Measurement of $B \to \pi\pi\ell\nu$ with Full Hadronic Reconstruction at Belle

Dissertation
zur
Erlangung des Doktorgrades
„Doctor rerum naturalium“
der Georg-August-Universität Göttingen

im Promotionsprogramm ProPhys
der Georg-August University School of Science (GAUSS)

vorgelegt von

César Augusto Beleño de la Barrera
aus
Sincelejo-Colombia

Göttingen, 2018

Betreuungsausschuss

Prof. Dr. Ariane Frey
Prof. Dr. Arnulf Quadt

Mitglieder der Prüfungskommission:

Referentin:	Prof. Dr. Ariane Frey II. Physikalisches Institut, Georg-August-Universität Göttingen
Koreferent:	Prof. Dr. Jochen Dingfelder Physikalisches Institut, Rheinische Friedrich-Wilhelms-Universität Bonn

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Baida Achkar
II. Physikalisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Wolfram Kollatschny
Institut für Astrophysik, Georg-August-Universität Göttingen

Prof. Dr. Stan Lai
II. Physikalisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Steffen Schumann
Institut für Theoretische Physik, Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 1.10.2018

En memoria de mis abuelos José Francisco de la Barrera y Pascual Beleño, a quienes no pude decirles el último adiós.

Measurement of $B \to \pi\pi\ell\nu$ with Full Hadronic Reconstruction at Belle

Abstract

This thesis presents an analysis of the exclusive $B \to \pi^+\pi^-\ell\nu_\ell$ decay, where ℓ represents an electron or a muon. The analysis uses the complete $\Upsilon(4S)$ data sample collected by the Belle detector, corresponding to 711 fb^{-1}. The events are selected by fully reconstructing one B meson into hadronic decay modes, consequently the properties of the other B meson, "the signal B", can be inferred via conservation laws. The branching ratio is extracted by using an extended binned maximum likelihood method and measured to be $\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell) = (2.24 \pm 0.19_{\text{stat}} \pm 0.13_{\text{syst}}) \times 10^{-4}$. This result is the first reported measurement of this decay. The dominant resonance in the dipion mass spectrum is the ρ^0 meson, observed with a branching ratio of $\mathcal{B}(B^+ \to \rho^0\ell^+\nu_\ell) = (1.71 \pm 0.16_{\text{stat}} \pm 0.10_{\text{syst}}) \times 10^{-4}$. From this measurement, the magnitude of the CKM matrix element V_{ub} is obtained as $|V_{ub}| = (3.56 \pm 0.17_{\text{stat}} \pm 0.10_{\text{syst}}\ ^{+0.38}_{-0.25\ \text{theo}}) \times 10^{-3}$, in agreement with the most precise measurement of this value from global fits to the $B \to \pi\ell\nu_\ell$ channel. Another resonance is reported with a product of branching ratios, $\mathcal{B}(B^+ \to f_2(1270)\ell^+\nu) \times \mathcal{B}(f_2(1270) \to \pi^+\pi^-) = (1.8 \pm 0.9_{\text{stat}}{}^{+0.2}_{-0.1\text{syst}} \times 10^{-5}$, with a 2.2σ statistical significance without taking into account systematic uncertainties.

CONTENTS

CHAPTER 1

INTRODUCTION

The standard model of particle physics is one of the most significant achievements of the 20th century in the field of physics. It successfully describes the fundamental interactions of elementary particles and predicts a wide range of phenomena. However, it contains some parameters that are not possible to calculate from first principles and therefore, must be determined empirically. Among these parameters are the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements that govern the rate of the weak transition between quark flavors. The most promising channels to extract the magnitude of these elements are semileptonic decays. For instance, in the nuclear β reaction ${}^{14}_{6}\mathrm{C} \to {}^{14}_{7}\mathrm{N} + e^- + \bar{\nu}_\ell$, a neutron in the Carbon-14 atom transforms into a proton by emitting an electron and an anti-neutrino. The Carbon-14 becomes another element, Nitrogen-14. At quark level, this decay involves the $u \to d$ transition, and thus can be used to measure the magnitude of the $|V_{ud}|$. Among the measured CKM matrix elements, $|V_{ub}|$ has the largest uncertainty, at the 4% level [1], so a precise determination of this element will allow for stringent tests of the CKM mechanism of the Standard Model that also describes CP violation [2].

The value of $|V_{ub}|$ can be extracted from charmless semileptonic B decays. The term charmless refers to hadronic states resulting from combinations of the light quarks u or d. Two approaches are applied to obtain this value; they use different theoretical assumptions and experimental methodology. The inclusive approach considers all possible final states with charmless mesons, while the exclusive approach focuses on a particular decay channel. Presently, there exists a tension at the 3σ level between the measurements of $|V_{ub}|$ from both approaches. Given that the measured exclusive modes do not saturate the total inclusive charmless semileptonic decay rate, measurements of other channels are necessary to improve the understanding of the inclusive modeling. Since only a quarter of the exclusive charmless semileptonic B decays have been measured [1], the lack of knowledge on the remaining states constitute a source of systematic uncertainty for the determination of $|V_{ub}|$ in the inclusive approach.

The present analysis explores the exclusive $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay and reports the first measurement of its branching ratio. This channel is crucial as it broadens the number of exclusive channels and opens the door to future measurements involving charmless mesons with masses above 1 GeV.

This thesis starts with an overview of the theoretical framework of the Standard Model, B physics, and theoretical developments on the $B \to \pi\pi\ell\nu_\ell$ decay in Chapter 2. Chapter 3 describes the experimental setup and components of the Belle detector, the methods used for particle identification and the trigger system. Chapter 4 presents a synopsis of the analysis tools employed in this thesis. Chapter 5 outlines the data and monte carlo samples used for this research, together with a description of the signal reconstruction

and background rejection strategy. Chapter 6 explains the signal extraction procedure, the implementation of an extended binned maximum likelihood method and the strategy to validate the results obtained therein. Chapter 7 discusses the sources of systematic uncertainties for the measurement. Chapter 8 presents the main results from this thesis, that is the measurement of branching ratios, the study of control samples, a strategy to search for resonances in the dipion mass spectrum, and a measurement of the $|V_{ub}|$ CKM matrix element. Finally, Chapter 9 closes with a summary of this thesis and suggestions for further work on this channel.

Throughout this thesis, the physical quantities, such as mass or energy, are expressed in terms of natural units according to $\hbar = c = 1$, where $\hbar$ is the reduced Planck constant, and c is the speed of light in the vacuum. The dimensions of the detector are given in SI units.

CHAPTER 2

PHYSICS FOUNDATIONS

In this chapter, an overview of the fundamental concepts of the particle physics necessary to study the $B \to \pi\pi\ell\nu$ decay is given. The chapter starts with a brief description of the Standard Model of particle physics (section 2.1), followed by an introduction to B physics and semileptonic decays (section 2.2), and finally recent theoretical developments on the $B \to \pi\pi\ell\nu$ decay are discussed in section 2.3.

2.1 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics is the modern description of the fundamental laws of nature at the subatomic level. It is a well-tested theory, being able to successfully explain most of the results at accelerators and cosmic ray experiments, as well as to predict a wide range of phenomena. The SM can be defined as a relativistic local quantum field theory in four space-time dimensions [3] whose properties derive from the symmetries of the $SU(3) \otimes SU(2) \otimes U(1)$ gauge group. Matter is composed by fundamental particles subjected to the action of four forces: gravitational, electromagnetic, weak and strong. Out of these forces, only gravity is not included in the SM. The three remaining forces are mediated by spin-1 gauge bosons [1], also known as force-carriers, namely photons (γ) for the electromagnetic interaction, gluons (g) for the strong interaction and the massive weak vector bosons $W^{\pm}$ and Z^0 for the weak interaction.

The elementary particles are classified as quarks or leptons, depending on whether or not they interact via the strong force, respectively. In the SM, quarks and leptons are described as chiral fermions [2], coming in three families or generations, with an associated partner known as antiparticle. Each generation contains a pair of particles ordered according to a mass hierarchy, from which the first generation corresponds to the lightest and hence the most stable particles that make up the ordinary matter. The particles in the second and third generation are only accessible at high energy scales, conditions that are only met at particle accelerators, cosmic rays or in the early universe. In the lepton sector, a pair is formed by a charged particle and its corresponding neutral particle or neutrino: electron (e^-) and electron neutrino (ν_e), muon (μ^-) and muon neutrino (ν_u), and tau (τ^-) and tau neutrino (ν_τ). Charged leptons have an integer value for the electric charge (Q) given in terms of the electron charge (e) and because they do not

[1] A boson is an integer-spin particle described by the Bose-Einstein statistics.

[2] Fermions are half-integer spin particles constrained by the Pauli exclusion principle and they behave according to Fermi-Dirac statistics.

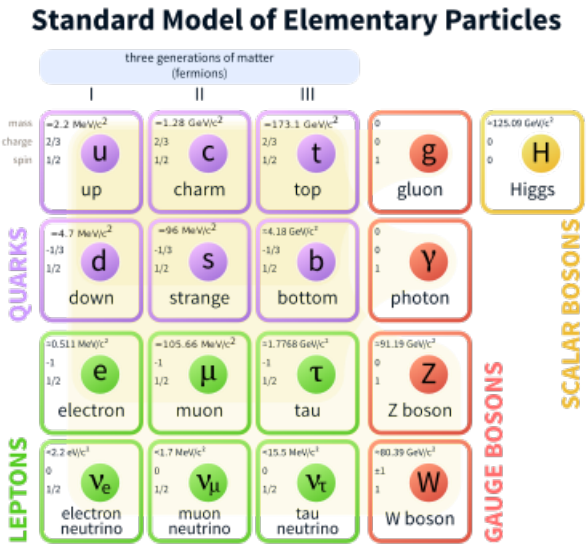


Figure 2.1: Schematic representation of the building blocks of the SM, showing the three generations of fermions (quarks and leptons), the gauge bosons responsible for the interactions and the Higgs boson. Figure available via Wikimedia Commons under the creative commons CC0 1.0 universal public domain dedication.

interact strongly, they do not form bound states and thus are found freely in nature. In the quark sector, the particle pair for each generation has a semi-integer value for the electric charge, one with $Q = +2/3e$ and the other with $Q = -1/3e$. The quark-pair per family is as follows: up (u) and down (d), charm (c) and strange (s), and top or truth (t) and bottom or beauty (b). Quarks, with the exception of the top, are always found forming bound states or hadrons, either as a quark-antiquark state referred to as meson or as a combination of three quarks (or three antiquarks) known as baryons. Examples of mesons are the $\pi^+(u\bar{d})$ and the $B^0(d\bar{b})$, and as for baryons the proton (uud) and the neutron (udd).

