust axis. moment w.r.t. the thrust axis. Wolfram moment.

Figure 6.12: The distributions of variables used for continuum suppression MVA for electrons. The data and MC comparison is shown. The distribution in the signal $B \to X_u \ell \nu$ MC is shown additionally in red, with a 50 times increased yield.
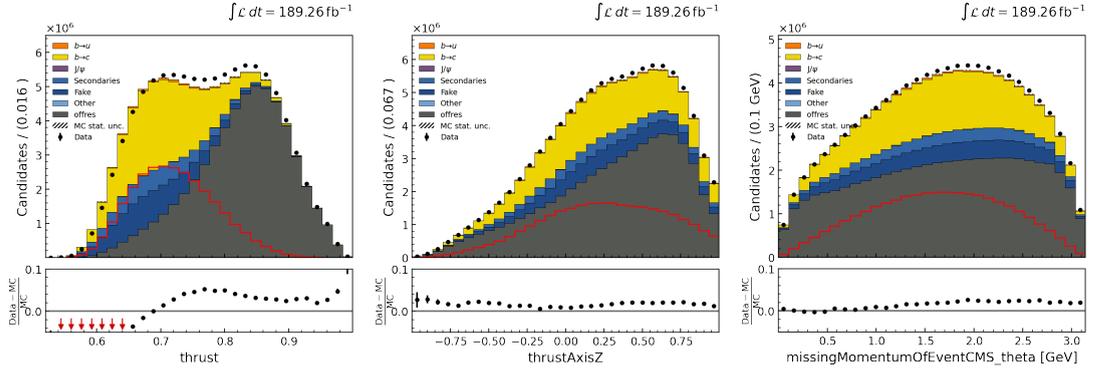
(a) The thrust of the events. (b) The z-axis projection of the thrust of the events. (c) The θ direction of the missing momentum in the event.
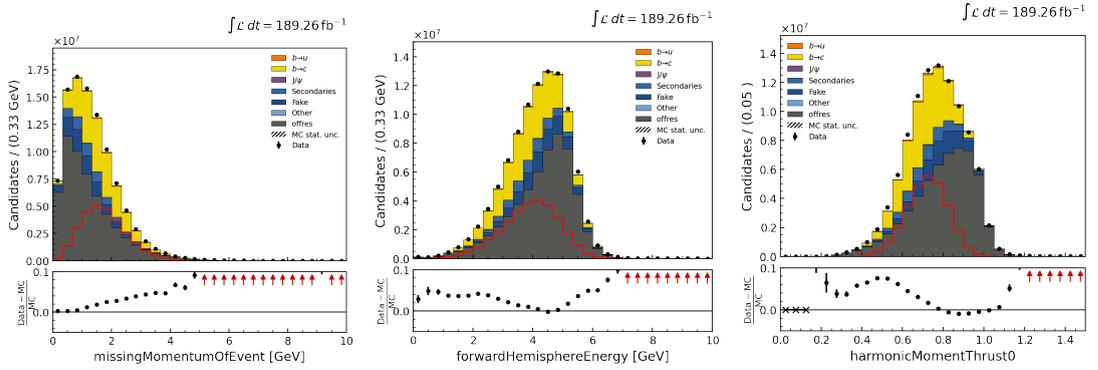
(d) The missing momentum of the event in the CMS frame. (e) The forward energy of the event. (f) The 0th order harmonic moment w.r.t. the thrust axis.
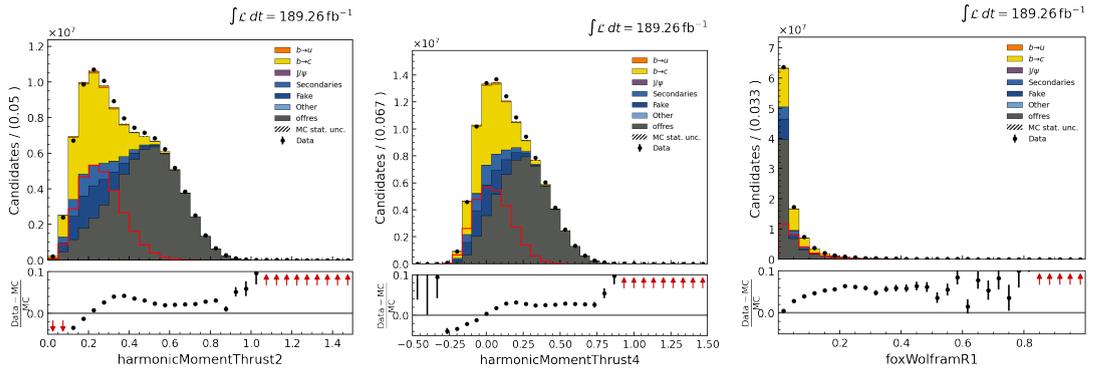
(g) The 2nd order harmonic moment w.r.t. the thrust axis. (h) The 4th order harmonic moment w.r.t. the thrust axis. (i) The 1st order Fox-Wolfram moment.

Figure 6.13: The distributions of variables used for continuum suppression MVA for muons. The data and MC comparison is shown. The distribution in the signal $B \to X_u \ell \nu$ MC is shown additionally in red, with a 50 times increased yield.
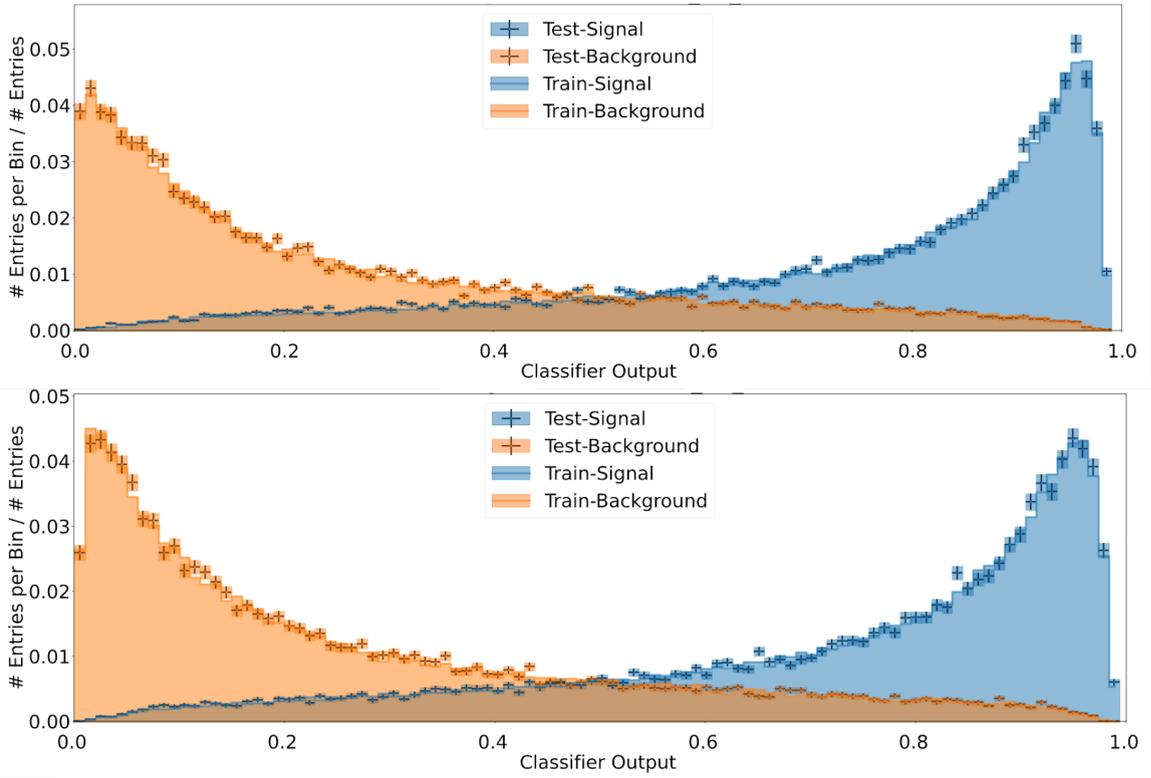
Figure 6.14: The classifier output distributions for electrons (top) and muons (bottom). The background distributions are shown in orange, while the signal distributions are shown in blue. The distributions of the training samples are shown with filled histograms, while the test sample distributions are shown with points with statistical error bars.

it was found that for both the muon and the electron sample the optimal classifier cut is at 0.4. All the events that pass the cut of `classifier` > 0.4 are kept for further analysis. The efficiency of the selected cut was evaluated on the test sample. It was found that 88.0% (87.5%) of signal electron (muon) events pass the cut. For the background, 74.5% (75.2%) of events are rejected for the electron (muon) sample.

The agreement of the MVA classifier variable between data and MC is shown in Fig. 6.16. The good agreement between data and MC assures that the introduction of the cut on MVA classifier does not introduce any further discrepancies between data and MC.
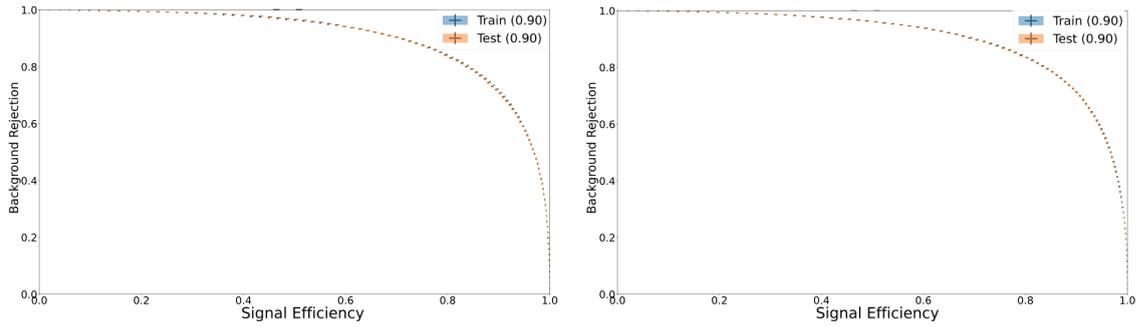
Figure 6.15: The ROC curve for the MVA background suppression training and testing samples for electrons (left) and muons (right). The value of AUC is shown in the legend for the train and test samples.
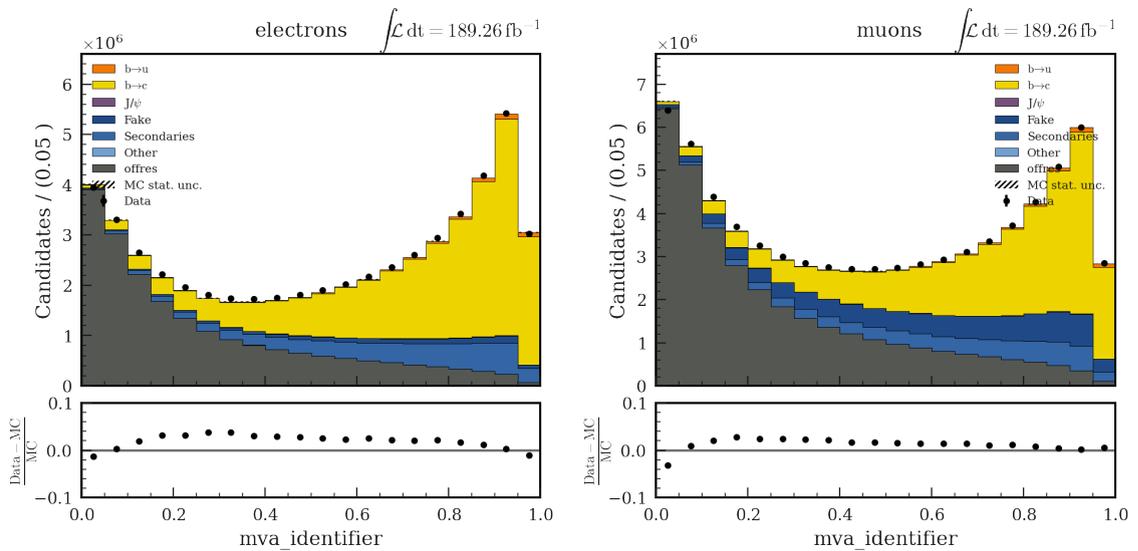
Figure 6.16: The comparison of MVA classifier variables in data and MC for electrons (left) and muons (right).

6.5 Cut flow

To verify the selection efficiency of the previously listed cuts for signal selection, a cut flow analysis is performed. For each of the cuts explained in the previous sections, the signal selection efficiency is calculated sequentially. For each of the selection requirements, the signal selection efficiency is calculated as:

$$\epsilon(\text{cut}_n) = \frac{N_{signal}(\text{cut}_n, \text{cut}_{n-1}, ..., \text{cut}_1)}{N_{signal}}, \tag{6.8}$$

where N_{signal} is the number of signal events in MC before any cuts are applied, and $N_{signal}(\text{cut}_n, \text{cut}_{n-1}, ..., \text{cut}_1)$ is the number of signal events in MC that pass all the cuts up to cut n. The results are summarized in Table 6.3. It is seen that the muon channel has higher selection efficiency than the electron channel. The difference comes from the particle ID cut, where the `muonID` cut has slightly higher efficiency than the `electronID` cut. This difference in the particle ID performance for electrons and muons is expected, since at the moment the TOP and the SVD subdetectors are not used in the electron identification. This comes from a problem on the software side of data reconstruction and it is expected to be rectified by the Belle II Physics Performance group in the near future. Also, the electrons are affected by Bremsstrahlung, as they pass through the detector, they emit radiation, which results in energy loss. Although corrections are introduced, as described in Section 6.3.4, they cannot completely compensate for all the Bremsstrahlung effects. This results in the p^* cut for electrons having lower efficiency.

Cut	Electron channel efficiency [%]	Muon channel efficiency [%]
$dr < 1.0$ cm and $|dz| < 3.0$ cm	99.7	99.9
$p_t > 0.05$ GeV	99.6	99.9
particle ID> 0.9	84.7	87.8
`thetaInCDCAcceptance`	84.5	87.7
$1.0 < p^* < 3.2$ GeV	78.3	83.8
`nGoodTracks`> 3	76.5	81.8
`nGoodClusters`> 2	74.3	79.1
J/Ψ veto	74.3	79.1
nTracks> 4	72.4	77.1
$4.0 < \mathrm{E}_{\mathrm{CMS}}^{\mathrm{vis}} < 10.0$ GeV	71.9	76.5
$0.2967 < \mathrm{p}_\theta^{\mathrm{missing}} < 2.705$ rad	66.7	70.7
R2< 0.5	66.1	70.0

Table 6.3: Cut flow table for signal selection efficiency. All the cuts are applied sequentially. The percentage efficiencies are given for electron and muon channels. The statistical uncertainty of the efficiencies is estimated to be 0.2%.

6.6 Data/MC comparison

The signal selection requirements described are applied to the data sample, the $B\bar{B}$ MC sample and the off-resonance sample. At this point, candidate ranking is also applied. Only one candidate per event is allowed. The ranking is done based on the `electronID` (`muonID`) variable for electron (muon) candidates. For each selected event, in the case of multiple potential signal candidates, only the candidate with the highest particle ID probability is kept. After all the signal selection criteria are applied, the agreement between data and MC is checked. This is done to ensure that the on-resonance data is well represented by the combination of $B\bar{B}$ MC and the off-resonance data. The fit procedure, which is explained in the next chapter, is optimized and verified using the $B\bar{B}$ MC and the off-resonance data, so it is crucial to ensure that they provide a good estimation of the on-resonance data.

6.6 Data/MC comparison

For this analysis we look at the lepton momentum in the center of mass frame between 1.0 and 3.0 GeV. The momentum region below 1.0 GeV is not of interest because it has higher contribution of fakes, which are not well simulated in MC and the data/MC agreement in that region is not indicative of the performance in the region of interest for this analysis, which is mostly above 2.0 GeV. The electron and muon spectra for the full range, $[1.0, 3.0]$ GeV, are shown in Figs. 6.17 and 6.18.

Uncertainties due to MC statistics, as well as the systematic uncertainties, are indicated with hatching on the MC histogram totals. The systematic uncertainties include the uncertainties due to the PID efficiency and fake rate corrections, the $b \rightarrow c$ branching fraction uncertainties, $b \rightarrow u$ branching fraction uncertainties, the hybrid model uncertainties and the tracking uncertainty. The systematic uncertainties are further described in Chapter 8, but first the fit to the momentum distributions and the signal extraction will be discussed in Chapter 7.
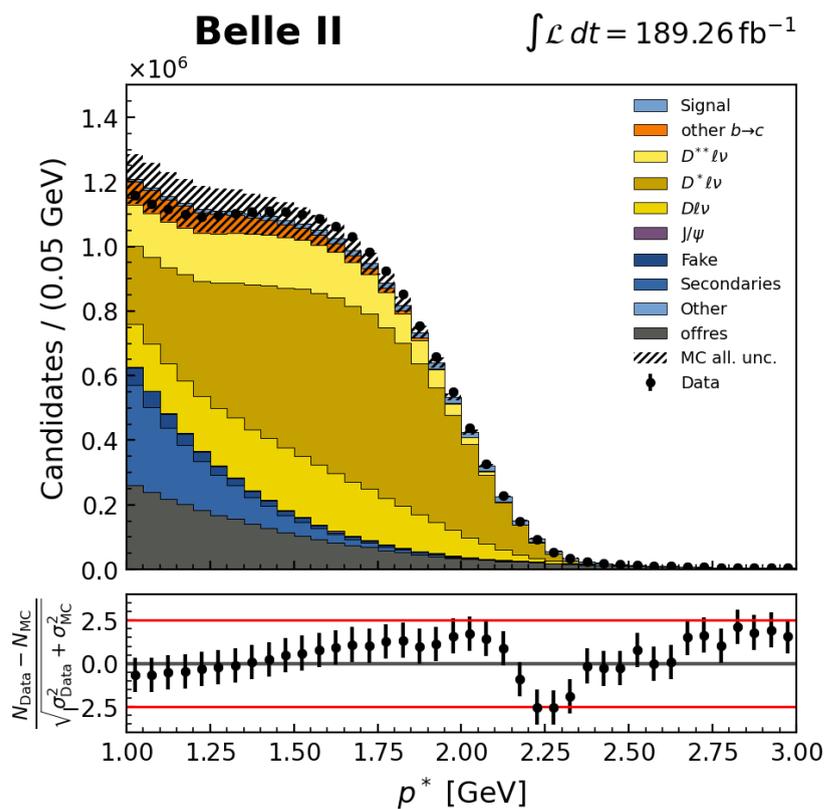
Figure 6.17: The electron momentum in the center-of-mass frame in the region between 1.0 and 3.0 GeV. Appropriately scaled off-resonance data was used instead of continuum MC.
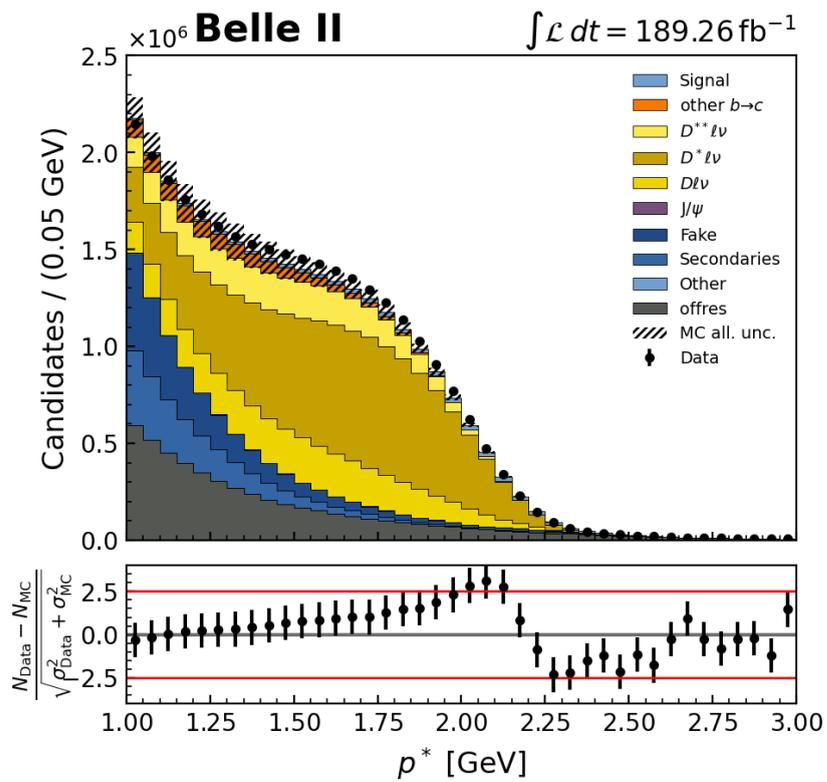
Figure 6.18: The muon momentum in the center-of-mass frame in the region between 1.0 and 3.0 GeV. Appropriately scaled off-resonance data was used instead of continuum MC.

7

Fit to the momentum spectrum and signal extraction

7.1 Fitting procedure

Knowing that the MC description of data is incomplete, a fit is needed to estimate the model parameters to obtain the best agreement between data and MC. A binned fit to the lepton momentum spectrum is performed, separately for the electron and the muon channel, to extract the signal $B \to X_u \ell \nu$ yield from the endpoint. The fit uses the MC and off-resonance momentum distributions as templates. These templates are considered to be the approximations of the probability density functions (PDF) describing various physics processes. The fit was performed between 1.0 and 3.0 GeV of p_ℓ^*. Bins of 50 MeV are used in the intervals [1.0, 2.1] GeV and [2.7, 3.0] GeV, while one wide bin was used between 2.1 and 2.7 GeV. The wide bin covers the lepton momentum endpoint and corresponds to the momentum range used for extraction of the signal $B \to X_u \ell \nu$ yield for the branching

fraction and $|V_{ub}|$ calculation. The use of the wide bin makes the fit less sensitive to the theoretical modelling of the signal shape in the lepton momentum endpoint region.

The chosen fitting approach was inspired by the fit used in the BaBar analysis [13]. The BaBar analysis used only the electron final states, while here the muons are included as well. Various other fitting approaches were attempted by the author, but this approach provided the most stable fit. BaBar's approach consists of two separate fits; a polynomial fit was performed separately to estimate the off-resonance contribution, which was afterwards subtracted from the on-resonance data. This subtracted dataset was then fitted using MC templates. This fitting approach was attempted here as well, but did not perform well and the fit had trouble converging. Different binning choices were also tested, e.g. where the sideband between 1.5 and 2.0 GeV was fitted separately, as a control region, and the region between 2.0 and 2.8 GeV was fitted in order to extract the signal yield, both using 50 MeV bin widths. This approach proved unstable, where the yields would be heavily suppressed in the control region and enhanced in the endpoint. Another approach with five fit templates was tested: the signal template, the $b \rightarrow c$ template, the fakes template, the $B\bar{B}$ background template and the off-resonance template. This approach had issues because the $B\bar{B}$ background template has a similar shape to the signal template, so the algorithm was not able to distinguish these two shapes. It was also attempted to perform a fit where the systematic uncertainties are included as nuisance parameters, but this proved to be too computationally demanding and often the code would run into memory issues.

For the chosen fitting approach, the MC and off-resonance datasets are divided into templates as follows:

- $b \rightarrow u$ **template**: describes the signal $B \rightarrow X_u \ell \nu$ decays.

7.1 Fitting procedure

- $b \to c$ **template**: describes the $B \to X_c \ell \nu$ background processes.

- **Background template**: describes both the continuum backgrounds, as obtained from the off-resonance data, as well as all the remaining backgrounds from $B\bar{B}$ processes and the mis-identified leptons (fakes). These processes were combined into one template due to their similar shape in p_ℓ^*, otherwise the fitting algorithm would not be able to distinguish between them and it would favour one process at the cost of the others.

The number of expected events in bin i is given by the expression:

$$\nu_i^{\mathrm{exp}} = \sum_j \nu^j p_i^j, \tag{7.1}$$

where p_i^j is the probability that an event from template j is found in bin i. In this analysis, the fit has three yields:

$$\nu^j = \left(N_{b \to u}, N_{b \to c}, N_{\mathrm{background}}\right), \tag{7.2}$$

and four discrete probability density functions:

$$p^j = \left(p_{b \to u}, p_{b \to c}, p_{\mathrm{background}}\right). \tag{7.3}$$

Additionally, the expected number of events in each bin i is described by a Poisson distribution, $\mathcal{P}(n_i | \nu_i)$, where n_i is the number of observed events in bin i. In the case of the large number of events, the Poisson distribution can be approximated by a Gaussian with mean n_i and the standard deviation $\sqrt{n_i}$. The optimal set of model parameters that maximizes the probability of seeing the observed data given the underlying model hypothesis is obtained

by maximizing the likelihood function:

$$\mathcal{L} = \prod_i \frac{1}{\sqrt{2\pi n_i}} \exp - \frac{(n_i - \nu_i)^2}{2n_i}. \tag{7.4}$$

This equation can be maximized by finding the minimum set of parameters that minimize the χ^2 function:

$$\chi^2 = \sum_i \left(\frac{n_i - \nu_i}{\sqrt{n_i}}\right)^2. \tag{7.5}$$

This χ^2 function can be generalized as:

$$\chi^2 = (\vec{n} - \vec{\nu})^T C^{-1} (\vec{n} - \vec{\nu}), \tag{7.6}$$

where C is the covariance matrix that describes the correlations across the bins.

A χ^2 template fit was performed using the `BinFit` package [96]. The `BinFit` package was developed for template fitting in the Belle and Belle II experiments. It uses the `Minuit`[97] package for performing the minimization of the cost function.

7.2   Fit validation

Before applying the fit to data, some validation tests are performed to ensure that the fit model is stable and unbiased. To simulate for statistical fluctuations, 2000 pseudo datasets are generated by sampling from a Poisson distribution for each bin, while the expectation values were set to the MC expectations. The fit was performed on each pseudo dataset. For
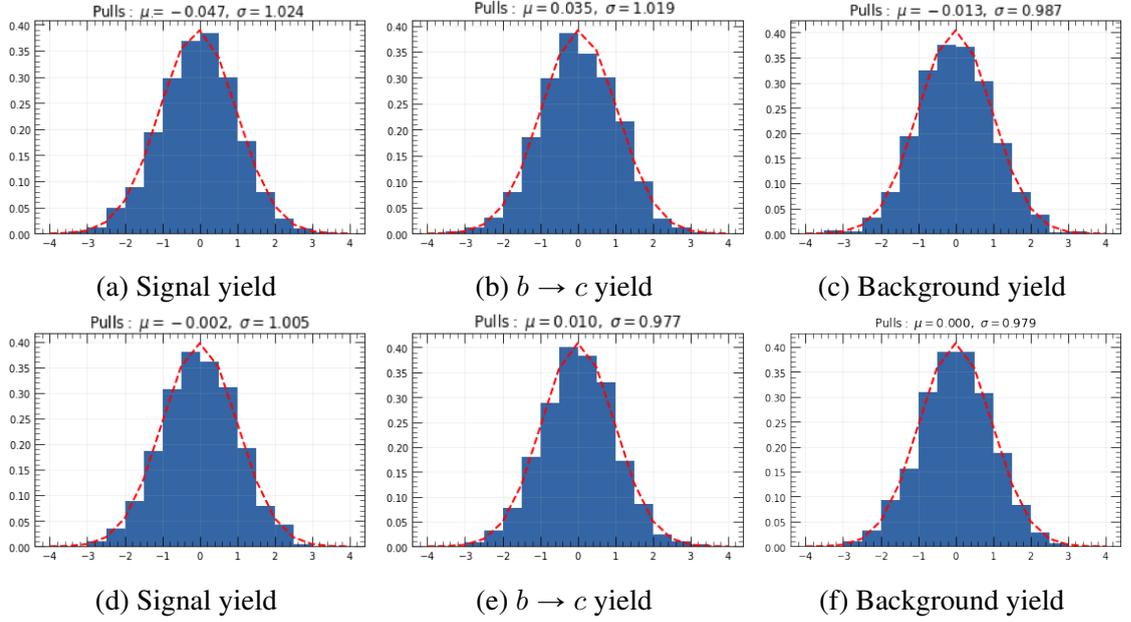
Figure 7.1: Pull distributions for each of the fit yield parameters for electrons (top) and muons (bottom). The mean and variance are obtained by fitting a Gaussian curve to the distributions.

each resulting yield, the pull was calculated:

$$\text{pull} = \frac{\nu^{\text{fit}} - \nu^{\text{expected}}}{\sigma_\nu^{\text{fit}}}. \tag{7.7}$$

For an unbiased fit, it is expected that the pull distribution is a Gaussian distribution centred around zero, with a standard deviation of 1. Figure 7.1 shows the pull distribution for each of the three yield parameters, for both electrons and muons. These distributions were fitted with a Gaussian curve, with the results of that fit indicated in the figure. All the pull distributions behave as expected, confirming that the fit is unbiased.