The last building block of the SM is the Higgs boson, which was discovered by the ATLAS [4] and CMS [5] experiments in 2012. In the SM all particles are massless, however they acquire mass upon interactions under the Higgs field that permeates the entire universe. Depending on the strength of the interaction with this field, the particles gain mass accordingly. Like the other fields in the SM, the Higgs field has an associated particle: the Higgs boson. Unlike the gauge bosons, which are spin-1 particles, the Higgs boson is a scalar particle, i.e. a spin-0 particle. A graphic representation of the building blocks of the SM is presented in Fig. 2.1. The SM is far from being a complete theory, in fact it contains 19 free parameters that must be determined from experiments. It cannot explain the dominance of matter over antimatter in the universe, neither the origin of dark matter and dark energy. Among other phenomena not accounted for by the SM, one can find the neutrino oscillations, where a deficit of two-thirds with respect to the SM prediction, is found in the solar electron neutrino flux. These flaws in the theory suggest that the SM is a low energy limit of a more fundamental theory, in other words, the SM is just an effective field theory of a more complete theory, which may include gravity.

The following subsections outline the fundamental aspects of the two constituent theories of the SM: quantum chromodynamics (QCD) for the strong force and the electroweak theory that unifies electromagnetism with the weak interaction.

2.1.1 Quantum chromodynamics

The strong interaction is the force responsible for binding quarks in hadrons, and also protons and neutrons in nuclei. The nuclear force is just an effective remnant of the strong interaction between quarks. In analogy to the electric charge in electromagnetism, the particles participating in the strong interaction carry a color charge or simply color. Historically, the concept of color was introduced as an additional degree of freedom to describe the total wave function of the Δ^{++} baryon [6]. The quark content of this baryon consist of three up quarks and total spin of $J = 3/2$; therefore, as a state composed of indistinguishable fermions, it should be totally antisymmetric under the interchange of two fermions to satisfy the Pauli exclusion principle. However, the wave functions for the spatial, flavor and spin components are symmetric, which makes necessary the introduction of an antisymmetric component to assure that the total wave function is antisymmetric. The quantum number for this component is the color, which comes in three varieties for quarks: "red", "blue" and "green", with their respective anti-color for the anti-quarks. The theory of strong interaction is named quantum chromodynamics (QCD) after the colors. Unlike the photon that does not have an electric charge, gluons carry a color and thus they interact with one another. Because color is conserved in all physical processes [7], when two quarks interact they emit or absorb a gluon that carries the difference between the two colors of the quarks; this implies that gluons are bi-colored carrying one unit of color and one unit of anti-color. In general, all combinations of quarks must be colorless, hence mesons comprise a quark of one color and an antiquark of the corresponding anticolor, whereas baryons must consist of quarks of different colors.

The arrangement for three species of color in a color/anti-color state gives nine possible combinations for gluons whose linear combinations result in an octet and a singlet. The singlet state composed by red/anti-red + blue/anti-blue + green/anti-green is color-neutral, which implies that it must be non-interacting and if it exists as a mediator, it should occur as a free particle. Moreover, it could be exchanged between two colorless baryons[3], suggesting that they interact via the strong force, contrary to the evidence [9]. Consequently, there are only eight gluons.

Color symmetry is described by the non-Abelian Lie group $SU(3)$, a special unitary group in three dimensions, that has eight generators corresponding to the eight gluons of QCD. The assumption of local gauge symmetry leads to the QCD lagrangian that can be written as

$$\mathcal{L}_{QCD} = -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} + i \sum_q \bar{\psi}_q \gamma^\mu D_\mu \psi_q - \sum_q m_q \bar{\psi}_q \psi_q, \tag{2.1}$$

where

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_s f_{abc} A^b_\mu A^c_\nu, \tag{2.2}$$

and

$$D_\mu = \partial_\mu - \frac{i}{2} g_s \lambda^a A^a_\mu, \tag{2.3}$$

in the above equations g_s is the dimensionless QCD coupling constant, the f^{abc} are the structure constants of the $SU(3)$ group, m_q is the quark mass and the index a runs from 1 to 8 for the eight gluon fields A^a_μ. The first term in eq. 2.1 describes the propagation of the gluon fields and also gluon self-interactions. The second term can be expanded in two components, one describing the propagation of free quarks and the other the interaction between gluons and quarks, and finally the third term is the mass term resulting

[3] In the quark model, both mesons and baryons are not affected by a color transformation; they are said to be color singlets [8].

from spontaneous symmetry breaking in the electroweak sector. The QCD coupling is defined as

$$g_s = \sqrt{4\pi\alpha_s}, \tag{2.4}$$

with α_s the QCD coupling strength that indicates how the effective color-charge between two quarks depends on the distance between them [10]. The contributions for this dependence come from two sources. On one hand we have the shielding effect of a virtual quark-antiquark pair that reduces the α_s value for increasing distance, and on the other hand, since gluons have a net color-charge, they produce the opposite effect called anti-screening. Consequently, the effects of the gluons exceeds the impact of the quark-antiquark pairs, increasing the magnitude of α_s with the increase in distance. This behavior of α_s leads to the confinement of quarks inside hadrons, and offers an explanation to the question of why no colored objects have ever been observed in nature. At short distance scales α_s vanishes, allowing quarks to behave as free particles. This phenomenon is referred to as asymptotic freedom. The lowest-order expression for α_s referred to as the running coupling constant is given by

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f)\ln(Q^2/\Lambda_{QCD})}. \tag{2.5}$$

Here Q represents the energy scale of the process also known as momentum transfer, n_f is the effective number of quark flavors at that energy, and Λ_{QCD} corresponds to energy boundary between asymptotically free quarks and hadrons. Typical values of Λ_{QCD} fall in the range between 200 MeV and 300 MeV [11]. This scale defines two energy regimes in QCD, a non-perturbative low-energy regime, where calculations from first-principles are not currently possible (such as the hadronisation processes) and a high-energy regime where perturbation theory can be used[4]. The energy scale of Λ_{QCD} also allows us to categorize the quarks according to their masses: the light quarks (u, d, s) with masses below the QCD energy scale, and the heavy quarks (c, b, t) with masses above this scale.

2.1.2 Electroweak Theory

The two remaining interactions accounted for by the SM are the electromagnetic and weak interactions. The electromagnetic interaction is responsible for the binding of the electrons to the atomic nucleus, the formation of molecules and chemical reactions; macroscopic phenomena such as friction and air resistance, among others. By comparison, the weak interaction causes radioactive decay of nuclei like the β-decay[5], the decay of light hadrons such as the pion, kaon and neutron, and gives rise to processes with neutrinos. Unlike the other interactions, the weak force does not produce bound states and violates both charge conjugation and parity symmetries[6]. The first experimental evidence of parity violation was carried out by Wu and collaborators in 1957 [12] in a β-decay of polarised cobalt-60, $^{60}\text{Co} \rightarrow ^{60}\text{Ni}^* + e^- + \bar{\nu}_e$. The ^{60}Co nuclei, with an inherent nuclear magnetic moment μ, were aligned in a strong magnetic field $\mathbf{B}$ and the β-decay electrons were detected at different polar angles with respect to this axis. Since both $\mathbf{B}$ and μ are axial vectors, the parity operation does not change their sign. In the mirror world of this system, the only quantity that changes sign is the momentum of the emitted electrons. Hence, if parity

[4] For high energy scales or short distances ($Q \gg \Lambda_{QCD}$), the value of α_s becomes sufficiently small to allow for high-order calculations in terms of α_s (perturbative corrections)

[5] The β-decay refers to the conversion of a neutron into a proton by the emission of an electron and an electron antineutrino

[6] The charge conjugation is a discrete symmetry transformation that replace particles with their corresponding antiparticles. The parity transformation corresponds to spatial inversion through the origin, $\mathbf{x} \rightarrow -\mathbf{x}$. Scalar quantities such as temperature, or axial vectors like the angular momentum remains unaltered under parity transformations, while pseudoscalar quantities such as helicity or vectors like the momentum change sign under parity operations.

were conserved in the weak interaction, the rate at which electrons were emitted at a certain direction relative to the **B**-field would be identical to the rate in the opposite direction. The empirical results showed a deficit of electrons in the direction of the applied field, and an excess in the opposite direction, thus providing a clear demonstration that parity is not conserved in the weak interaction [7]. One way of translating this observation to mathematical concepts was to decompose the wavefuntion of fermions into two components: left-handed and right-handed. As a consequence, the weak interaction distinguishes between left-handed and right-handed particles. Another aspect of the weak interaction is that it facilitates flavor changing interactions, indeed the β-decay, at the quark level, implies the flavor change of one down quark in the neutron to an up quark, changing the neutron for a proton by emitting an electron and an electron antineutrino ($d \rightarrow ue^{-}\bar{\nu}_e$). There are two ways in which the weak interaction takes place: the charged current (CC) mediated by the $W^{\pm}$ bosons where the electric charges of the interacting particles differ by one unit, and the neutral current (NC) mediated by the Z^0 boson, where the electric charges of the interacting particles are equal [11].

The electroweak sector of the SM is based on the symmetry group an $SU(2)_L \otimes U(1)_Y$, where the subindices L and Y refer to the left-handed fields and the hypercharge respectively. The gauge symmetry is broken by the vacuum, which triggers the spontaneous symmetry breaking (SSB) of the electroweak group into the electromagnetic subgroup, $SU(2)_L \otimes U(1)_Y \rightarrow U(1)_Q$, with Q the electric charge defined as

$$Q = \frac{Y}{2} + T_3, \tag{2.6}$$

and T_3 the diagonal generator of the group $SU(2)_L$ also known as weak isospin [13]. The SSB mechanism generates the masses of the weak gauge bosons and the fermions, and gives rise to the appearance of a physical scalar particle in the model, the Higgs boson [14]. The $SU(2)_L$ symmetry is associated with three gauge bosons ($W^{\mu i}, i = 1, 2, 3$) and a gauge coupling constant g', whereas the $U(1)_Y$ is associated with one gauge boson B^{μ} and the gauge coupling constant g. In the SM the couplings g and g' cannot be calculated and must be determined empirically. The W^i and B bosons are not the physical bosons that mediate the interactions, instead they mix to form the $W^{\pm}$ bosons of the CC interactions, and the photon (A) and Z^0 bosons of the NC interactions. The following relations show the connection between the bosons of the gauge groups and the physical bosons:

$$W^{\pm\mu} = \frac{1}{\sqrt{2}}(W_1^{\mu} \mp iW_2^{\mu}), \tag{2.7}$$

$$\begin{bmatrix} Z^0 \\ A \end{bmatrix} = \begin{bmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{bmatrix} \begin{bmatrix} W_3 \\ B \end{bmatrix}, \tag{2.8}$$

where the angle θ_W is the weak mixing angle, also known as the Weinberg angle, and is defined by the relationship

$$\theta_W = \tan^{-1}\left(\frac{g'}{g}\right). \tag{2.9}$$

The electroweak Lagrangian contains a vast number of terms that can be grouped in four components [15]:

$$\mathcal{L}_{\text{electroweak}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \tag{2.10}$$

Given the extension of the Lagrangian, the reader is referred to the textbooks [7, 9, 15] for further details. Some important aspects of the Lagrangian are discussed in the following.

The first term in eq. 2.10, $\mathcal{L}_{\text{gauge}}$, corresponds to the kinetic energy of the vector bosons and can be

written as

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{4}F_a^{\mu\nu}F^a_{\mu\nu}, \tag{2.11}$$

where $F_{\mu\nu}$ and $F^a_{\mu\nu}$ are the $U(1)$ and $SU(2)$ field tensors, respectively. They are given by

$$F_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{2.12}$$

$$F^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^q_\mu + g\epsilon_{abc}W^b_\mu W^c_\nu, \tag{2.13}$$

with $a = 1, 2, 3$ and ϵ_{abc} are the structure constants for $SU(2)$. Since $F^a_{\mu\nu}$ includes a quadratic term in W^a_ν, the Lagrangian $\mathcal{L}_{\text{gauge}}$ gives rise to cubic and quartic self-interactions among the gauge fields.

The second term in eq. 2.10, $\mathcal{L}_{\text{fermion}}$, describes the kinetic energy of the fermions and the fermion-boson interactions. In a simplified manner, it can be written as

$$\mathcal{L}_{\text{fermion}} = i\bar{\psi}\gamma^\mu D_\mu \psi, \tag{2.14}$$

where γ^μ are the four Dirac matrices and D_μ is the covariant derivative.

Since the weak interaction discriminates between left-handed and right-handed fermions, the former are assigned to a doublet and the latter to a singlet[7], the covariant derivative must be modified to account for this fact. Thus

$$\mathcal{L}_{\text{fermion}} = i\bar{\psi}_R\gamma^\mu(\partial_\mu + ig'YB_\mu)\psi_R + i\bar{\psi}_L\gamma^\mu[\partial_\mu + \frac{i}{2}(g'YB_\mu + gT_aW^a_\mu)]\psi_L, \tag{2.15}$$

where T_a are the generators of the group $SU(2)$.

The next term in eq. 2.10 is responsible for the SSB by which the three weak bosons acquire mass. To that end the scalar fields require a choice of representation in which the field with a nonzero vacuum value is electrically neutral, so the photon remains massless, but it must carry nonzero values of T_3 and Y so that the Z^0 boson acquires mass. The simplest representation is

$$\phi = \begin{pmatrix}\varphi^+ \\ \varphi^0\end{pmatrix} \qquad \text{and} \qquad \phi^\dagger = (-\varphi^-, \varphi^{0\dagger}). \tag{2.16}$$

Then the Higgs Lagrangian can be written as

$$\mathcal{L}_{\text{Higgs}} = \left|\left(i\partial_\mu - \frac{g}{2}T_aW^a_\mu - \frac{g'}{2}B_\mu\right)\phi\right|^2 - \frac{1}{2}\mu^2\phi^\dagger\phi - \frac{1}{4}\lambda(\phi^\dagger\phi)^2, \tag{2.17}$$

with $\mu^2, \lambda \in \mathbb{R}$ and $\lambda > 0$. The scalar potential, $V(\phi) = \frac{1}{2}\mu^2\phi^\dagger\phi + \frac{1}{4}\lambda(\phi^\dagger\phi)^2$, has a continuum minima for $\mu^2 < 0$, where the vacuum expectation value (VEV) satisfies $<\phi>= v = \sqrt{-\frac{\mu^2}{\lambda}}$. The ϕ can then be expanded around this VEV as $\phi(x) = \frac{1}{\sqrt{2}}(h(x) + v)$, and hence the Higgs Lagrangian becomes

[7] For each generation of fermions we have $\psi_L = \binom{\nu_L}{e_L^-}$ and $\psi_R = e_R^-$ for leptons, and $\psi_L = \binom{u'_L}{d'_L}$ and $\psi_R = u'_R$ or d'_R with $u' = u, c, t$ and $d' = d, s, b$ for quarks.

$$\mathcal{L}_{\text{Higgs}} = \frac{g^2}{4} W_-^\mu W_\mu^+ (h+v)^2 + \frac{1}{2}\partial^\mu h \partial_\mu h + \frac{\sqrt{g^2+g'^2}}{8} Z^{\mu 0} Z_\mu^0 (h+v)^2 + \frac{\mu^2}{2}(h+v)^2 + \frac{\lambda}{4}(h+v)^4, \tag{2.18}$$

where h is the Higgs field. The coefficients of the quadratic terms in the fields correspond to the square of their masses. In this way the masses for the W and Z bosons are $M_W = gv/2$ and $M_Z = \sqrt{g^2+g'^2}v/2$. The Higgs Lagrangian also includes the interactions between the Higgs field, the $W^\pm$ and Z^0 bosons, as well as the Higgs self-interactions. However, this Lagrangian has neither interaction terms nor mass terms for the photon, and hence the photon remains massless.

The last term in eq. 2.10, the Yukawa Lagrangian ($\mathcal{L}_{\text{Yukawa}}$), couples the Higgs field to the fermions to provide masses of the quarks and charged leptons. The expression for this lagrangian is given by

$$\mathcal{L}_{\text{Yukawa}} = g_f(\bar{\psi}_L \phi \psi_R + \bar{\psi}_R \phi \psi_L), \tag{2.19}$$

where g_f is a constant known as the Yukawa coupling of the fermion to the Higgs field. The Yukawa coupling cannot be predicted from the SM, it must be determined empirically. Nonetheless, after the Higgs field has acquired a VEV from SSB, the masses of the fermions (m_f) can be related to the Yukawa coupling by

$$g_f = \sqrt{2}\frac{m_f}{v}. \tag{2.20}$$

Hence, the larger the magnitude of the Yukawa coupling is, the more massive the particle becomes. It should be emphasized that the Higgs mechanism does not give mass to composite particles such as protons or nuclei; they owe their masses to the binding energy that holds together the quarks inside them.

2.1.3 The Cabibbo-Kobayashi-Maskawa (CKM) mechanism

In the 1960's, the measured decay rate for leptons suggested that the strength of the coupling of the $W^\pm$ bosons to the neutrino-lepton pairs for all generations of leptons was the same, a property known as lepton universality. An analogous situation was expected in the quark sector, where the $W^\pm$ boson couples to up-type and down-type quark pairs. In this way, universality would imply that decays such as $n \to pe^-\bar{n}u_e$ and $\Lambda \to pe^-\bar{n}u_e$, with $d \to u$ and $s \to u$ quark transitions respectively, will have the same matrix elements contrary to the evidence where the latter is significatively smaller than the former. This, and other observations, led Nicola Cabibbo in 1963 to formulate the concept of quark mixing, in which the weak eigenstates of quarks that couple to the W bosons are quantum superpositions of the mass eigenstates of quarks [16]. The relationships between these two basis are contained in a 2×2 unitary matrix in terms of a single parameter, the Cabibbo angle θ_{12}. This angle allows for mixing between the first and second generation of quarks. The extension to three generations was done by Makoto Kobayashi and Toshihide Maskawa in 1973, with the argument that CP violation arise naturally in the theory with the inclusion of a third generation, while it was impossible with just two generations [17]. The 3×3 matrix transformation that relates the weak eigentstates (d', s', b') to the mass eigenstates (d, s, b) is referred to

as the Cabibbo-Kobayashi-Maskawa (CKM) matrix (V_{CKM}) and is defined as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.21}$$

As of 2018 the magnitudes of the CKM matrix elements are [1]

$$V_{CKM} = \begin{pmatrix} 0.97434^{+0.00011}_{-0.00012} & 0.22506 \pm 0.00050 & 0.00357 \pm 0.00015 \\ 0.22492 \pm 0.00050 & 0.97351 \pm 0.00013 & 0.0411 \pm 0.0013 \\ 0.00875^{+0.00032}_{-0.00033} & 0.0403 \pm 0.0013 & 0.99915 \pm 0.00005 \end{pmatrix}. \tag{2.22}$$

An unitary $N \times N$ matrix of quark mixing has $N(N-1)/2$ real parameters and $(N-1)(N-2)/2$ phases, hence V_{CKM} can be parametrized by three rotation angles and one complex phase δ as follows

$$V_{CKM} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.23}$$

where $s_{ij} = \sin\phi_{ij}$ and $c_{ij} = \cos\phi_{ij}$. The phase δ enters the wavefunction as $\exp[i(\omega t + \delta)]$, which is not invariant under time reversal ($t \to -t$) and therefore introduces CP-violating effects in the SM. Another widely used parametrization of the CKM matrix results by expanding it in terms of $\lambda = |V_{us}| \approx 0.225$, which is known as Wolfenstein parametrization [18]. It is given in terms of four real parameters λ, A, ρ, η by

$$V_{CKM} \approx \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^5). \tag{2.24}$$

This parametrization is convenient because it shows a hierarchical pattern in the magnitudes of the CKM matrix elements, where the off-diagonal terms are relatively small and the diagonal terms are close to unity. Consequently, the weak interaction of quarks of different generations are suppressed relative to those of the same generation. The suppression is maximum for the couplings between the first and third generation quarks, ub and td. Furthermore, the complex components in the CKM matrix resides solely on V_{ub} and V_{td} at order λ^5, which implies that η must be non-zero to violate CP in the quark sector.