A linearity test was also performed. This is to check the correlation between the ex-

(a) Electron channel   (b) Muon channel
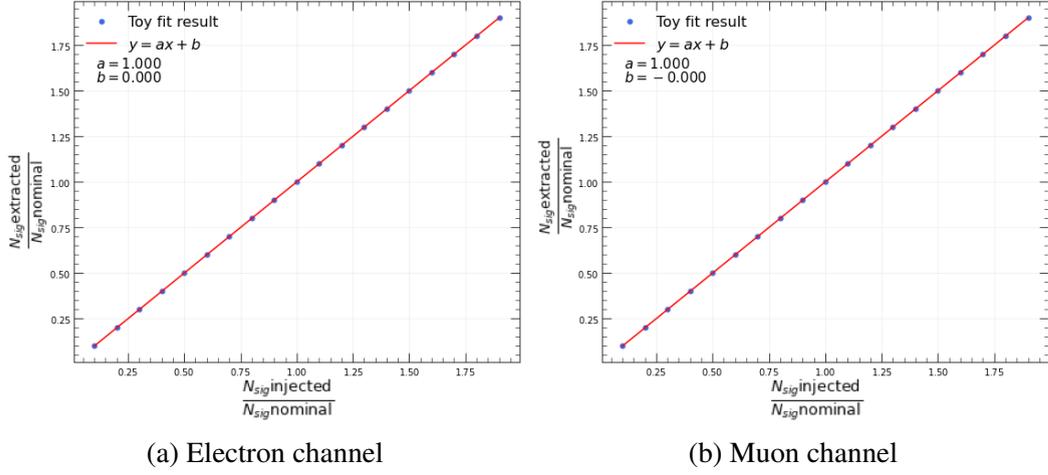
Figure 7.2: Linearity test results for the signal extraction for electrons (left) and muons (right). A linear polynomial is fitted to the extracted yield parameters and results are shown. The bottom panel shows the fit pulls, defined as the difference between the data and the fit result divided by the fit uncertainty.

pected and extracted yields. For this purpose, pseudo datasets were generated by scaling the expected signal yield from 0.5 to 2 times the nominal value. The extracted signal yields were fitted with a linear polynomial to verify the linear dependence. The result is shown in Fig. 7.2. The obtained fit has slope of 1 and it intercepts the y-axis at 0, which is consistent with an unbiased fit.

To further verify the fitting procedure, a fit to so-called Asimov data was performed. The Asimov dataset is a simulated dataset where the expected values perfectly agree with the MC predictions. In the case of a correct fitting procedure, the fit to the Asimov data exactly corresponds to the MC expectation. The results of the Asimov fit are shown in Fig. 7.3. It can be seen that the fit result agrees with the Asimov dataset in each bin. Additionally, the pull distribution is calculated, as shown in the bottom panel. The pulls are calculated for each bin as the difference between the data and the fitted prediction divided

(a) Electron channel   (b) Muon channel
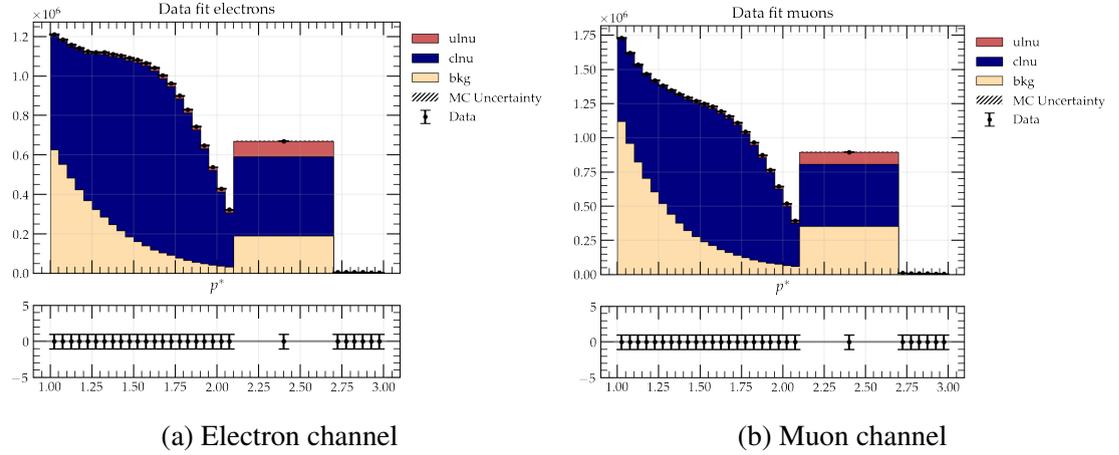
Figure 7.3: Post-fit lepton momentum distribution for electrons (left) and muons (right) for fits on Asimov dataset.

by the total error. It behaves as expected, having the value of zero for each bin.

7.3   Fit results

Finally, the above-described fitting procedure is applied on Belle II data. The data-MC comparison is shown before the fit and after the fit in Fig. 7.4 for electrons and Fig. 7.5 for muons. The uncertainties shown represent only the statistical MC uncertainties, as the fit procedure does not include any systematic uncertainties. Initially, there is a slight disagreement between data and MC in the pre-fit distribution due to imperfect modelling and reconstruction efficiencies in MC. The data and the fitted result show improved agreement, as can be seen in the pull distribution, where bin-wise pulls do not exceed 5σ, accounting only for the statistical uncertainty. The fitted yields are shown in Table 7.1.
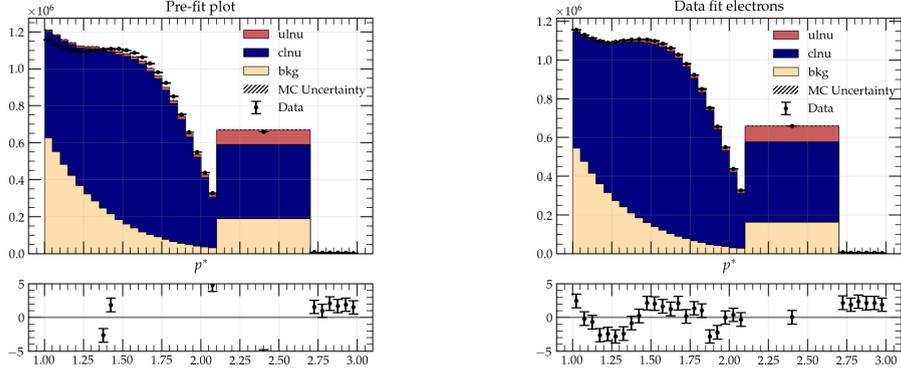
(a) Electron channel − BLNP model

(b) Electron channel − BLNP model
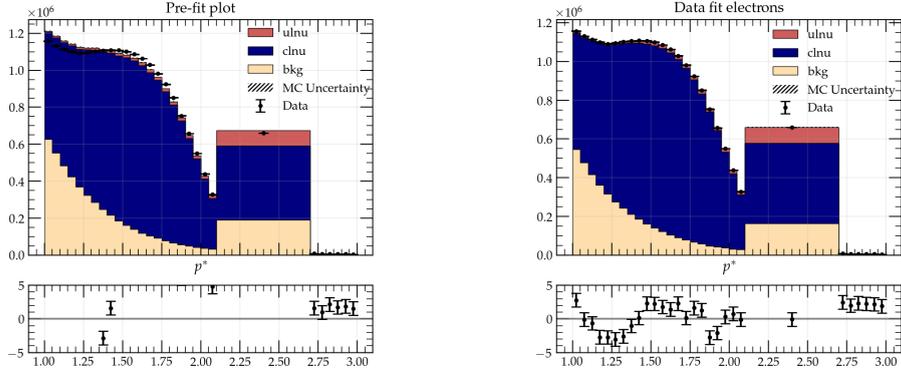
(c) Electron channel − DFN model

(d) Electron channel − DFN model

Figure 7.4: Pre-fit and post-fit electron momentum distributions, for fits on the Belle II dataset using the BLNP and DFN signal modelling. The bottom panel shows the fit pulls.

To calculate the partial branching fraction for $B \to X_u \ell \nu$ decay in the endpoint, the signal yield from the wide bin is needed. The fit result contains the total yields, as well as the bin-by-bin probabilities, as described in Eq. 7.1. From there, the number of signal events in the wide bin can be obtained as:

$$\nu^{b\to u}_{2.1 < p^*_\ell < 2.7 \text{ GeV}} = N_{b\to u} \cdot p^{b\to u}_{2.1 < p^*_\ell < 2.7 \text{ GeV}}, \tag{7.8}$$

where $N_{b\to u}$ is the total fitted signal yield and $p^{b\to u}_{2.1 < p^*_\ell < 2.7 \text{ GeV}}$ is the fitted fraction of the
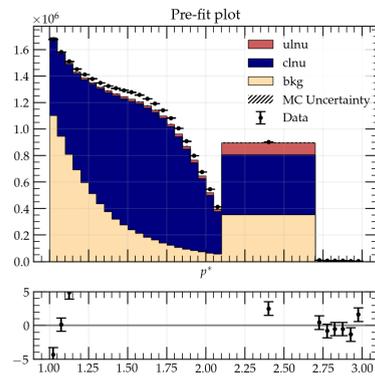
(a) Muon channel − BLNP model
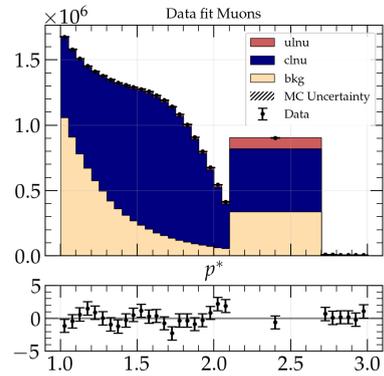
(b) Muon channel − BLNP model
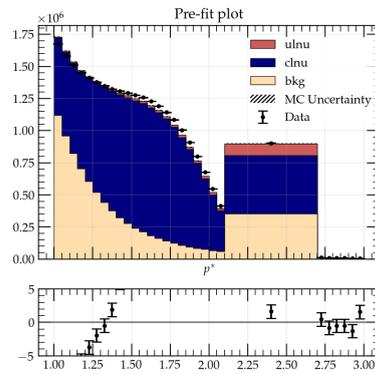
(c) Muon channel − DFN model
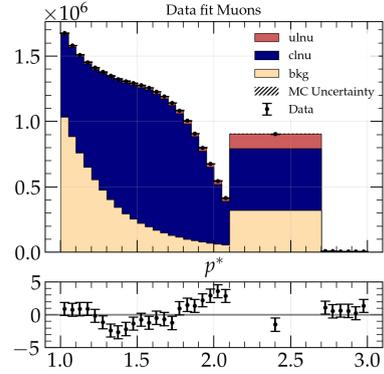
(d) Muon channel − DFN model

Figure 7.5: Pre-fit and post-fit muon momentum distributions, for fits on the Belle II dataset using the BLNP and DFN signal modelling. The bottom panel shows the fit pulls.

7.3 Fit results

Electrons	BLNP model		DFN model	
Template	fitted yield	pre-fit yield	fitted yield	pre-fit yield
background	4236217 ± 12379	4907685	4230522 ± 12165	4907685
$X_c\ell\nu$	16958618 ± 16418	16356704	16969793 ± 16111	16356704
$X_u\ell\nu$	390602 ± 9592	387195	385125 ± 9252	394432
Muons	BLNP model		DFN model	
Template	fitted yield	pre-fit yield	fitted yield	pre-fit yield
background	7476039 ± 13477	8117678	7470087 ± 13451	8117678
$X_c\ell\nu$	18784501 ± 17985	17968689	18801716 ± 17802	17968689
$X_u\ell\nu$	528336 ± 11200	420340	517046 ± 11058	428577

Table 7.1: The results of the fit showing the yields for different fitted yields. The pre-fit MC yields are given for comparison. The quoted uncertainty is the fit uncertainty that contains only the statistical uncertainty.

signal events that are found in the wide bin, with momentum $2.1 < p_\ell^* < 2.7 \text{ GeV}$.

To minimize the bias of the analyst, the signal yield in the endpoint is blinded, as described in Section 5.1.1. The statistical uncertainty of the endpoint signal yield is obtained from the fit. It is estimated by setting the background and $c\ell\nu$ template values to 0 in the wide bin and re-calculating the fit uncertainty of the signal yield. This procedure sets any correlation between different templates to zero and gives a purely statistical uncertainty of the signal yield in the wide bin. The statistical uncertainty of the signal yield in the wide bin is thus found to be 0.1%. Further, the systematic uncertainties of the endpoint signal yield are calculated, as described in the next chapter.

8

Systematic Uncertainties

In this chapter the sources of systematic uncertainties are explained, as well as the estimation of the systematic uncertainty of the endpoint signal yield. The signal yield in the endpoint is obtained as the signal yield of the wide bin of the fit, as described in Chapter 7. Since the aim of this analysis is to calculate the partial branching fraction in the endpoint region, the systematic uncertainties for the signal yield are not estimated in the full momentum range, but only in the endpoint, for the wide bin of the fit, between 2.1 and 2.7 GeV. To estimate the systematic uncertainties, the fit was repeated with varied MC weights, according to different systematic sources. The procedure was done separately for the electron and the muon channel, as well as for BLNP and DFN signal modelling.

8.1 Particle ID and fake rates

The particle ID efficiency and fake rate corrections are provided by the Belle II Physics Performance group, together with the correction uncertainties, as described in Section 6.3.5.

The systematic uncertainties from the particle ID and the fake rate corrections are separately evaluated. The procedure for obtaining these two systematic uncertainties is the same. Fifty repeated fits are performed with PID weights varied within their uncertainties. The weight variations are obtained through the `PIDVar` package [93]. The endpoint signal yield is obtained from each of these 50 fit results. The systematic uncertainty is obtained as the average difference between the nominal fit yield and the varied fit yields. For the electron channel, the obtained particle ID efficiency systematic is 8.6% (8.3%) and the fake rate systematic is 0.22% (0.23%) for the BLNP (DFN) model. For the muon channel, the obtained particle ID efficiency systematic is 5.27% (5.09%) and the fake rate systematic is 0.72% (0.72%) for the BLNP (DFN) model. The particle ID efficiency systematic is the dominant systematic uncertainty in both channels. The uncertainties on the particle ID efficiency corrections are illustrated in Fig. 6.9. These uncertainties are expected to get smaller as Belle II collects more data.

8.2 Tracking efficiency

The Belle II Physics Performance group provides the per-track uncertainty of 0.3% for the tracking efficiency. In this analysis, only one track (the electron or the muon candidate) per event is selected. For each channel, the fit is performed with the weight of $(1 + 0.003)$ and with $(1 - 0.003)$ to simulate for the tracking uncertainty. The systematic uncertainty was estimated as the average difference between the nominal fit endpoint yield and the up and down variations. The obtained systematic uncertainty is 0.25% for the electron channel and 0.32% for the muon channel, for the BLNP (DFN) model.

8.3 $B \to X_c \ell \nu$ branching fraction uncertainties

In MC, the $B \to X_c \ell \nu$ decays are generated according to their PDG [8] averaged branching fractions. These are the averaged branching fractions obtained from the previous experimental measurements and they are provided with an uncertainty. Additionally, a discrepancy in the measured values is present between the inclusive and exclusive measurements. The branching fraction obtained from the inclusive measurements is higher than the sum of the measured exclusive branching fraction. In MC, this gap is filled using the $B \to D^{(*)} \pi \pi \ell \nu$, $B \to D_s^{(*)} K \ell \nu$ and $B \to D^{(*)} \eta \ell \nu$ decays. The $B \to D^{(*)} \eta \ell \nu$ decays are artificially added, since they have not been previously measured, and they are assigned a 100% uncertainty. All the branching fractions and their uncertainties are given in Table 8.1. The fit is performed with the $B \to X_c \ell \nu$ weights varied up and down according to their branching fraction uncertainties, for each final state separately. The systematic uncertainty is obtained as the average difference between the nominal fit endpoint yield and the up and down variations.

The values of the systematic uncertainties are given in Table 8.2 for the electron and muon channels. The systematic uncertainty from the $D^{**} \ell \nu$ decays (where D^{**} can be D_0^*, D_1, D_1', $D_2^{(*)}$) are high due to the large uncertainties on their branching fractions. These systematic uncertainties can be decreased by having more precise measurements of these decay channels. Also, better understanding of the discrepancy between the inclusive and exclusive $B \to X_c \ell \nu$ branching fractions will eliminate the need for the artificially added $B \to D^{(*)} \eta \ell \nu$ decay channel with 100% assigned uncertainty.

$\mathcal{B}(B \to X_c \ell \nu)$	Value B^+	Value B^0
$B \to D\ell^+\nu_\ell$	$(2.5 \pm 0.1) \times 10^{-2}$	$(2.3 \pm 0.1) \times 10^{-2}$
$B \to D^*\ell^+\nu_\ell$	$(5.4 \pm 0.1) \times 10^{-2}$	$(5.1 \pm 0.1) \times 10^{-2}$
$B \to D_0^*(\to D\pi)\ell^+\nu_\ell$	$(4.2 \pm 0.8) \times 10^{-3}$	$(3.9 \pm 0.7) \times 10^{-3}$
$B \to D_1^*(\to D^*\pi)\ell^+\nu_\ell$	$(4.2 \pm 0.8) \times 10^{-3}$	$(3.9 \pm 0.8) \times 10^{-3}$
$B \to D_1(\to D^*\pi)\ell^+\nu_\ell$	$(4.2 \pm 0.3) \times 10^{-3}$	$(3.9 \pm 0.3) \times 10^{-3}$
$B \to D_2^*(\to D^*\pi)\ell^+\nu_\ell$	$(1.2 \pm 0.1) \times 10^{-3}$	$(1.1 \pm 0.1) \times 10^{-3}$
$B \to D_2^*(\to D\pi)\ell^+\nu_\ell$	$(1.8 \pm 0.2) \times 10^{-3}$	$(1.7 \pm 0.2) \times 10^{-3}$
$B \to D_1(\to D\pi\pi)\ell^+\nu_\ell$	$(2.4 \pm 1.0) \times 10^{-3}$	$(2.3 \pm 0.9) \times 10^{-3}$
$B \to D\pi\pi\ell^+\nu_\ell$	$(0.6 \pm 0.6) \times 10^{-3}$	$(0.6 \pm 0.6) \times 10^{-3}$
$B \to D^*\pi\pi\ell^+\nu_\ell$	$(2.2 \pm 1.0) \times 10^{-3}$	$(2.0 \pm 1.0) \times 10^{-3}$
$B \to D\eta\ell^+\nu_\ell$	$(4.0 \pm 4.0) \times 10^{-3}$	$(4.0 \pm 4.0) \times 10^{-3}$
$B \to D^*\eta\ell^+\nu_\ell$	$(4.0 \pm 4.0) \times 10^{-3}$	$(4.0 \pm 4.0) \times 10^{-3}$
$B \to X_c\ell^+\nu_\ell$	$(10.8 \pm 0.4) \times 10^{-2}$	$(10.1 \pm 0.4) \times 10^{-2}$

Table 8.1: Branching fractions for $B \to X_c \ell \nu$ background processes that were used. Values are taken from [8].

8.4 $B \to X_u \ell \nu$ branching fraction uncertainties

The $B \to X_u \ell \nu$ branching fraction systematics are estimated in the same manner as the $B \to X_c \ell \nu$ branching fraction systematics. The $B \to X_u \ell \nu$ branching fractions and their uncertainties were already given in Table 5.3. The systematic uncertainties are given in Table 8.2.

8.5 Hybrid modelling uncertainty

The systematic uncertainty due to the hybrid modelling is estimated as the difference between the BLNP and DFN models. The difference in the lepton momentum spectrum shape between these two models is illustrated in Fig. 8.1. The systematic uncertainty value is obtained as the difference between the nominal fitted endpoint yields with the DFN and the

BLNP model. The value is shown in Table 8.2.
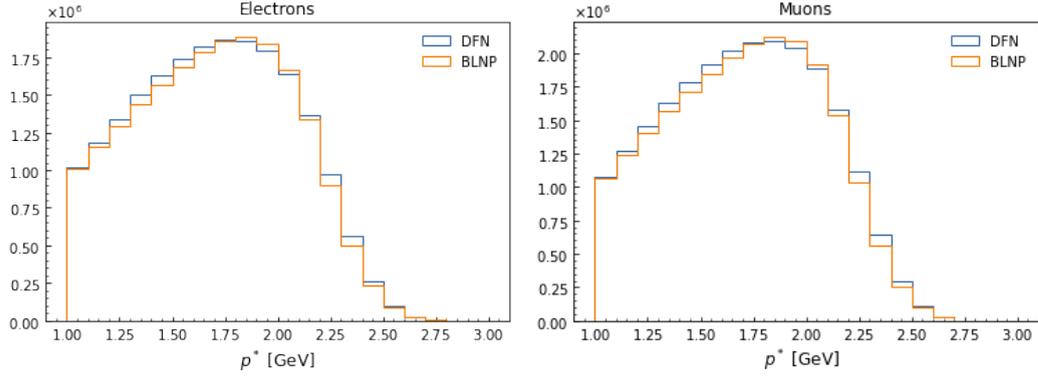
Figure 8.1: The comparison of BLNP and DFN non-resonant $B \rightarrow X_u \ell \nu$ models' momentum distributions for electrons(left) and muons(right). The vertical axis is shown with arbitrary scaling.

8.6 Uncertainty due to limited statistics of the MC sample

The MC sample was generated to correspond to the luminosity of 100 fb^{-1}, as described in Chapter 5. To compare the MC predictions to data, the MC sample needed to be reweighted to match the data luminosity of 189 fb^{-1}. This reweighting introduces a systematic uncertainty that is dependent on the size of the MC sample. This systematic uncertainty was estimated by performing repeated fits using 50 toy-MC samples. The original MC dataset was resampled by selecting a random number of events within the Poissonian statistical uncertainty of $\sqrt{N}$ and allowing for replacement. The estimated systematic uncertainties for electrons and muons, for the BLNP and the DFN signal modelling, are given in Table 8.2. While a larger MC sample would reduce this uncertainty, it proved to be too computationally expensive to perform the whole analysis with the larger MC sample. Since this systematic uncertainty is not a dominant one, it was decided to be more beneficial to keep

the smaller MC sample.

8.7 Uncertainty due to limited statistics of the off-resonance sample

The recorded off-resonance sample is much smaller than the recorded on-resonance sample, with the luminosity of 18 fb^{-1}. The off-resonance sample was reweighted to match the on-resonance luminosity, as described in Section 6.3. The systematic uncertainty due to the limited off-resonance sample statistics is estimated similarly to the MC statistics described in the previous section. The off-resonance data was randomly resampled with replacement, creating 50 new toy samples. The fit was repeated for each of the toy samples. The systematic uncertainty was estimated as the average difference between the nominal signal yield in the endpoint and the toy yields. The estimated systematic uncertainties for electrons and muons, for the BLNP and the DFN signal modelling, are given in Table 8.2.

8.8 Total systematic uncertainty on the signal yield

Looking at Table 8.2, it is evident that the largest contribution to the systematic uncertainty comes from the PID efficiency, especially for electrons. This is due to the fact that the PID at Belle II is not yet well optimized. The improvements to PID software are expected in the near future. These improvements will fix the issue where some of the subdetectors (SVD and TOP for electrons, SVD for muons) are not used for lepton identification. Also, at the moment, some of the phase space regions have large PID correction uncertainties due to the limited statistics of the data used in the calculation of these corrections.

Systematics source	BLNP		DFN	
	e^-	μ^-	e^-	μ^-
PID efficiency	8.6	4.1	8.3	3.9
PID fake rate	0.22	0.75	0.23	0.76
Tracking	0.25	0.32	0.25	0.32
Hybrid modelling	0.67	0.057	0.67	0.057
$\pi^0\ell^-\bar{\nu}$ BF	0.00029	0.0041	0.0014	0.0046
$\pi^+\ell^-\bar{\nu}$ BF	6.8e-05	0.0082	0.0028	0.0093
$\rho^0\ell^-\bar{\nu}$ BF	0.0067	0.026	0.014	0.028
$\rho^+\ell^-\bar{\nu}$ BF	0.014	0.051	0.027	0.055
$\eta\ell^-\bar{\nu}$ BF	0.00027	0.0027	0.0017	0.0028
$\eta'\ell^-\bar{\nu}$ BF	0.07	0.25	0.13	0.27
$\omega\ell^-\bar{\nu}$ BF	0.0072	0.027	0.014	0.029
$X_u^+\ell^-\bar{\nu}$ BF	0.028	0.13	0.066	0.14
$X_u^0\ell^-\bar{\nu}$ BF	0.031	0.13	0.063	0.14
$D\ell\nu$ BF	1.3	1.2	1.2	1.2
$D^*\ell\nu$ BF	3.6	3.2	3.6	3.2
D_1^0	4.5	3.9	4.4	3.8
D_1^+	3.8	3.3	3.7	3.3
D_0^{*0}	2.4	2.1	2.3	2.1
D_0^{*+}	2.1	1.8	2.0	1.8
$D_1'^0$	2.5	2.1	2.4	2.1
$D_1'^+$	2.1	1.8	2.0	1.8
D_2^{*0}	1.5	1.3	1.4	1.3
D_2^{*+}	1.4	1.2	1.4	1.2
$D^0\eta$	4.8	4.3	4.7	4.2
$D^+\eta$	4.8	3.3	4.7	4.3
$D^{*0}\eta$	4.5	3.9	4.4	4.8
$D^{*+}\eta$	4.5	4.0	4.4	4.0
$D\pi\pi$	0.28	0.25	0.28	0.77
$D^*\pi\pi$	2.2	1.9	2.2	1.9
D_sK	0.33	0.28	0.32	0.28
D_s^*K	0.43	0.37	0.42	0.36
MC statistics	0.33	1.3	0.4	1.4
Off-resonance statistics	0.14	0.48	0.14	0.46

Table 8.2: The systematic uncertainties from different sources, estimated for the electron and the muon channel, for BLNP and DFN signal models. All the uncertainties are cited as percent uncertainties.