The unitarity of the CKM matrix imposes six orthogonality constraints on its rows and columns:

$$\begin{aligned} \sum_{\alpha} V_{\alpha i} V^*_{\alpha j} &= 0, \ i \neq j \\ \sum_{i} V_{\alpha i} V^*_{\beta i} &= 0, \ \alpha \neq \beta. \end{aligned} \tag{2.25}$$

where the Greek subscripts run over the up-type quarks (u, c, t), while the Latin ones run over the down-type quarks (d, s, b). The relations in eq. 2.25 can be graphically represented in the complex plane under the name of unitarity triangles [19] as illustrated in Fig. 2.2. Each of these triangles is named after the columns or rows whose orthogonality it represents. The length of the sides of the triangles, as well as their angles, are invariant under phase transformations and thus are physical observables that can be measured in suitable experiments. All six triangles have the same area that is a measure of the size of CP violation in the SM. If the CKM matrix were not truly unitary, it would indicate the presence of additional

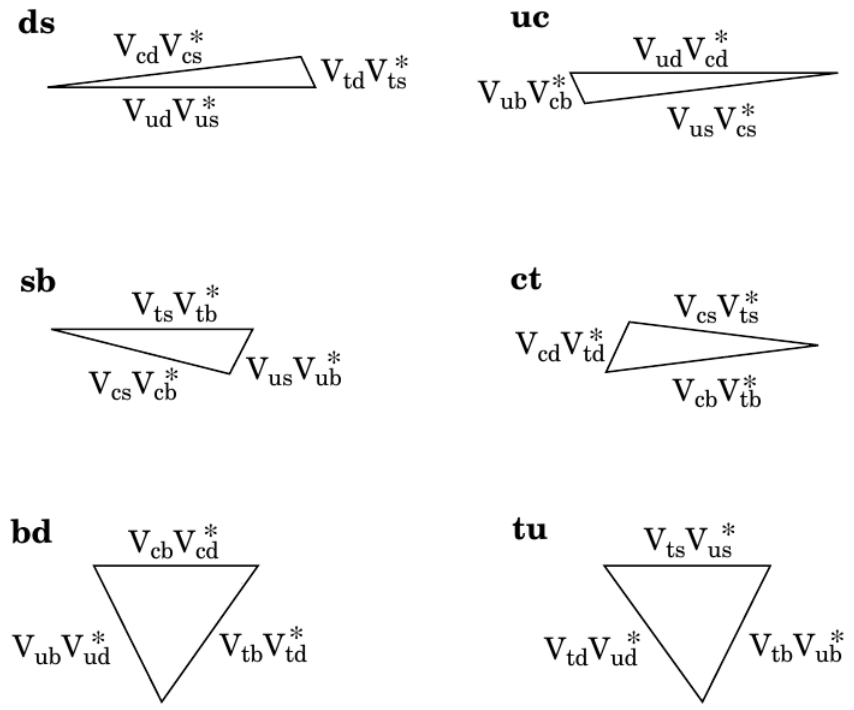


Figure 2.2: The unitarity triangles for three generations of quarks. The triangles are labeled by the pair of columns or rows whose orthogonality is represented.

quark generations or another more fundamental problem in the SM.

Out of the six unitarity triangles, the **bd**-triangle is the most commonly used which is derived from the relation

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{2.26}$$

Dividing this expression by the best known factor ($V_{cd}V_{cb}^*$) results in a triangle of unit base and sides with lengths of $|V_{ud}V_{ub}^*/V_{cd}V_{cb}^*|$ and $|V_{td}V_{tb}^*/V_{cd}V_{cb}^*|$. One can redefine the sides of this triangle as $-V_{ud}V_{ub}^*/V_{cd}V_{cb}^* = \bar{\rho} + i\bar{\eta}$ and $V_{td}V_{tb}^*/V_{cd}V_{cb}^* = 1 - \bar{\rho} - i\bar{\eta}$. The angles for this triangle are given by

$$\begin{aligned}
\alpha &= \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \\
\beta &= \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{ud}V_{ub}^*}\right), \\
\gamma &= \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right).
\end{aligned} \tag{2.27}$$

Combination of independent measurements can experimentally constrain the sides and angles of the triangle and test its unitarity. For instance, measurements of mass splitting (ΔM_d, ΔM_s) from $B_{(s)}$ meson mixing, $|V_{ub}/V_{cb}|$ from semileptonic B decays and CP violation either in the kaon system (ϵ_K) or the B meson system ($\sin(2\beta)$) can be used in a global fit in the $(\bar{\rho}, \bar{\eta})$ plane. The result of the global fit, carried out by the CKMfitter group [20], is presented in Fig. 2.3 and validates the unitarity of the CKM matrix.

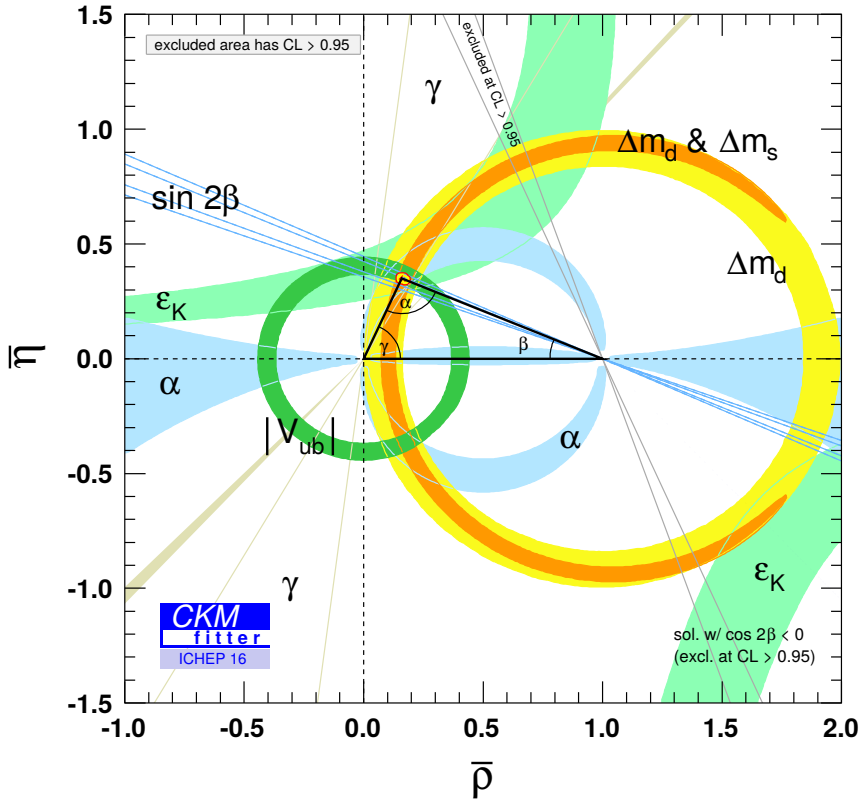


Figure 2.3: Experimental constraints on the unitarity triangle as provided by the CKMfitter group for the ICHEP conference 2016 [20].

2.2 Phenomenology of B mesons

The bottom or beauty quark (b) is the heaviest quark that hadronizes to form a bound state, which then exhibits a rich variety of decay modes. This property makes b-flavored hadrons an interesting research probe whose studies fall into five categories:

1. test of the electroweak sector of the SM,
2. study of the dynamics of the heavy quark interactions,
3. the origin of CP violation,
4. the measurement of the magnitude of the CKM matrix elements $|V_{cb}|$, $|V_{ub}|$, $|V_{td}|$ and $|V_{ts}|$,
5. and the observation of rare processes which can probe physics beyond the SM (BSM).

The b quark was discovered at Fermilab in 1977 by the E288 experiment headed by Leon Lederman [22]. They measured a $\mu^+\mu^-$ resonance from collisions of 400 GeV protons on a nuclear fixed target, and

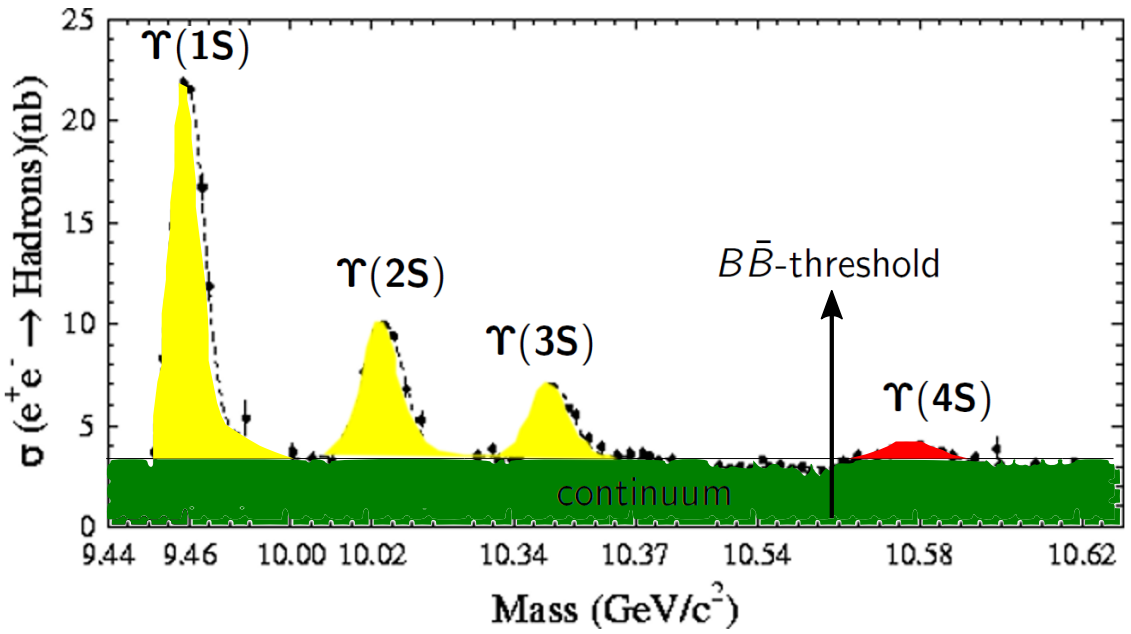


Figure 2.4: Cross section of the annihilation of e^+e^- collisions to bottomonium resonances as measured by the CLEO collaboration. The green shaded area underneath the resonances corresponds to continuum events. The figure is taken from [21] and modified for illustrative purposes.

associated this resonance with a $b\bar{b}$ bound state[8] which was then named Upsilon (Υ). This resonance was later called $\Upsilon(1S)$ and confirmed by two e^+e^- experiments, Doris at DESY (Hamburg) in 1978 [23, 24] and CESR at LNS (Cornell) in 1979 [25, 26]. They also observed other $b\bar{b}$ narrow resonances: the $\Upsilon(2S)$ and the $\Upsilon(3S)$. These particles decay via the strong interaction[9], with a b and $\bar{b}$ annihilating each other by emitting gluons or virtual photons. Henceforth, the upsilon resonances are not useful in determining further properties of the b quark, such a task can only be achieved in the study of their weak decays [27]. For that purpose, it was necessary to study a b-flavored meson, where the heavy mass of the b quark dominates the dynamics of the meson[10]. In 1980 the CESR experiment observed another upsilon resonance [28, 29], $\Upsilon(4S)$, which is just 20 MeV above the mass threshold to produce two of the lightest b-flavored particle, the B meson. The family of Υ resonances as measured by the CLEO collaboration is presented in Fig. 2.4, where the green shaded area corresponds to the nonresonant background originating from the annihilation of e^+e^- into a pair of light quarks (u, d, s, c) and is generically known as *continuum* background. The first experimental evidence of B mesons was inferred in 1981 from the measurement of the inclusive semileptonic decay rate [30, 31], where the presence of a single lepton in the final state was a signature of a weak decay. It was not until 1983 when the existence of the B meson was completely

[8] A particle consisting of a heavy quark and its antiquark is referred to as quarkonium, so in this particular case a $b\bar{b}$ state is known as bottomonium.

[9] Typical lifetime for particles whose decays are mediated by the strong, electromagnetic and weak interactions are of the order of 10^{-23} s, 10^{-16} s and 10^{-8} s, respectively. As consequence, the strong force gives no measurable lifetime, instead the particle width can be measured.

[10] Quarkonia particles are flavorless, since the total flavor quantum number of the system is zero. Hence, the upsilon resonances have a total bottom of zero.

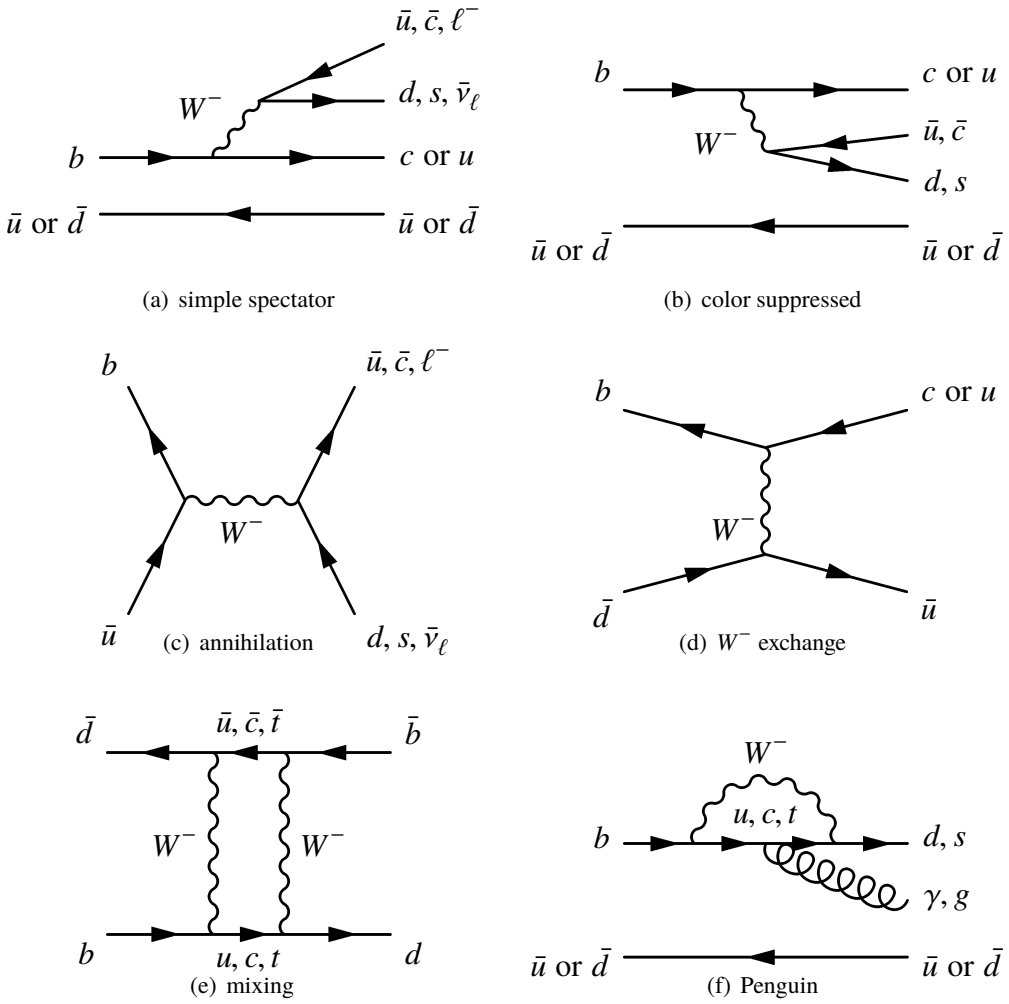


Figure 2.5: Diagrams for B meson decays, where the index ℓ refers to the three flavors of charged leptons (e^-, μ^- and τ^-).

established by the CLEO collaboration with the observation of fully hadronic decays [32].

The b-flavored meson species that result from binding a $\bar{b}$ quark and an u, d, s or c quarks are referred to as the $B_u(B^+)$, $B_d(B^0)$, B_s and B_c mesons, respectively, as depicted in Fig. 2.6. They all decay exclusively via the weak interaction. Excited or high mass states of b-flavored mesons decay strongly to these four mesons. Due to the much heavier mass of the s and c quarks, in comparison with the u and d quarks, the B_s and B_c mesons cannot be produced at the $\Upsilon(4S)$ resonance. Hereafter, the term B mesons is used to refer indistinctly to both the B^+ and B^0 mesons. The decay of B mesons occurs primarily through the CKM favored $b \to c$ transitions. Nevertheless, one can distinguish various ways in which a virtual $W^\pm$ boson mediates a B meson decay [33], depicted as Feynman diagrams in Fig. 2.5. In the simple spectator diagram, the virtual $W^\pm$ materializes into either a lepton and anti-neutrino or a $\bar{u}d$ or $\bar{c}s$ quark pair. The terminology spectator comes from considering the decay of the B meson as due to the decay of the b quark, while ignoring the presence of the spectator antiquark. This representation serves only as an approximation to the B decay. In order to predict physical observables, the confinement and hadronization of the b quark inside the B meson must be taken into account, as they are included as non-spectator effect corrections. The spectator diagram is the dominant decay mode of B mesons.

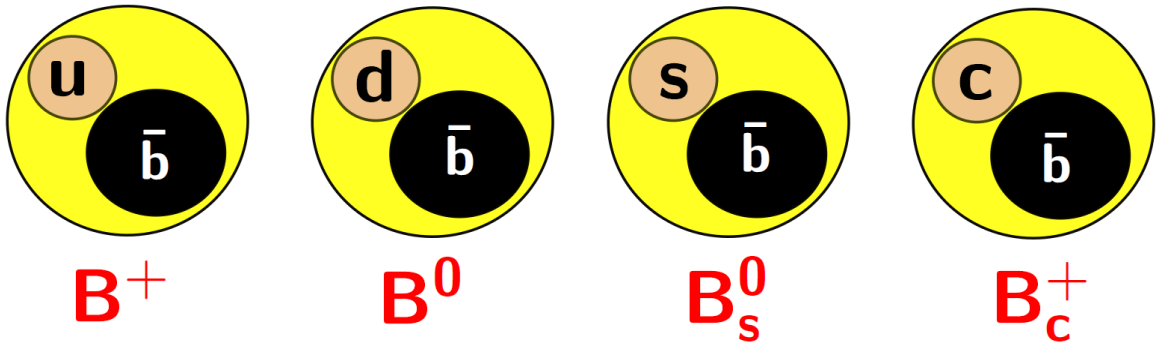


Figure 2.6: Possible combinations of a bottom anti-quark with other quarks to form a ground state B meson.

The color suppressed diagram arises when the virtual $W^{\pm}$ decays into a quark-antiquark pair that happen to have the same colors as the initial b quark. Therefore, these modes are expected to be a factor of three lower in amplitude and nine times lower in rate, with respect to the non-color suppressed spectator diagram. The annihilation diagram shown in Fig. 2.5(c) takes place when the b quark and spectator anti-quark are in the same space-time region and annihilate by coupling to a virtual $W^{\pm}$. The decay amplitude for this decay is proportional to a numerical factor known as the decay constant f_B times the CKM matrix element $|V_{ub}|$.

The neutral mesons have the possibility of oscillating to their antiparticles before they decay. This behavior is described by box diagrams in the SM as shown in Fig. 2.5(e). Although the up-quarks are exchanged in the mixing of B mesons, the t quark has a major relevance due to its mass, as the amplitude of this process is proportional to the mass of the exchanged fermion. Decays involving this mode are useful when extracting the CKM matrix elements $|V_{td}|$ and $|V_{ts}|$. The last diagram corresponds to penguin processes, which proceeds through higher order weak interactions involving loops. In this case, a b quark decays into another down-quark (s or d) by radiating a photon or a gluon, hence penguin processes proceed via flavor changing neutral current (FCNC). As in the box diagram, the relative size of the amplitude depends on the mass of the exchanged fermion. Decays that proceed through such a loop are often referred to as rare decays and they are sensitive to contributions from new particles in BSM physics.