8.8 Total systematic uncertainty on the signal yield

Another significant source of systematic uncertainty comes from the D^{**} decays, namely $B \to D\eta\ell\nu$ and $B \to D^*\eta\ell\nu$, which have 100% branching fraction uncertainty. Theoretical and experimental input is needed for better understanding of the discrepancy between the inclusive and exclusive $B \to X_c\ell\nu$ branching fraction measurements.

The total signal yield relative systematic uncertainty in the region between 2.1 and 2.7 GeV is 15.5% for the electron channel using the BLNP signal modelling and 15.1% using the DFN signal modelling. For the muon channel, the relative systematic uncertainty is 11.9% using the BLNP signal modelling and 12.4% using the DFN signal modelling. The next chapter explains how the partial branching fraction and the $|V_{ub}|$ value are calculated, and how these systematic uncertainties affect those measurements.

9

Branching Fraction Calculation and V_{ub} matrix element extraction

9.1 Signal selection efficiency

In order to calculate the partial branching fraction from the measured signal yield, the signal selection efficiency is needed. It can be estimated based on the MC simulation of the signal decays. The signal selection efficiency is calculated separately for the two signal models, for electrons and muons. The calculation is based on the true and reconstructed MC information. The partial signal selection efficiency, for a given momentum range Δp, is found as:

$$\epsilon(\Delta p) = \frac{N_{u\ell\nu}^{\mathrm{reco}}(\Delta p)}{N_{u\ell\nu}^{\mathrm{true}}(\Delta p)},\tag{9.1}$$

where $N_{u\ell\nu}^{\mathrm{reco}}$ is the number of signal events in MC after all the selections are applied; $N_{u\ell\nu}^{\mathrm{true}}$ is the number of signal events at the generator level, so-called true MC events.

	BLNP	
Momentum range	e^-	μ^-
$[1.0 - 3.0]$ GeV	0.43510 ± 0.00024	0.47128 ± 0.00026
$[2.1 - 2.7]$ GeV	0.5085 ± 0.0005	0.5469 ± 0.0005
	DFN	
Momentum range	e^-	μ^-
$[1.0 - 3.0]$ GeV	0.44317 ± 0.00025	0.47994 ± 0.00026
$[2.1 - 2.7]$ GeV	0.5149 ± 0.0005	0.5528 ± 0.0005

Table 9.1: The signal selection efficiency over the full lepton momentum range and in the momentum endpoint, given for the electron and muon channel, for BLNP and DFN signal modelling. The quoted uncertainties are the MC statistical uncertainties.

The signal selection efficiency was calculated for each momentum bin used in the fit, for the electron and muon channel, for both the BLNP and DFN signal models. The calculated total efficiencies and the efficiencies for the endpoint wide bin are shown in Table 9.1. Figure 9.1 shows how the efficiencies vary bin-by-bin. The uncertainties on the efficiency are estimated by taking into account the MC statistical uncertainties.

9.2 $B \to X_u \ell \nu$ partial branching fraction measurement

The partial branching fraction (BF) of the $B \to X_u \ell \nu$ decay is calculated as:

$$\Delta \mathcal{B}(\Delta p) = \frac{N_{u\ell\nu}(\Delta p)}{2 N_{B\bar{B}} \epsilon(\Delta p)}, \tag{9.2}$$

where $N_{u\ell\nu}(\Delta p)$ is the signal yield in the given momentum region, the $N_{B\bar{B}}$ is the number of $B\bar{B}$ pairs in the data sample, and the $\epsilon(\Delta p)$ is the signal selection efficiency. The partial branching fraction is calculated in the momentum region $\Delta p = [2.1, 2.7]$ GeV. The number of $B\bar{B}$ pairs in the data sample, $N_{B\bar{B}}$, for the sample used for this analysis is measured by
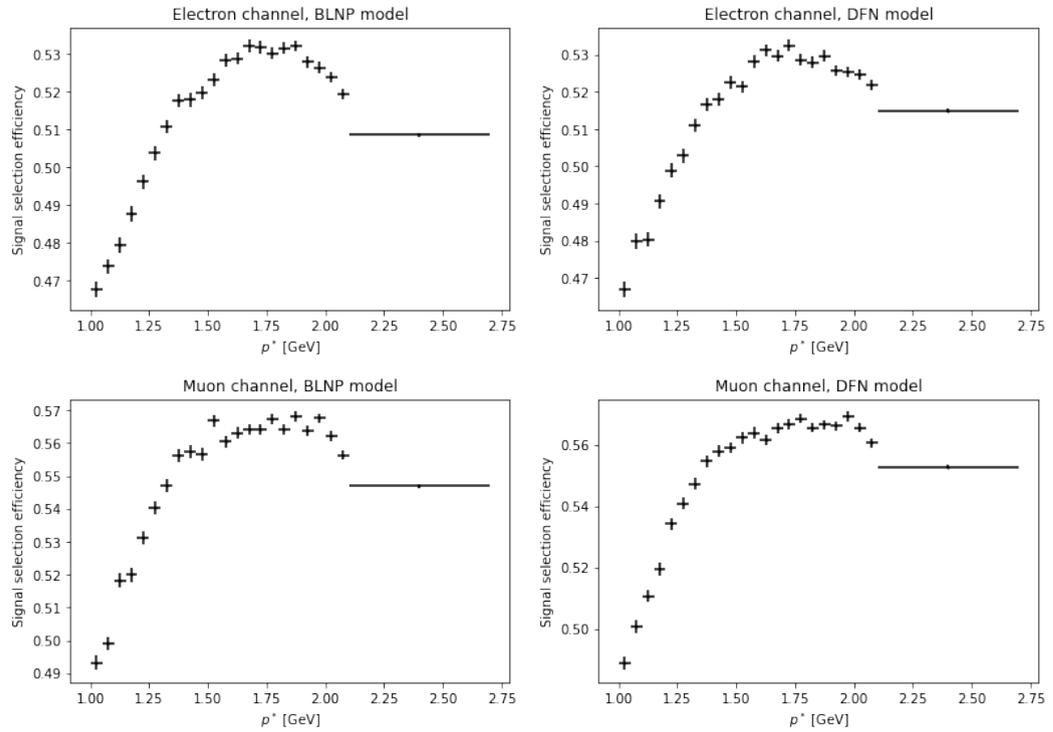
Figure 9.1: The signal selection efficiency shown for each bin of the fit. The horizontal error bars indicate the bin width, while the vertical errors are calculated based on the MC statistical uncertainties. The top row shows the electron channel, while the bottom row shows the muon channel. The left hand plots are using the BLNP signal modelling, while the right hand plots use the DFN signal modelling.

the Belle II collaboration and it is estimated to be:

$$N_{B\bar{B}} = (198.0 \pm 3.0_{(stat.+syst.)}) \times 10^6.$$ (9.3)

The factor of 2 in the denominator in Eq. 9.2 is used to account for the fact that each $\Upsilon(4S)$ decay event has two B mesons, thus doubling the probability of observing the desired B meson decay.

The calculation of the partial branching fraction is done separately for the electron and muon final states. For both, the partial branching fraction is estimated based on the BLNP and the DFN signal models. The signal yield, $N_{u\ell\nu}(\Delta p)$, is obtained from the fit result in the wide momentum bin, as described in Chapter 7.

The result for the partial branching fraction is presented blinded in this thesis, as outlined in Chapter 5. This was chosen in order to avoid showing an unofficial result before it has passed all the stages of collaboration approval. The relative statistical and systematic uncertainties are shown. These experimental uncertainties are presented in Table 9.2. The uncertainties shown are the final ones, obtained using the Belle II data. While this does not allow us to compare our result with the previous measurements, it is still possible to estimate how competitive the result will be by comparing the uncertainties.

The estimation of the theoretical uncertainties based on the BLNP and DFN signal models is outside the scope of this thesis. This estimation requires input from the authors of these theoretical models. Namely, the uncertainties coming from the shape function (SF) (see Section 2.4.2) contribute to the total partial branching fraction uncertainty. The BaBar publication [13] used the same theoretical models and estimated the theoretical uncertainty contributions in the wide electron momentum bin, between 2.1 and 2.7 GeV. Given the use

of the same models and the same lepton momentum range, it is expected that the theoretical uncertainties due to shape function modelling are the same for this analysis as for the BaBar analysis. Here, those theoretical uncertainties quoted by BaBar are cited, with the assumption that the theoretical uncertainties for the muon channel are the same as the ones for the electron channel. The blinded partial branching fractions with the estimated uncertainties are:

$$\Delta \mathcal{B}(B \to X_u e \nu)_{\text{BLNP model}} = XXX \pm 0.14\%_{\text{stat}} \pm 15.5\%_{\text{syst}} \pm {}^{+3.5}_{-3.1}\%_{\text{SF}}, \tag{9.4}$$

$$\Delta \mathcal{B}(B \to X_u e \nu)_{\text{DFN model}} = XXX \pm 0.13\%_{\text{stat}} \pm 15.2\%_{\text{syst}} \pm {}^{+2.7}_{-2.7}\%_{\text{SF}}, \tag{9.5}$$

$$\Delta \mathcal{B}(B \to X_u \mu \nu)_{\text{BLNP model}} = XXX \pm 0.16\%_{\text{stat}} \pm 12.0\%_{\text{syst}} \pm {}^{+3.5}_{-3.1}\%_{\text{SF}}, \tag{9.6}$$

$$\Delta \mathcal{B}(B \to X_u \mu \nu)_{\text{DFN model}} = XXX \pm 0.16\%_{\text{stat}} \pm 12.5\%_{\text{syst}} \pm {}^{+2.7}_{-2.7}\%_{\text{SF}}. \tag{9.7}$$

These values confirm that the partial branching fraction measurements in both the electron and muon channel are limited by systematic uncertainty, as expected. As can be seen in Table 9.2, the systematic uncertainties are dominated by the PID efficiency uncertainties, especially in the electron channel. The next most significant source of systematics is the uncertainty from the $B \to D^{**} \ell \nu$ branching fractions.

9.2.1 Comparison to previous measurements of the partial branching fraction

The estimated relative uncertainties on the partial branching fraction are directly comparable to the most recent BaBar result [13], which assessed the same region of the lepton momentum, between 2.1 and 2.7 GeV. The total experimental uncertainty for the BaBar re-

sult for electrons, using the DFN model, was 5.45%. This quoted uncertainty combines the statistical and systematic uncertainty of the measurement. At the moment, our result, with 15.5% relative experimental uncertainty for the DFN model for electrons, is not directly competitive with the BaBar result. The PID selection efficiency dominates the systematic uncertainty. This PID systematic uncertainty is expected to decrease with the better understanding of the Belle II PID system. As Belle II gathers more data, the PID corrections will become more precise, resulting in a lower PID uncertainty. The BaBar result cites the PID systematic uncertainty at 0.5%. It is expected that, during its full operation, Belle II will have similar PID systematic uncertainty.

For the BLNP model, the BaBar measurement quotes 5.54% relative experimental uncertainty. Again, this value is much lower than our relative uncertainty for electron using the BLNP signal model, which is at 15.2%. It should be emphasized that the quoted BaBar results came 17 years and 9 years after that experiment commenced and concluded data-taking, respectively. By this point, the BaBar experiment was at a mature stage, with the detector systematics being well understood. In contrast, the Belle II experiment is currently in a period of relative infancy.

The previously quoted BaBar measurement does not include muon final states; our result is therefore unique among the B-factory measurements. Similar to the electron channel, the muon channel relative uncertainties have a large contribution from the PID systematics. Again, these values are expected to improve as Belle II gathers more data.

Direct comparisons to other measurements of the partial branching fraction are not possible due to differences in the selected phase space region. The BaBar measurement [12] chose a similar electron momentum region, between 2.0 and 2.6 GeV. In this region, their cited statistical uncertainty is 7.2% and the systematic uncertainty is 11.3%. The dataset used

Systematics source	BLNP		DFN	
	e^-	μ^-	e^-	μ^-
PID efficiency	8.6	4.1	8.3	3.9
PID fake rate	0.22	0.75	0.23	0.76
Tracking	0.25	0.32	0.25	0.32
Hybrid modelling	0.67	0.057	0.67	0.057
$\pi^0\ell^-\bar{\nu}$ BF	0.00029	0.0041	0.0014	0.0046
$\pi^+\ell^-\bar{\nu}$ BF	6.8e-05	0.0082	0.0028	0.0093
$\rho^0\ell^-\bar{\nu}$ BF	0.0067	0.026	0.014	0.028
$\rho^+\ell^-\bar{\nu}$ BF	0.014	0.051	0.027	0.055
$\eta\ell^-\bar{\nu}$ BF	0.00027	0.0027	0.0017	0.0028
$\eta'\ell^-\bar{\nu}$ BF	0.07	0.25	0.13	0.27
$\omega\ell^-\bar{\nu}$ BF	0.0072	0.027	0.014	0.029
$X_u^+\ell^-\bar{\nu}$ BF	0.028	0.13	0.066	0.14
$X_u^0\ell^-\bar{\nu}$ BF	0.031	0.13	0.063	0.14
$D\ell\nu$ BF	1.3	1.2	1.2	1.2
$D^*\ell\nu$ BF	3.6	3.2	3.6	3.2
$D_1^0\ell\nu$	4.5	3.9	4.4	3.8
$D_1^+\ell\nu$	3.8	3.3	3.7	3.3
$D_0^{*0}\ell\nu$	2.4	2.1	2.3	2.1
$D_0^{*+}\ell\nu$	2.1	1.8	2.0	1.8
$D_1'^0\ell\nu$	2.5	2.1	2.4	2.1
$D_1'^+\ell\nu$	2.1	1.8	2.0	1.8
$D_2^{*0}\ell\nu$	1.5	1.3	1.4	1.3
$D_2^{*+}\ell\nu$	1.4	1.2	1.4	1.2
$D^0\eta\ell\nu$	4.8	4.3	4.7	4.2
$D^+\eta\ell\nu$	4.8	3.3	4.7	4.3
$D^{*0}\eta\ell\nu$	4.5	3.9	4.4	4.8
$D^{*+}\eta\ell\nu$	4.5	4.0	4.4	4.0
$D\pi\pi\ell\nu$	0.28	0.25	0.28	0.77
$D^*\pi\pi\ell\nu$	2.2	1.9	2.2	1.9
$D_sK\ell\nu$	0.33	0.28	0.32	0.28
$D_s^*K\ell\nu$	0.43	0.37	0.42	0.36
MC statistics	0.33	1.3	0.4	1.4
Off-resonance statistics	0.14	0.48	0.14	0.46
$N_{B\bar{B}}$	1.5	1.5	1.5	1.5
$\epsilon(\Delta p)$	0.1	0.1	0.1	0.1
Statistics	0.1	0.13	0.08	0.13
Total	15.5	12.0	15.2	12.5

Table 9.2: The relative experimental uncertainties on the $\Delta\mathcal{B}(B \to X_u\ell\nu)$ in the momentum range $[2.1, 2.7]$ GeV from different sources, estimated for the electron and the muon channel, for BLNP and DFN signal models. All the uncertainties are cited as percent uncertainties.

for this measurement contains $88 \times 10^6 B\bar{B}$ pairs from the early operation of the BaBar experiment. This result has a large statistical uncertainty due to using a smaller dataset. This result only includes the electron channel.