Depending on the type of particles present in the final state, the B meson decays can be classified as hadronic, leptonic or semileptonic decays. These decays contain only hadrons, only leptons, or combination of hadrons and leptons, respectively. Hadronic decays pose serious theoretical challenges, since the dynamics of the hadrons implies the use of difficult calculations in QCD, for which the uncertainties in the knowledge of certain parameters may result in unreliable predictions. The available theoretical description of hadronic decays exploits effective theories where the QCD Lagrangian is simplified with the use of approximate flavor symmetries, or the existence of several scales, which allows for an expansion in $\Lambda_{\mathrm{QCD}}/m_b$. These approximations result in a theoretical model valid for a specific region of momenta of the particles involved. On the experimental side, hadronic decays are easy to reconstruct and deal with. Leptonic decays lie in a completely opposite situation as hadronic decays. On one hand, due to the lack of strong interaction effects among leptons, the theoretical description of final state leptons is straightforwardly calculated from QED, while the impact of the annihilation of the quark and anti-quark pairs within the meson is contained in a single parameter: the decay constant of the meson. However, because of their tiny decay rate, leptonic decays are difficult to detect experimentally. Finally, the theoretical description and experimental detection of semileptonic decays lie in an intermediate

situation between that for hadronic and for leptonic decays.

2.2.1 Semileptonic B decays

In the SM formalism, charged-current semileptonic decays at parton level proceed via exchange of a charged weak boson, W^+, between a quark and a lepton current. The quark current is given by

$$J_\mu = \bar{\mathcal{U}}_L \gamma_\mu V_{\mathrm{CKM}} \mathcal{D}_L, \tag{2.28}$$

where

$$\mathcal{U}_L = \begin{pmatrix} u_L \\ c_L \\ t_L \end{pmatrix}, \qquad \mathcal{D}_L = \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix}, \tag{2.29}$$

involving the left-handed up- and down-type quarks. The lepton current can be written as

$$j_\mu = \bar{e}_L \gamma_\mu \bar{\nu}_{e,L} + \bar{\mu}_L \gamma_\mu \bar{\nu}_{\mu,L} + \bar{\tau}_L \gamma_\mu \bar{\nu}_{\tau,L}. \tag{2.30}$$

At energies much less than the W mass, the two interactions can be combined into a local effective Hamiltonian of the form

$$H_{\mathrm{eff}}^{\mathrm{sl}} = \frac{4G_{\mathrm{F}}}{\sqrt{2}} \left(\bar{U}_L \gamma_\mu V_{\mathrm{CKM}} \mathcal{D}_L \right) \times \left(\bar{e}_L \gamma_\mu \bar{\nu}_{e,L} + \bar{\mu}_L \gamma_\mu \bar{\nu}_{\mu,L} + \bar{\tau}_L \gamma_\mu \bar{\nu}_{\tau,L} \right) + \mathrm{h.c.}, \tag{2.31}$$

where G_{F} is the Fermi coupling constant defined as

$$G_{\mathrm{F}} = \frac{g^2}{4\sqrt{2} M_W^2}, \tag{2.32}$$

and the change of $\bar{\mathcal{U}}_L$ by $\bar{U}_L$ has to do with the fact that the top quark is much more massive than the $W^\pm$ boson and that it does not hadronize. Consequently, when considering weak decays of hadrons, the effective Hamiltonian does not contain the top quark anymore and thus

$$U_L = \begin{pmatrix} u_L \\ c_L \\ 0 \end{pmatrix}. \tag{2.33}$$

The expression for the effective Hamiltonian indicates that all CKM matrix element, with the exception of those involving the top quark (V_{tq}), are accessible in semileptonic decays [34]. The differential B meson decay rates take the form

$$d\Gamma \quad \propto \quad G_{\mathrm{F}}^2 \left| V_{qb} \right|^2 \left| j^\mu \langle X | \, \bar{q}_L \gamma_\mu b_L \, | B \rangle \right|^2, \tag{2.34}$$

where $\langle X|$ represents the hadronic final state and $|B\rangle$ the initial state given by the B meson. In semileptonic decays, the leptonic part of the effective Hamiltonian factorizes from the quark current, so all QCD corrections are contained in the hadronic matrix element $\langle X | \, \bar{q}_L \gamma_\mu b_L \, | B \rangle$.

Semileptonic B decays can be studied in two different approaches. The *inclusive* considers all possible contributions to the $q \to q'$ quark transitions. The *exclusive* approach focuses in the explicit reconstruction of the final state hadron. These approaches are based on different experimental techniques and theoretical concepts, rendering complementary information for semileptonic decays and a way to test consistency between the results from the two approaches. As this thesis focuses on an exclusive decay channel, the

remainder of this chapter is dedicated to the theoretical methods used in exclusive semileptonic B decays. In exclusive semileptonic decays the hadronic matrix elements are parametrized as form factors, which are non-perturbative functions of the momentum transfer q^2 [35]. Their calculation is based on different assumptions, dependent on the type of quark transition that the decay exhibits. These transitions are of two classes, $b \to c$ and $b \to u$ quark transitions; decays involving the latter are called charmless B decays. Decays proceeding through a $b \to c$ quark transition are studied under the Heavy Quark Effective Theory (HQET), which assumes that the masses of both the b and c quarks are large compared to Λ_{QCD}. In the limit of infinite-mass, hadronic systems, which differ only in the flavor or spin quantum numbers of the heavy quark, have the same configuration of the spectator quark and gluons in the initial and final state [36]. Nonetheless, as the quark masses are finite, corrections to the calculations must be added in powers of the inverse of the heavy quark mass, $1/m_Q$, and $\alpha_s(m_Q)$. This method has proven to be successful in the precise determination of the CKM matrix element $|V_{cb}|$.

For charmless semileptonic B decays, the scenario is different as the up and down quarks are light, and the HQET is no longer applicable. For the heavy to light quark transitions, different theoretical calculations have been developed to describe the form factors, these include:

1. **Quark models**. These models assume wave functions for the mesons and use them to compute the hadronic matrix elements [37]. They are calculated at a particular value of q^2, either $q^2 = 0$ or $q^2 = q^2_{\text{max}}$. The q^2 dependence of the form factors is determined as a separate step in the calculation as an extrapolation based on phenomenological motivations. One of such models is the Isgur-Scora-Grinstein-Wise (ISGW) quark model [38], which is based on a nonrelativistic constituent-quark potential model, with an assumed Coulomb-plus-linear form for $V(r)$. Here, the form factor is calculated at $q^2 = q^2_{\text{max}}$. This model was updated to include relativistic corrections under the name ISGW2 [39]. Due to the unclear connection to QCD, the theoretical uncertainties from this model are not under control, and hence impossible to estimate.

2. **QCD sum rules**. The main idea behind this method is to perform a perturbative calculation of suitable correlation functions in a kinematic region where perturbation theory can be applied. These perturbative results are then extended to the nonperturbative regions by utilizing unitarity and analyticity, that is, by using a dispersion relation [34]. These calculations are valid for low values of q^2 and have proven to be a powerful tool to derive form factors with a minimal number of phenomenological assumptions [40]. One of these methods is the Light Cone Sum Rule (LCSR), which implements light-cone distribution amplitudes for both the heavy and light mesons.

3. **Lattice QCD (LQCD)**. This method calculates nonperturbative QCD processes on a discrete lattice characterized by a spacing a via a numerical simulation [34]. In the limit when $a \to 0$ and the number of grids is infinite, the continuum limit, LQCD can provide an accurate solution to QCD; in practice the calculations are carried out for different values of a and then extrapolated to the continuum limit. However, these calculations are limited by the amount of computational power. To ease the computational burden a quenching approximation is used, which consists of removing all Feynman diagrams with closed fermion loops. Most of the LQCD results available are quenched, but in some cases the use of the quenching approximation results in unknown effects on the results, thus, when possible, an unquenched calculation is desired. One limitation in the implementation of the LQCD calculation is the simulation of heavy quarks in the lattice, since the lattice spacing has to be chosen to be smaller than the inverse of the heavy quark, and the simulation of a hadron of typical size $1/\Lambda_{QCD}$ already requires very large lattices. Another complication is in the simulation of objects with large momentum. This makes impossible to compute form factors

for decays of heavy mesons in regions where the light meson has a large momentum in the rest frame of the decaying meson. This condition restricts LQCD calculations to high values of q^2.

2.3 The decay $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$

The main motivation in the study of charmless semileptonic B decays is the determination of the CKM matrix element $|V_{ub}|$. However, for that to be possible the form factors from a reliable calculation, such as LCSR or LQCD, must be available. This is not yet possible for all exclusive channels. At present, there is a discrepancy at the 3σ level between the values of $|V_{ub}|$ obtained from the inclusive and exclusive approach. According to the Heavy Flavor Averaging Group (HFLAV) [41], $|V_{ub}| = (4.52 \pm 0.15_{\text{exp}} \pm 0.13_{\text{theo}}) \times 10^{-3}$ from the inclusive determination, whereas $|V_{ub}| = (3.65 \pm 0.14) \times 10^{-3}$ from the most precise exclusive measurement that comes from the $B \to \pi\ell\nu_\ell$ channel. This discrepancy has given rise to different conjectures, one of them being the revision of the theoretical inputs used in the description of inclusive charmless semileptonic B decays. Currently, there are only five exclusive channels ($B \to (\pi, \eta, \rho, \omega, \eta')\ell\nu_\ell$) that have been measured, representing about 22% of the total rate for inclusive charmless semileptonic B decays. This fact demands the study of other exclusive channels in order to understand the composition of the inclusive rate and consequently improve the inclusive determination of $|V_{ub}|$.