The BaBar result where the missing neutrino momentum was reconstructed [49], as summarized in Chapter 3, quoted the partial branching fraction in the electron energy region between 2.0 and 3.5 GeV. This result uses the same data set as the previously quoted measurement [12]. The statistical uncertainty on the partial branching fraction is 9.3% and the systematic uncertainty is 9.6%. The high statistical and systematic uncertainties are due to the additional missing neutrino reconstruction.

The most recent tagged inclusive measurement from Belle [50] gives the partial branching fraction in the lepton energy region above 1.0 GeV, utilizing the full data sample recorded at Belle of 711 fb^{-1}. This measurement includes the largest phase space region of all previous measurements. The statistical uncertainty is 7.2% and the systematic uncertainty is 11.3%. The tagged approach has low signal selection efficiency, leading to high statistical uncertainty, as well as introducing additional systematics.

9.3   Extrapolation of the total branching fraction

From the measured partial branching fraction, it is possible to obtain the total branching fraction of the $B \to X_u \ell \nu$ decay. If the partial branching fraction is measured in the momentum range Δp, the total branching fraction is obtained as:

$$\mathcal{B}(B \to X_u \ell \nu) = \frac{\Delta \mathcal{B}(\Delta p)}{f_u(\Delta p)}.$$

(9.8)

Here, $f_u(\Delta p)$ is the theoretically predicted fraction of leptons from the $B \to X_u \ell \nu$ decay that have the momentum in the range Δp. In this work, the values for f_u are obtained from the truth information of signal MC, as these MC samples were generated according to the DFN and BLNP models:

$$f_u(\Delta p) = \frac{N^{\text{gen}}_{B \to X_u \ell \nu}(\Delta p)}{N^{\text{gen}}_{B \to X_u \ell \nu}},$$ (9.9)

where $N^{\text{gen}}_{B \to X_u \ell \nu}$ is the total number of generated $B \to X_u \ell \nu$ events for the given model, and $N^{\text{gen}}_{B \to X_u \ell \nu}(\Delta p)$ is the number of generated $B \to X_u \ell \nu$ events that have leptons with momentum in the range of Δp. Given that the partial branching fraction is blinded, the total branching fraction will not be presented in this thesis.

9.4 $|V_{ub}|$ **calculation**

The value of the V_{ub} CKM matrix element can be calculated, based on the measured partial branching fraction. The $|V_{ub}|$ quantity is calculated as:

$$|V_{ub}| = \sqrt{\frac{\Delta \mathcal{B}(\Delta p)}{\tau_b \Delta \zeta(\Delta p)}}.$$ (9.10)

Here, $\tau_b = (1.5673 \pm 0.0029) \times 10^{-12}$ s [8] is the mean lifetime of the $B^0/B^{\pm}$ mesons. $\Delta \zeta(\Delta p)$ is the theoretical prediction for partial decay rate normalized by $|V_{ub}|^2$:

$$\Delta \zeta(\Delta p) = \Gamma_{\text{theory}} \times \frac{f_u(\Delta p)}{|V_{ub}|^2},$$ (9.11)

where Γ_{theory} is the total predicted decay rate and $f_u(\Delta p)$ is the fraction of the predicted decays that have the lepton momentum in the range Δp, as given in Eq. 9.9. The $|V_{ub}|^2$ denominator in Eq. 9.11 divides out the $|V_{ub}|^2$ factor intrinsic to the predicted value of Γ_{theory}.

The necessary theoretical values for the partial and total branching fractions are provided by De Fazio and Neubert [32] for the DFN model and by Bosh, Lange, Neubert and Paz [35] for the BLNP model. Based on these models, the values for the predicted normalized partial decay rate, $\Delta\zeta(\Delta p)$, for the electron momentum range between 2.1 and 2.7 GeV, were estimated by BaBar[13]. Those values are reused here.

The experimental uncertainties from the partial branching fraction measurement are propagated to the $|V_{ub}|$ value. These experimental uncertainties are shown in Table 9.3. The theoretical uncertainties consist of the shape function uncertainty and the model uncertainties for $\Delta\zeta(\Delta p)$. These are cited as estimated by the BaBar collaboration [13]. The blinded $|V_{ub}|$ values quoted with the relative uncertainties are:

$$|V_{ub}|_{\text{BLNP model},\Delta\mathcal{B}(B\to X_u e\nu)} = XXX \pm 7.8\%_{\text{experiment}} \pm {}^{+5.0}_{-4.5}\%_{\text{SF}} \pm {}^{+7.5}_{-6.1}\%_{\text{theory}}, \quad (9.12)$$

$$|V_{ub}|_{\text{DFN model},\Delta\mathcal{B}(B\to X_u e\nu)} = XXX \pm 7.6\%_{\text{experiment}} \pm {}^{+7.7}_{-5.7}\%_{\text{SF}} \pm {}^{+4.5}_{-3.9}\%_{\text{theory}}, \quad (9.13)$$

$$|V_{ub}|_{\text{BLNP model},\Delta\mathcal{B}(B\to X_u \mu\nu)} = XXX \pm 6.0\%_{\text{experiment}} \pm {}^{+5.0}_{-4.5}\%_{\text{SF}} \pm {}^{+7.5}_{-6.1}\%_{\text{theory}}, \quad (9.14)$$

$$|V_{ub}|_{\text{DFN model},\Delta\mathcal{B}(B\to X_u \mu\nu)} = XXX \pm 6.3\%_{\text{experiment}} \pm {}^{+7.7}_{-5.7}\%_{\text{SF}} \pm {}^{+4.5}_{-3.9}\%_{\text{theory}}. \quad (9.15)$$

9.4.1 Comparison to previous $|V_{ub}|$ measurements

The estimated relative uncertainties of the $|V_{ub}|$ can be compared to the previous measurements. The BaBar's measurement [13] quotes the $|V_{ub}|$ values for the DFN and the BLNP signal models. The $|V_{ub}|$ estimation using these two models has the same relative experimental uncertainty of 2.8%. Similarly to the comparison of the partial branching fraction precisions, the measurement described in this thesis is dominated by the PID systematic

Systematics source	BLNP		DFN	
	e^-	μ^-	e^-	μ^-
PID efficiency	4.3	2.1	4.2	2.0
PID fake rate	0.11	0.38	0.11	0.38
Tracking	0.13	0.16	0.13	0.16
Hybrid modelling	0.34	0.028	0.34	0.028
$\pi^0\ell^-\bar{\nu}$ BF	0.00014	0.0021	0.00069	0.0023
$\pi^+\ell^-\bar{\nu}$ BF	3.4e-05	0.0041	0.0014	0.0047
$\rho^0\ell^-\bar{\nu}$ BF	0.0033	0.013	0.0069	0.014
$\rho^+\ell^-\bar{\nu}$ BF	0.0072	0.025	0.014	0.028
$\eta\ell^-\bar{\nu}$ BF	0.00013	0.0014	0.00083	0.0014
$\eta'\ell^-\bar{\nu}$ BF	0.035	0.12	0.067	0.14
$\omega\ell^-\bar{\nu}$ BF	0.0036	0.013	0.0069	0.015
$X_u^+\ell^-\bar{\nu}$ BF	0.014	0.065	0.033	0.072
$X_u^0\ell^-\bar{\nu}$ BF	0.015	0.064	0.032	0.071
$D\ell\nu$ BF	0.64	0.6	0.61	0.58
$D^*\ell\nu$ BF	1.8	1.6	1.8	1.6
$D_1^0\ell\nu$	2.2	1.9	2.2	1.9
$D_1^+\ell\nu$	1.9	1.7	1.9	1.6
$D_0^{*0}\ell\nu$	1.2	1.1	1.2	1.0
$D_0^{*+}\ell\nu$	1.0	0.92	1.0	0.91
$D_1'^0\ell\nu$	1.2	1.1	1.2	1.1
$D_1'^+\ell\nu$	1.0	0.92	1.0	0.91
$D_2^{*0}\ell\nu$	0.73	0.64	0.71	0.63
$D_2^{*+}\ell\nu$	0.69	0.61	0.68	0.6
$D^0\eta\ell\nu$	2.4	2.1	2.3	2.1
$D^+\eta\ell\nu$	2.4	1.7	2.3	2.2
$D^{*0}\eta\ell\nu$	2.2	2.0	2.2	2.4
$D^{*+}\eta\ell\nu$	2.2	2.0	2.2	2.0
$D\pi\pi\ell\nu$	0.14	0.13	0.14	0.38
$D^*\pi\pi\ell\nu$	1.1	0.97	1.1	0.97
$D_s K\ell\nu$	0.16	0.14	0.16	0.14
$D_s^* K\ell\nu$	0.21	0.18	0.21	0.18
MC statistics	0.17	0.66	0.2	0.68
Off-resonance statistics	0.072	0.24	0.072	0.23
$N_{B\bar{B}}$	0.8	0.8	0.8	0.8
$\epsilon(\Delta p)$	0.05	0.05	0.05	0.05
Statistics	0.05	0.07	0.05	0.07
Total	7.8	6.0	7.6	6.3

Table 9.3: The experimental uncertainties on the $|V_{ub}|$ from different sources, estimated for the electron and the muon channel, for BLNP and DFN signal models. All the uncertainties are cited as percent uncertainties.

uncertainties. Once the understanding of that source of systematic uncertainty is improved, it is expected that the Belle II measurement can be competitive with BaBar.

The BaBar measurement using the electron momentum region between 2.0 and 2.6 GeV [12] quotes the total experimental relative uncertainty on $|V_{ub}|$ at 5.6%. The BaBar measurement involving the missing neutrino reconstruction [49] quotes the relative experimental uncertainty of 6.6%. These measurements have slightly better precision than what is obtained in this thesis.

9.5 $B \to X\ell\nu$ cross-check

To validate the analysis procedure before the unblinding of the $B \to X_u\ell\nu$ BF and $|V_{ub}|$ value, the hadron-flavour agnostic $B \to X\ell\nu$ mode was used as a cross-check. The fitted yield of the $b \to u$ and $b \to c$ templates was added together to obtain the total $B \to X\ell\nu$ yield. Only the BLNP model for the $B \to X_u\ell\nu$ inclusive events was used for convenience. Based on the total yield, the partial branching fraction was calculated for each bin of the fit in the momentum region below the wide bin, between 1.0 and 2.1 GeV. The partial branching fraction for each bin was calculated as:

$$\Delta\mathcal{B}(\Delta p) = \frac{N_{b \to u} + N_{b \to c}}{2\epsilon(\Delta p)N_{B\bar{B}}}.$$

(9.16)

Here, $\epsilon(\Delta p)$ denotes the selection efficiency for $B \to X\ell\nu$ events for a given momentum bin. The results were obtained separately for the electron and muon final states. The systematic uncertainty was not estimated for this cross-check. Only the fit uncertainty, corresponding to the statistical uncertainty, the $N_{B\bar{B}}$ and the selection efficiency uncertainties
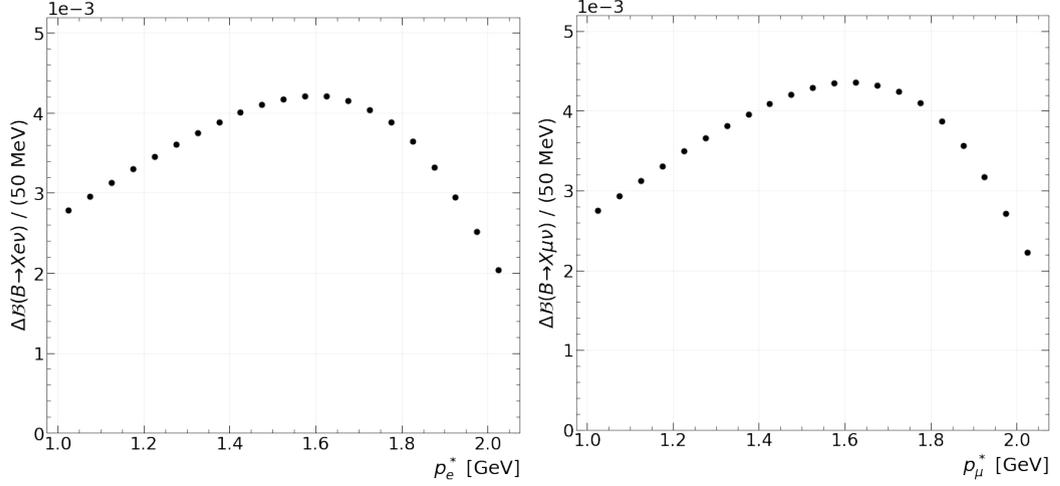
Figure 9.2: Partial BF as a function of the lepton momentum in the $\Upsilon(4S)$ rest frame for inclusive $B \to X\ell\nu$ decays, for the electron (left) and muon (right) final states.

were estimated. The results for the partial branching fractions for each bin are given in Fig. 9.2. The result of the partial branching fractions for the electron mode can be compared to BaBar's result [13], shown in Fig. 9.3. The BaBar result was obtained for the wider momentum region, between 0.8 and 2.8 GeV. The results show good agreement in the overlapping momentum region, with the values of the partial branching fraction and the shapes of the curves agreeing well between the two results.