The exclusive $B \to \rho\ell\nu_\ell$ channel, a resonant mode in the $B \to \pi^+\pi^-\ell\nu_\ell$ decay, has also been used to determine $|V_{ub}|$, but the central values obtained from this channel are usually lower than those from $B \to \pi\ell\nu_\ell$ [42]. For instance, the BaBar collaboration analysis [43] measures $|V_{ub}| = (2.76 \pm 0.21) \times 10^{-3}$. Recent discussions about this peculiarity are found in the literature [44–46], with the consensus that previous analyses did not take into account the broad ρ width. This is important as different analyses only consider a limited range in the $\pi\pi$ invariant mass and the exclusive channels suffer from large model dependence. The latest measurement [47] carried out by Belle on this channel determines $|V_{ub}| = (3.68 \pm 0.14^{+0.29}_{\text{exp}-0.34\ \text{theo}}) \times 10^{-3}$, which is in agreement with the value extracted from the $B \to \pi\ell\nu_\ell$ channel. For the latter, a mass window around the nominal mass of the ρ meson of twice the decay width is used. There has also been suggestions from theorists to treat more rigorously the background underneath the mass of the ρ meson, especially those coming from other contributions from $\pi\pi$ states, such as s-wave or other p-wave contributions[11]. To avoid dealing with the identification of resonant or nonresonant contributions, the decay $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ can be studied as a four-body decay, to provide a model-independent determination of $|V_{ub}|$ in the future [46]. Extending the accessible region of phase space beyond of that dominated by the ρ meson will help theorists and experimentalists to better understand the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$, especially to improve the description of the $B \to \rho\ell\nu_\ell$ decay beyond the approximation of narrow width and flat nonresonant background [48]. This information will also improve the calculation of the $B \to \pi\pi$ form factors, that are also essential hadronic input for other processes such as the rare flavor-changing-neutral-current decay $B \to \pi\pi\ell^+\ell^-$ and to hadronic decays such as $B \to \pi\pi\pi$ [49].

The kinematics of the four-body decay $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ is described by the following phase space variables [46]:

1. the effective mass squared of the pion pair $s = M^2_{\pi\pi}$;
2. the effective mass squared of the dilepton system also known as momentum transfer $q^2 = M^2_{\ell\nu_\ell}$;

[11] In a partial wave formalism a $\pi\pi$ state can be viewed as a superposition of different configurations states characterized by an angular momentum, for example the ρ meson is a vector meson with angular momentum $j = 1$, that corresponds to a p-wave configuration of the $\pi\pi$ mass spectrum.

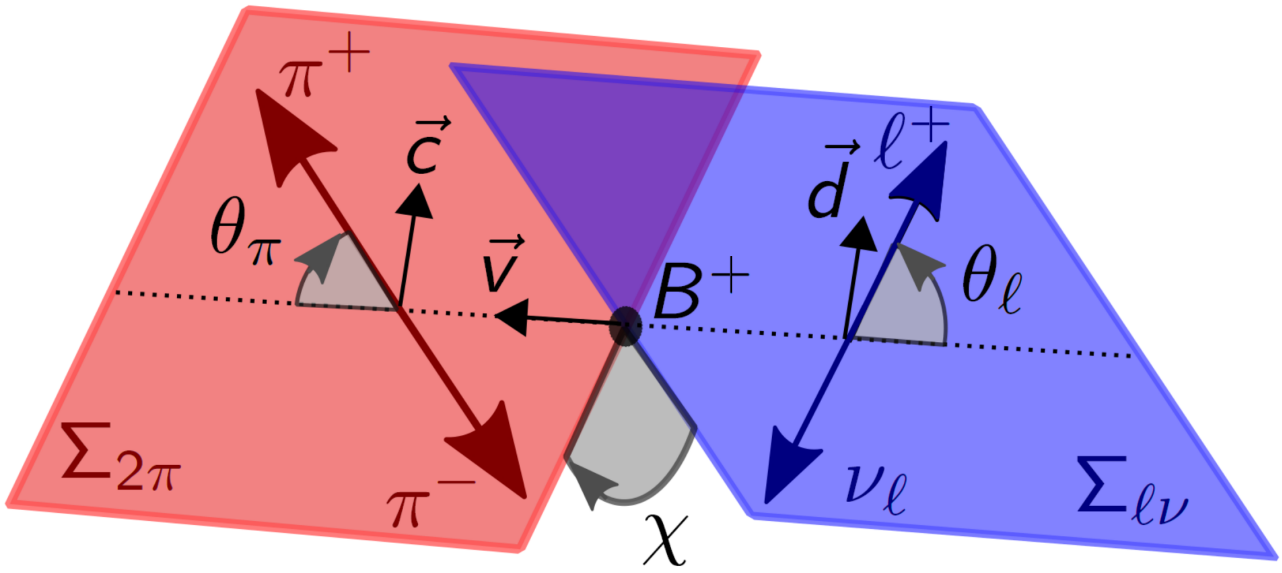


Figure 2.7: Kinematics in the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$. The $\pi^+\pi^-$ moves along the $z-$axis in the B^+ rest frame. The angle θ_π (θ_ℓ) is defined in the $\pi^+\pi^-$ (lepton pair) rest frame as the angle between the $z-$axis and the flight direction of π^+ (ℓ^-), respectively. The azimuth angle χ is the angle between the $\pi^+\pi^-$ decay and lepton pair planes. The unitary vectors $\vec{v}$, $\vec{c}$ and $\vec{d}$, used in the calculation of the phase space variables, are also shown.

3. the angle θ_π of the π^+ in the $\pi^+\pi^-$ center-of-mass frame $\Sigma_{2\pi}$ with respect to the dipion line of flight in the B^+ rest frame Σ_B;

4. the angle θ_ℓ of the charged lepton ℓ in the dilepton center-of-mass system $\Sigma_{\ell\nu}$ with respect to the dilepton line of flight in Σ_B;

5. the angle χ between the dilepton and dipion planes.

The angles θ_π and θ_ℓ are polar, while χ is azimuthal. An illustration of the choice of the angles and the vectors involved is presented in Fig. 2.7. To specify the above variables more precisely, let p_+, p_-, p_ℓ and p_ν be the four-momenta of the π^+, π^-, ℓ^+ and ν_ℓ, respectively. The three-momentum of the π^+ in $\Sigma_{2\pi}$ and that of the lepton in $\Sigma_{\ell\nu}$ are denoted by $\vec{p}_+$ and $\vec{p}_\ell$, respectively. Furthermore, let $\vec{v}$ be a unit vector along the direction of flight of the dipion in Σ_B and $\vec{c}(\vec{d})$ a unit vector along the projection of $\vec{p}_+(\vec{p}_\ell)$ perpendicular to $\vec{v}(-\vec{v})$. Hence,

$$\begin{aligned} \vec{c} &= (\vec{p}_+ - (\vec{v}\cdot\vec{p}_+)\vec{p}_+) \,/\, \left[(\vec{p}_+)^2 - (\vec{v}\cdot\vec{p}_+)^2\right]^{1/2}, \\ \vec{d} &= (\vec{p}_\ell - (\vec{v}\cdot\vec{p}_\ell)\vec{p}_\ell) \,/\, \left[(\vec{p}_\ell)^2 - (\vec{v}\cdot\vec{p}_\ell)^2\right]^{1/2}. \end{aligned} \tag{2.35}$$

With these definitions the phase space variables become [50]

$$\begin{aligned} s &= (p_+ + p_-)^2, & q^2 &= (p_\ell + p_\nu)^2, \\ \cos\theta_\pi &= \vec{v}\cdot\vec{p}_+/\,|\vec{p}_+|, & \cos\theta_\ell &= -\vec{v}\cdot\vec{p}_\ell/\,|\vec{p}_\ell|, \\ \cos\chi &= \vec{c}\cdot\vec{d}, & \sin\chi &= (\vec{c}\times\vec{v})\cdot\vec{d}. \end{aligned} \tag{2.36}$$

Defining the combination of four-vectors $P = p_+ + p_-$, $Q = p_+ - p_-$ and $L = p_\ell + p_\nu$, the hadronic matrix elements can be written in terms of form factors according to

$$\begin{aligned}&\left\langle \pi^+(p_+)\pi^-(p_-)\right| \bar{u}\gamma_\mu(1-\gamma_5)b \left|B^-(p_B)\right\rangle \\ &= -\frac{i}{m_B}\left(P_\mu F + Q_\mu G + L_\mu R\right) - \frac{H}{m_B^3}\epsilon_{\mu\nu\rho\sigma}L^\nu P^\rho Q^\sigma,\end{aligned} \tag{2.37}$$

where p_B and m_B correspond to the four-momentum and mass of the B meson, respectively, and the Levi-Civita symbol follows the convention $\epsilon_{0123} = 1$. The first three terms correspond to the axial current part, while the last term corresponds to the vector current [46]. The dimensionless form factors F, G, H and R are analytic functions of s, q^2, and $t - u$[12]. An alternative set of form factors is given by

$$\begin{aligned}F_1 &= X \cdot F + \sigma_\pi (PL)\cos\theta_\pi G, & F_2 &= G,\\ F_4 &= -(PL)F - q^2 R - \sigma_\pi X \cos\theta_\pi G, & F_3 &= H,\end{aligned} \tag{2.38}$$

where,

$$\begin{aligned}(PL) \equiv P\cdot L &= \frac{m_B^2 - s - q^2}{2},\\ \sigma_\pi &= \sqrt{1 - \frac{4M_\pi^2}{s}},\\ X &= \frac{1}{2}\lambda^{1/2}(m_B^2, s, q^2),\end{aligned} \tag{2.39}$$

with $\lambda(a,b,c) = a^2 + b^2 + c^2 - 2(ab + ac + bc)$. Upon integrating over the angles χ and θ_ℓ, the decay rate can be written as

$$d\Gamma = G_F^2 |V_{ub}|^2 N(s, q^2) J_3(s, q^2, \theta_\pi) ds dq^2 d\cos\theta_\pi, \tag{2.40}$$

with

$$J_3(s, q^2, \theta_\pi) = \frac{2 + z_\ell}{3}|F_1|^2 + z_\ell |F_4|^2 + \frac{(2 + z_\ell)\sigma_\pi^2 s q^2}{3}\left(|F_2|^2 + \frac{X^2}{m_B^4}|F_3|^2\right)\sin^2\theta_\pi, \tag{2.41}$$

and

$$\begin{aligned}N(s, q^2) &= \frac{(1 - z_\ell)^2 \sigma_\pi X}{2(4\pi)^5 m_B^5},\\ z_\ell &= \frac{m_\ell^2}{q^2}.\end{aligned} \tag{2.42}$$