Based on the partial BF between 1.0 and 2.1 GeV, the total $B \to X\ell\nu$ BF was estimated using Eq. 9.8. The obtained results are:

$$\mathcal{B}(Xe\nu) = (10.00 \pm 0.23)\% \tag{9.17}$$

$$\mathcal{B}(X\mu\nu) = (10.13 \pm 0.24)\%. \tag{9.18}$$
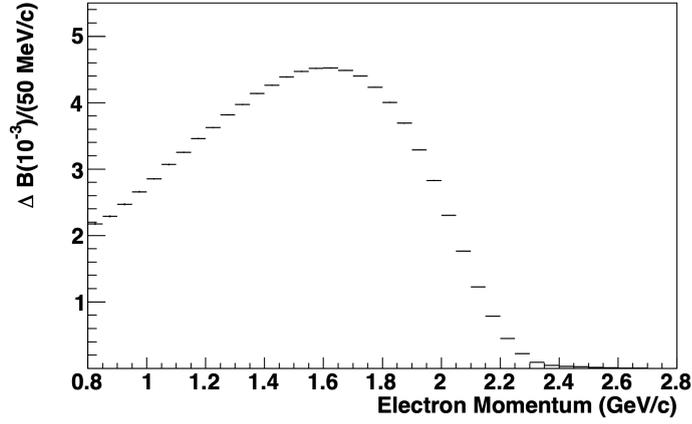
Figure 9.3: BaBar's result [13] for the partial BF as a function of the electron momentum in the $\Upsilon(4S)$ rest frame for inclusive $B \to Xe\nu$ decays.

Again, the uncertainties contain the statistical uncertainty obtained from the fit, the MC statistical uncertainty, the $N_{B\bar{B}}$ and the selection efficiency uncertainties. Given that the values are cited without the systematic uncertainties, they are comparable to the average $B \to X\ell\nu$ BF value from the PDG [8]:

$$\mathcal{B}(X\ell\nu) = (10.49 \pm 0.20)\% \quad \text{(PDG average)}. \tag{9.19}$$

The agreement between the results of this cross check with the previous measurements attests to the validity of the used analysis approach.

9.6 Unblinding strategy

Before the unblinding of the final $B \to X_u\ell\nu$ partial branching fraction and the $|V_{ub}|$ value, the analysis needs to follow the Belle II collaboration's unblinding guidelines, as described in Section 5.1.1. The analysis has been presented to the Belle II Semileptonic Working Group and has received conditional approval. The analysis has been documented

in an internal Belle II note. The next step is to present the analysis to the review committee, consisting of three senior members of the Belle II collaboration. Since this analysis is based on the MC14 campaign, it will be updated to the newer MC15 campaign that became recently available, before presenting to the review committee. The results presented here are based on 189 fb^{-1} of Belle II data. A larger dataset of 402 fb^{-1} on-resonance data is now available. The analysis will be updated to include the full available dataset. The update to the new MC sample and the full recorded dataset is expected to lower the statistical and systematic uncertainties, especially the dominant PID systematic uncertainty. Once the approval from the review committee is obtained, the results will be unblinded. The analysis then proceeds to the collaboration-wide review, where all the collaboration members are invited to provide their feedback and raise any concerns. Following the collaboration-wide review, the result becomes official and can be presented to the public at conferences and can be submitted for journal publication.

10

Conclusion

In this thesis, the first measurement of the inclusive $B \to X_u \ell \nu$ partial branching fraction with Belle II data was presented. The analysis utilized 189 fb^{-1} of Belle II data. An untagged analysis approach was used, whereby the companion B meson in the event is not reconstructed. This analysis approach enabled a high signal selection efficiency for the signal mode, which is suppressed in the Standard Model. The signal consisted of the charged electrons and muons, as was described in Chapter 6. To minimize the effect of the dominant $B \to X_c \ell \nu$ decay, the signal yield was extracted from the lepton momentum endpoint, between 2.1 and 2.7 GeV. The continuum backgrounds were suppressed using a multivariate training. The training was performed using the FastBDT model with 20 variables that provide discrimination between $B\bar{B}$ and continuum events.

A fitting algorithm was developed for the extraction of the signal yield. A binned fit was used, where the MC and off-resonance data provided the fit templates. Two different models were used to simulate the signal contributions, the BLNP and the DFN model. The

fit was performed on the lepton momentum, between 1.0 and 3.0 GeV. The regions between 1.0 and 2.1 GeV and 2.7 and 3.0 GeV were divided into 50 MeV bins. The data in the endpoint, between 2.1 and 2.7 GeV, was combined into a single wide bin. The signal yield for the partial branching fraction measurement was extracted from the wide endpoint bin. The theoretical predictions for the $B \to X_u \ell \nu$ decay in the lepton momentum endpoint are less reliable. Thus, by combining the data in the endpoint into a singe wide bin, the potential impact of the theoretical modelling on the systematic uncertainty was minimized.

The systematic uncertainties on signal yield were estimated for various sources. The dominant systematic uncertainty proved to be from the particle identification. As Belle II is a young experiment, the particle identification is still being optimized. This leads to the large particle identification uncertainties. These uncertainties are expected to decrease as the Belle II experiment collects more data.

Uniquely, unlike the previous inclusive untagged measurements performed at Belle and BaBar with electron channel only, this analysis is an inclusive untagged analysis that includes both electron and muon final states. The relative uncertainties on the partial branching fraction and the $|V_{ub}|$ values were presented for the electron and muon channel, separately for the BLNP and the DFN signal model. The estimated statistical uncertainty on the partial branching fraction of 0.14% for electrons and 0.16% for muons is competitive with the previous experiments. The systematic uncertainties on the partial branching fraction (15.5% for electrons and 12.5% for muons, for the BLNP model) are larger than those from the previous experiments due to the dominant particle identification uncertainty.

The analysis was presented blinded, as discussed in Chapter 5. The thesis demonstrated that the signal selection procedure is well optimized, as verified by the presented cut flow analysis (see Section 6.5). The fitting procedure was verified using the Asimov fits (see

Conclusion

Section 7.2). The dominant sources of systematic uncertainties were identified and their effect on the measurement was estimated (see Chapter 8). The branching fraction of the flavour-agnostic $B \to X\ell\nu$ decay was estimated in order to validate the analysis procedure. The result is in agreement with the current world average from the other measurements, as presented in the PDG (see Section 9.2). This demonstrates that the analysis is stable and ready for unblinding.

The next step in this analysis is to obtain the unblinding permission from the Belle II collaboration. The analysis will be upgraded to use the more recent MC simulation and to utilize the full Belle II on-resonance dataset of 402 fb^{-1}. It is expected that this will lead to the reduction of the systematic uncertainties coming from particle identification.

This analysis has been one of the first untagged analyses at the Belle II experiment. It has contributed to the validation of the `basf2` software for Belle II analyses. During its course, along with other similar analyses, it demonstrated the need for better simulation of the continuum events, as well as the need for the off-resonance data to replace the inadequately simulated continuum. As it examined a broad lepton momentum region, it demonstrated the need for more precise and uniform particle identification. These issues are under investigation by the Belle II collaboration and are expected to be resolved in the near future.

This analysis demonstrates the capability of the Belle II experiment to contribute to the longstanding puzzle of tensions between the inclusive and exclusive $|V_{ub}|$ measurements. Together with the exclusive semileptonic $B \to X_u\ell\nu$ analyses at Belle II, it will enable better understanding of the previously seen discrepancies between the inclusive and exclusive analysis approaches. With the planned dataset of 50 ab^{-1}, the Belle II experiment

is expected to improve the precision of the current V_{ub} measurements. This improvement will test the validity and consistency of the Standard Model and probe new frontiers in the search for the origin of the CP violation in the universe.

Bibliography

1. Belle II Collaboration. *Inclusive $B \rightarrow X_u e \nu_e$ Endpoint Analysis: Approved Plots for ICHEP2020* https://docs.belle2.org/record/2057. July 2020. (February 26, 2024).

2. Cheaib, R. & BELLE II collaboration. *Towards First Results on V_{ub} and V_{cb} with the Belle II Experiment* in *Proceedings of 40th International Conference on High Energy Physics — PoS(ICHEP2020)* (Sissa Medialab, Prague, virtual meeting, March 2021), 382.

3. Feynman, R. P. Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction. *Physical Review* **80,** 440–457 (November 1950).

4. Glashow, S. L. Partial-Symmetries of Weak Interactions. *Nuclear Physics* **22,** 579–588. ISSN: 0029-5582 (February 1961).

5. Salam, A. & Ward, J. C. Electromagnetic and Weak Interactions. *Physics Letters* **13,** 168–171. ISSN: 0031-9163 (November 1964).

6. Weinberg, S. A Model of Leptons. *Physical Review Letters* **19,** 1264–1266 (November 1967).

7. Aad, G. *et al.* Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC. *Physics Letters B* **716,** 1–29. ISSN: 0370-2693 (September 2012).

8. Particle Data Group *et al.* Review of Particle Physics. *Progress of Theoretical and Experimental Physics* **2022,** 083C01. ISSN: 2050-3911 (August 2022).

9. Fulton, R. *et al.* Observation of *B* -Meson Semileptonic Decays to Noncharmed Final States. *Physical Review Letters* **64,** 16–20. ISSN: 0031-9007 (January 1990).

10. Albrect, H. *et al.* Observation of Semileptonic Charmless B Meson Decays. *Physics Letters B* **234,** 409–416. ISSN: 0370-2693 (January 1990).

11. Dingfelder, J. & Mannel, T. Leptonic and Semileptonic Decays of B Mesons. *Rev. Mod. Phys.* **88,** 035008–035008 (2016).

12. BABAR Collaboration. Measurement of the Inclusive Electron Spectrum in Charmless Semileptonic B Decays near the Kinematic End Point and Determination of $|V_{ub}|$. *Physical Review D* **73,** 012006 (January 2006).

13. BABAR Collaboration. Measurement of the Inclusive Electron Spectrum from B Meson Decays and Determination of |Vub|. *Physical Review D* **95,** 072001. ISSN: 2470-0010, 2470-0029. arXiv: 1611.05624 (April 2017).

14. Martin, B. R. & Shaw, G. *Particle Physics* ISBN: 978-1-118-68166-4 (John Wiley & Sons, March 2013).

15. Burgess, C. & Moore, G. *The Standard Model: A Primer* ISBN: 978-0-521-86036-9 (Cambridge University Press, 2007).

16. Higgs, P. W. Broken Symmetries and the Masses of Gauge Bosons. *Physical Review Letters* **13,** 508–509 (October 1964).

17. Zwicky, F. Republication of: The Redshift of Extragalactic Nebulae. *General Relativity and Gravitation* **41,** 207–224. ISSN: 0001-7701 (January 2009).

18. Riess, A. G. *et al.* Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *The Astronomical Journal* **116,** 1009. ISSN: 1538-3881 (September 1998).

19. Perlmutter, S. *et al.* Measurements of Ω and Λ from 42 High-Redshift Supernovae. *The Astrophysical Journal* **517,** 565. ISSN: 0004-637X (June 1999).

20. Wikimedia Commons. *Standard Model of Elementary Particles* https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg. September 2019.

21. Super-Kamiokande Collaboration. Evidence for Oscillation of Atmospheric Neutrinos. *Physical Review Letters* **81,** 1562–1567 (August 1998).

22. SNO Collaboration. Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory. *Physical Review Letters* **89,** 011301 (June 2002).

23. Kobayashi, M. & Maskawa, T. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics* **49,** 652–657. ISSN: 0033-068X (February 1973).

24. Christenson, J. H., Cronin, J. W., Fitch, V. L. & Turlay, R. Evidence for the 2π Decay of the K_2^0 Meson. *Physical Review Letters* **13,** 138–140 (July 1964).

25. Chau, L.-L. & Keung, W.-Y. Comments on the Parametrization of the Kobayashi-Maskawa Matrix. *Physical Review Letters* **53,** 1802–1805 (November 1984).

26. Wolfenstein, L. Parametrization of the Kobayashi-Maskawa Matrix. *Physical Review Letters* **51,** 1945–1947 (November 1983).

27. Bevan, A. J. *et al.* The Physics of the B Factories. *The European Physical Journal C* **74,** 3026. ISSN: 1434-6044, 1434-6052 (November 2014).

28. Höcker, A. & Ligeti, Z. CP Violation and the CKM Matrix. *Annual Review of Nuclear and Particle Science* **56,** 501–567. ISSN: 0163-8998, 1545-4134 (November 2006).

29. Dingfelder, J. & Mannel, T. Leptonic and Semileptonic Decays of B Mesons. *Reviews of Modern Physics* **88,** 035008 (September 2016).

30. Eichten, E. & Hill, B. An Effective Field Theory for the Calculation of Matrix Elements Involving Heavy Quarks. *Physics Letters B* **234,** 511–516. ISSN: 0370-2693 (January 1990).

31. Uraltsev, N. Theoretical Uncertainties in Γsl (b → u). *International Journal of Modern Physics A* **14,** 4641–4652. ISSN: 0217-751X (November 1999).

32. De Fazio, F. & Neubert, M. $B \to X_u\ell\nu$ Decay Distributions to Order α_s. *Journal of High Energy Physics* **1999,** 017–017. ISSN: 1029-8479. arXiv: `hep-ph/9905351` (June 1999).

33. Lange, B. O., Neubert, M. & Paz, G. Theory of Charmless Inclusive B Decays and the Extraction of V$_{ub}$. *Physical Review D* **72,** 073006. ISSN: 1550-7998, 1550-2368. arXiv: `hep-ph/0504071` (October 2005).

34. Gambino, P., Giordano, P., Ossola, G. & Uraltsev, N. Inclusive Semileptonic B Decays and the Determination of |Vub|. *Journal of High Energy Physics* **2007,** 058. ISSN: 1126-6708 (October 2007).

35. Bosch, S. W., Lange, B. O., Neubert, M. & Paz, G. Factorization and Shape-Function Effects in Inclusive B-meson Decays. *Nuclear Physics B* **699,** 335–386. ISSN: 0550-3213 (November 2004).

36. Fermilab Lattice and MILC collaborations. V_{ub} from $B \to \pi\ell\nu$ Decays and $(2+1)$-Flavor Lattice QCD. *Physical Review D* **92,** 014024 (July 2015).

37. Berkelman, K. *A Personal History of CESR and CLEO: The Cornell Electron Storage Ring and Its Main Particle Detector Facility* ISBN: 978-981-238-697-7 (World Scientific, 2004).

38. Andrews, D. *et al.* Observation of a Fourth Upsilon State in e^+e^- Annihilations. *Physical Review Letters* **45,** 219–221 (July 1980).

39. Albrecht, H. *et al.* Argus: A Universal Detector at DORIS II. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **275,** 1–48. ISSN: 0168-9002 (February 1989).

40. Brodzicka, J. *et al.* Physics Achievements from the Belle Experiment. *Progress of Theoretical and Experimental Physics* **2012,** 04D001. ISSN: 2050-3911 (January 2012).

41. Aubert, B. *et al.* The BaBar Detector: Upgrades, Operation and Performance. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **729,** 615–701. ISSN: 0168-9002 (November 2013).

42. BABAR Collaboration. Measurement of the CP Asymmetry Amplitude $\sin 2\beta$ with B^0 Mesons. *Physical Review Letters* **89,** 201802 (October 2002).

43. Belle Collaboration. Improved Measurement of Mixing-Induced CP Violation in the Neutral B Meson System. *Physical Review D* **66,** 071102 (October 2002).

44. Belyaev, I., Carboni, G., Harnew, N., Matteuzzi, C. & Teubert, F. The History of LHCb. *The European Physical Journal H* **46,** 3. ISSN: 2102-6467 (March 2021).

45. LHCb Collaboration. First Observation of the Decay $B_s^0 \rightarrow K^- \mu^+ \nu_\mu$ and a Measurement of V_{ub}/V_{cb}. *Physical Review Letters* **126,** 081804 (February 2021).

46. Aaij, R. *et al.* Determination of the Quark Coupling Strength |Vub| Using Baryonic Decays. *Nature Physics* **11,** 743–747. ISSN: 1745-2481 (September 2015).

47. Bartelt, J. *et al.* Measurement of Charmless Semileptonic Decays of B Mesons. *Physical Review Letters* **71,** 4111–4115 (December 1993).

48. Gardi, E. *Inclusive Distributions near Kinematic Thresholds* June 2006. arXiv: `hep-ph/0606080`.

49. BABAR Collaboration. Determination of V_{ub} from Measurements of the Electron and Neutrino Momenta in Inclusive Semileptonic B Decays. *Physical Review Letters* **95,** 111801 (September 2005).

50. Belle Collaboration. Measurements of Partial Branching Fractions of Inclusive $B \rightarrow X_u \ell^+ \nu_\ell$ Decays with Hadronic Tagging. *Physical Review D* **104,** 012008 (July 2021).

51. Belle Collaboration. Measurement of Differential Branching Fractions of Inclusive $B \rightarrow X_u \ell^+ \nu_\ell$ Decays. *Physical Review Letters* **127,** 261801 (December 2021).

52. Heavy Flavor Averaging Group Collaboration. Averages of b-Hadron, c-Hadron, and τ-Lepton Properties as of 2021. *Physical Review D* **107,** 052008 (March 2023).

53. Athar, S. B. *et al.* Study of the q^2 Dependence of $B \rightarrow \pi \ell \nu$ and of $B \rightarrow \rho(\omega)\ell\nu$ and Extraction of V_{ub}. *Physical Review D* **68,** 072003 (October 2003).

54. CLEO Collaboration. Study of Exclusive Charmless Semileptonic B Decays and V$_{ub}$. *Physical Review Letters* **99,** 041802 (July 2007).

55. BABAR Collaboration. Branching Fraction and Form-Factor Shape Measurements of Exclusive Charmless Semileptonic B Decays, and Determination of V$_{ub}$. *Physical Review D* **86,** 092004 (November 2012).

56. Belle Collaboration. Measurement of the Decay $B^0 \rightarrow \pi^- \ell \nu$ and Determination of V$_{ub}$. *Physical Review D* **83,** 071101 (April 2011).

57. Belle Collaboration. Measurements of Branching Fractions and q^2 Distributions for $B \rightarrow \pi \ell \nu$ and $B \rightarrow \rho \ell \nu$ Decays with $B \rightarrow D^{(*)} \ell \nu$ Decay Tagging. *Physics Letters B* **648,** 139–148. ISSN: 0370-2693 (May 2007).

58. Belle Collaboration. Study of Exclusive $B \rightarrow X_u \ell \nu$ Decays and Extraction of V$_{ub}$ Using Full Reconstruction Tagging at the Belle Experiment. *Physical Review D* **88,** 032005 (August 2013).

59. BABAR Collaboration. Measurements of $B \rightarrow \{\pi, \eta, \eta^{'}\} \ell \nu_\ell$ Branching Fractions and Determination of V$_{ub}$ with Semileptonically Tagged B Mesons. *Physical Review Letters* **101,** 081801 (August 2008).

60. BABAR Collaboration. Measurement of the $B \rightarrow \pi \ell \nu$ Branching Fraction and Determination of V$_{ub}$ with Tagged B Mesons. *Physical Review Letters* **97,** 211801 (November 2006).

61. Akai, K., Furukawa, K. & Koiso, H. SuperKEKB Collider. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. Advances in Instrumentation and Experimental Methods (Spe-*

cial Issue in Honour of Kai Siegbahn) **907,** 188–199. ISSN: 0168-9002 (November 2018).

62. Abe, T. *et al. Belle II Technical Design Report* November 2010. arXiv: 1011.0352 [hep-ex, physics:physics].

63. Kou, E *et al.* The Belle II Physics Book. *Progress of Theoretical and Experimental Physics* **2019,** 123C01. ISSN: 2050-3911 (December 2019).

64. Belle II Collaboration. *Belle II Archives* https://www.belle2.org/archives/. (November 13, 2023).

65. Spruck, B. *et al. Belle II Pixel Detector Commissioning and Operational Experience* in *Proceedings of The 28th International Workshop on Vertex Detectors — PoS(Vertex2019)* **373** (SISSA Medialab, September 2020), 015.

66. Bilka, T. *Belle II Vertex Detector Performance* in *Proceedings of The 28th International Workshop on Vertex Detectors — PoS(Vertex2019)* (Sissa Medialab, Lopud, Croatia, February 2020), 001.

67. Dutta, D. *et al.* Belle II Silicon Vertex Detector. *Journal of Instrumentation* **12,** C02074–C02074. ISSN: 1748-0221 (February 2017).

68. Tanigawa, H. *et al.* Performance of the Belle II Silicon Vertex Detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **972,** 164129. ISSN: 01689002 (August 2020).

69. Taniguchi, N. Central Drift Chamber for Belle-II. *Journal of Instrumentation* **12,** C06014–C06014. ISSN: 1748-0221 (June 2017).

70. Nakagiri, K. *Central Drift Chamber for the Belle II Experiment* 2020.

71. Nishimura, K. The Time-of-Propagation Counter for BelleII. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **639,** 177–180. ISSN: 01689002 (May 2011).

72. Nishida, S. *et al.* Aerogel RICH for the Belle II Forward PID. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. RICH2013 Proceedings of the Eighth International Workshop on Ring Imaging Cherenkov Detectors Shonan, Kanagawa, Japan, December 2-6, 2013* **766,** 28–31. ISSN: 0168-9002 (December 2014).

73. Iwasaki, Y., Cheon, B., Won, E. & Varner, G. *Level 1 Trigger System for the Belle II Experiment* in *2010 17th IEEE-NPSS Real Time Conference* (May 2010), 1–9.

74. Itoh, R. *et al.* The Performance of Belle II High Level Trigger in the First Physics Run. *EPJ Web of Conferences* **245,** 01040. ISSN: 2100-014X (2020).

75. Belle II Collaboration. *Belle II Luminosity* https://www.belle2.org/research/luminosity/. (February 26, 2024).

76. Lange, D. J. The EvtGen Particle Decay Simulation Package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines* **462,** 152–155. ISSN: 0168-9002 (April 2001).

77. Was, Z. *Tauola: Modelling Tau Decays* https://indico.cern.ch/event/300387/session/7/contribution/33/material/slides/0.pdf. 2014.

BIBLIOGRAPHY

78. Kleiss, R. & Burkhardt, H. BBBREM — Monte Carlo Simulation of Radiative Bhabha Scattering in the Very Forward Direction. *Computer Physics Communications* **81,** 372–380. ISSN: 0010-4655 (July 1994).

79. Jadach, S., Płaczek, W. & Ward, B. F. L. BHWIDE 1.00: O(α) YFS Exponentiated Monte Carlo for Bhabha Scattering at Wide Angles for LEP1/SLC and LEP2. *Physics Letters B* **390,** 298–308. ISSN: 0370-2693 (January 1997).

80. Balossini, G. *et al.* Mini–Review on Monte Carlo Programs for Bhabha Scattering. *Nuclear Physics B - Proceedings Supplements. Proceedings of the 9th DESY Workshop on Elementary Particle Theory* **183,** 168–173. ISSN: 0920-5632 (October 2008).

81. Davidson, N., Przedzinski, T. & Was, Z. PHOTOS Interface in C++. *Computer Physics Communications* **199,** 86–101. ISSN: 00104655 (February 2016).

82. Agostinelli, S. *et al.* Geant4—a Simulation Toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506,** 250–303. ISSN: 0168-9002 (July 2003).

83. Ramirez, C., Donoghue, J. F. & Burdman, G. Semileptonic b $\rightarrow$ u Decay. *Physical Review D* **41,** 1496–1503. ISSN: 0556-2821 (March 1990).

84. Prim, M. T. *eFFORT: Tools for reweighting semileptonic B decays* `https://github.com/b2-hive/eFFORT`. November 2022. (November 13, 2023).

85. Bourrely, C., Lellouch, L. & Caprini, I. Model-Independent Description of $B \rightarrow \pi \ell \nu$ Decays and a Determination of V_{ub}. *Physical Review D* **79,** 013008 (January 2009).

86. Scora, D. & Isgur, N. Semileptonic Meson Decays in the Quark Model: An Update. *Physical Review D* **52,** 2783–2812 (September 1995).

87. Kuhr, T., Pulvermacher, C., Ritter, M., Hauth, T. & Braun, N. The Belle II Core Software. *Computing and Software for Big Science* **3,** 1. ISSN: 2510-2044 (November 2018).

88. Brun, R. & Rademakers, F. ROOT — An Object Oriented Data Analysis Framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. New Computing Techniques in Physics Research V* **389,** 81–86. ISSN: 0168-9002 (April 1997).

89. Moll, A. The Software Framework of the Belle II Experiment. *Journal of Physics: Conference Series* **331,** 032024. ISSN: 1742-6596 (December 2011).

90. Project Jupyter. *Jupyter* `https://jupyter.org`.

91. Zhou, Q.-D. *Correction for tracking momentum bias based on invariant mass peak studies* `https://docs.belle2.org/record/1947/files/BELLE2-NOTE-PH-2020-030-v3.pdf`. July 2020. (December 5, 2023).

92. Ji, Q., Jia, S. & Shen, C. *Energy resolution and bias with $E_\gamma < 2.0$ GeV using symmetric decays of $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ at Belle II* `https://docs.belle2.org/record/2348/files/BELLE2-NOTE-PH-2021-021.pdf`. May 2021. (December 5, 2023).

93. Sutcliffe, W., Junkerkalefeld, H. & Eliachevitch, M. *PIDvar: Offline Particle ID Efficiency and Fake Rate Corrections for Belle II* `https://gitlab.desy.de/william.sutcliffe/pidvar`. 2019. (February 26, 2023).

94. Fox, G. C. & Wolfram, S. Event Shapes in e^+e^- Annihilation. *Nuclear Physics B* **149,** 413–496. ISSN: 0550-3213 (March 1979).

BIBLIOGRAPHY

95. Keck, T. FastBDT: A Speed-Optimized Multivariate Classification Algorithm for the Belle II Experiment. *Computing and Software for Big Science* **1,** 2. ISSN: 2510-2044 (September 2017).

96. Sutcliffe, W. *BinFit* https://stash.desy.de/users/sutclw/repos/binfit. 2020. (November 13, 2023).

97. James, F. & Roos, M. Minuit - a System for Function Minimization and Analysis of the Parameter Errors and Correlations. *Computer Physics Communications* **10,** 343–367. ISSN: 0010-4655 (December 1975).