The form factors can be expanded in terms of partial wave functions as follows

[12] The Mandelstam variables t and u are defined as $t = (p_B - p_+)^2$ and $u = (p_B - p_-)^2$

$$
\begin{aligned}
F_1 &= X \sum_{l\geq 0} P_l(\cos\theta_\pi) f_l, & F_2 &= \sum_{l\geq 1} P_l'(\cos\theta_\pi) g_l, \\
F_4 &= \sum_{l\geq 0} P_l(\cos\theta_\pi) \widetilde{r}_l, & F_3 &= \sum_{l\geq 1} P_l'(\cos\theta_\pi) h_l, \\
\widetilde{r}_l &= -((PL) f_l + q^2 r_l),
\end{aligned}
\tag{2.43}
$$

here, the partial waves f_l, g_l, h_l and r_l are functions of s and q^2. The expression for the decay width can be simplified by assuming $z_\ell \ll 1$, and that only $S-$ and $P-$waves contribute. Since the partial waves do not depend on $\cos\theta_\pi$, the decay rate can be integrated over $\cos\theta_\pi$ to obtain

$$
\frac{d\Gamma}{ds dq^2} = G_{\mathrm{F}}^2 \left|V_{ub}\right|^2 N(s,q^2) J_2(s,q^2), \tag{2.44}
$$

where

$$
\begin{aligned}
J_2(s,q^2) &= \int_{-1}^{1} d\cos\theta_\pi J_3(s,q^2,\theta_\pi) \\
&= \frac{4X^2}{3}\left(\left|f_0(s,q^2)\right|^2 + \frac{1}{3}\left|f_1(s,q^2)\right|^2\right) + \frac{8}{9}\sigma_\pi^2 s q^2 \left(\left|g_1(s,q^2)\right|^2 + \frac{X^2}{m_B^4}\left|h_1(s,q^2)\right|^2\right) + \cdots .
\end{aligned}
\tag{2.45}
$$

Here, $f_0(s,q^2)$ is the S-wave form factor and $f_1(s,q^2)$, $g_1(s,q^2)$ and $h_1(s,q^2)$ correspond to the P-wave form factors. This parametrization indicates that a measurement of the decay rate for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ exclusive channel can be carried out as a function of the momentum transfer q^2 and the invariant mass $M_{\pi\pi}$.

CHAPTER 3

EXPERIMENTAL DESIGN

The study of relatively rare phenomena such as the Cabibbo suppressed charmless semileptonic decays of B mesons requires a large sample size of these mesons. For this purpose, high-luminosity e^+e^- colliders have been developed, such that with the right choice of the beam energy, they produce copious amount of B mesons through the process $e^+e^- \to \Upsilon(4S) \to B\bar{B}$. These machines, also known as "B-factories", are the CLEO experiment at the CESR collider at Cornell, USA, the Belle experiment at the KEKB[1] collider at Tsukuba, Japan, and the BaBar experiment at the PEP-II collider at Stanford, USA. BaBar and Belle collected data for over a decade, reaching a total accumulated luminosity of 530.82 fb^{-1} and 1052 fb^{-1}, respectively.

The original design goal of the B-factories was to observe the CP violation in the B meson system, which led to the 2008 Nobel prize in physics being granted to Kobayashi, Maskawa and Nambu. Other physics objectives include precision measurements of the magnitude of elements of the CKM matrix and tests of its unitarity, observation of rare B meson decays and searches for physics beyond the standard model (SM). Besides topics pertaining to B physics, the Belle experiment has also contributed in other research areas, such as charm physics, tau lepton physics, hadron spectroscopy and two-photon physics. The size of the Belle collaboration reached a maximum in 2012, with about 470 members from 72 institutions in 16 countries [35]. The upgrade of the Belle experiment, Belle II, is scheduled to start data acquisition in February 2019, and expected to reach a total accumulated luminosity of 50 ab^{-1}. The physics program covers a wide range of topics, including direct CP violation, lepton flavor violation and dark matter [51]. The Belle detector is designed to measure final state particles, stable enough within the dimensions of the apparatus, to reconstruct B mesons and hence study the physical parameters of interest.

This chapter is organized as follows: section 3.1 describes the KEKB accelerator, section 3.2 presents the main features of the components of the Belle detector, section 3.3 discusses the methods used to identify particles based on their signatures in the detector components, and finally section 3.4 gives an introduction to the trigger system and data processing at the Belle collaboration.

3.1 The KEKB Accelerator

The KEKB accelerator is an asymmetric energy electron-positron collider with only one collision or interaction point (IP) around which the Belle detector is placed. A schematic diagram of this accelerator

[1] KEK is the Japanese acronym for High Energy Accelerator Research Organization.

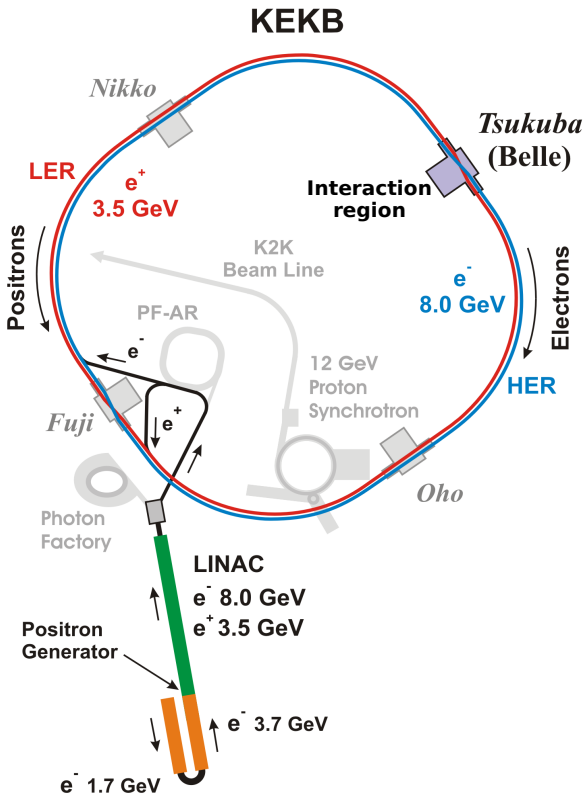


Figure 3.1: Schematic diagram of the KEKB accelerator: electrons and positrons are accelerated in the LINAC and injected in the HER and LER, respectively. The two beams cross at the interaction region where the Belle detector is located. Figure available via Wikimedia Commons under the creative commons CC0 1.0 universal public domain dedication.

is presented in Fig. 3.1. The energy asymmetry adds a forward boost factor ($\beta\gamma = 0.425$) of the center of mass system (CMS) frame relative to the laboratory frame that results in the displacement of the two B meson decay vertices, allowing measurements of their time evolution. In contrast, at a symmetric collider like CESR, the B mesons are produced almost at rest, and therefore they decay almost in the same position. This hinders establishing the exact moment at which each meson decays and thus the observation of CP violation. The KEKB accelerator consists of a linear accelerator (LINAC) and two 3016 m circumference rings; the LINAC injects 8 GeV electrons in the High Energy Ring (HER) and 3.5 GeV positrons in the Low Energy Ring (LER), which are characterized by high beam currents ($I_{\mathrm{HER}} = 1.1$ A and $I_{\mathrm{LER}} = 2.6$ A) and small beam sizes at the IP (in the horizontal direction $\sigma_x = 90$ μm and in the vertical direction $\sigma_y = 1.9$ μm) [52]. The choice of energies for the beams results in a center-of-mass energy that coincides with the mass of the $\Upsilon(4S)$ resonance, $\sqrt{s} = 10.58$ GeV. Contrary to other B-factories where the beams collide head on, at KEKB the beams collide at a crossing angle of 22 mrad, thus avoiding parasitic collisions and the use of bending magnets for beam separation, hence increasing the luminosity [53]. The luminosity was designed to be 1.0×10^{34} cm^{-2}s^{-1} and the KEKB accelerator exceeded this value on June 2009, achieving a world-record instantaneous luminosity of 2.11×10^{34} cm^{-2}s^{-1} [54]. The Belle detector accumulated a total integrated luminosity of 1040 fb^{-1} during the years 1999 until 2010, with data collected at the energy of the $\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(3S)$, $\Upsilon(4S)$ and $\Upsilon(5S)$ resonances as well as

Resonance	On-peak luminosity [fb^{-1}]	Off-peak luminosity [fb^{-1}]	Number of resonances
$\Upsilon(1S)$	5.7	1.8	102×10^6
$\Upsilon(2S)$	24.9	1.7	158×10^6
$\Upsilon(3S)$	2.9	0.25	11×10^6
$\Upsilon(4S)$	711	89.4	772×10^6 $B\bar{B}$
$\Upsilon(5S)$	121.4	1.7	7.1×10^6 $B_S\bar{B}_S$
Scan		27.6	

Table 3.1: Integrated luminosity accumulated by Belle, categorized by CM energy [55].

off-resonance data taken 60 MeV below the resonance peak of the $\Upsilon(4S)$ to study non-$B\bar{B}$ background. The corresponding integrated luminosity for each CM energy is listed in Table 3.1.

3.2 The Belle Detector

The Belle detector is a general purpose magnetic spectrometer built around a 1.5 T superconducting solenoid and iron structure. It is located at the interaction region (surrounding the IP) and has a cylindrical onion-like structure. Given the asymmetry of the collider, most of the components of the detector are oriented towards the forward direction. The adopted coordinate system is defined such that the z-axis is in the direction of the magnetic field, opposite to the positron beam direction, and the origin coincides with the IP. Because of the crossing angle between the beams, the z-axis is not exactly the same as the electron direction. The x and y axes correspond to the horizontal and vertical directions, respectively. It is also convenient to define a spherical coordinate system (r, θ, ϕ), with θ measured with respect to the z-axis. A depiction of the Belle experimental apparatus with its components can be seen in Fig. 3.2. In the following I briefly describe every component of the Belle detector. A more detailed explanation can be found in Ref. [56].

3.2.1 Silicon Vertex Detector (SVD)

The SVD is the innermost component of the Belle detector, whose main purpose is to provide information about the spatial distribution of charged particles. This information is of prime importance for time-dependent CP violation measurements, where the B meson decay vertices have to be precisely determined. Since the B mesons have a short decay length (around 200 μm at KEKB), the SVD has to be located as close to the IP as possible. The fundamental module of the SVD is the double sided silicon strip detector (DSSD) sketched in Fig. 3.3. Each DSSD has a thickness of 300 μm and consists of an n-bulk silicon with implantations of p-doped strips on one side, parallel to the beam, for measurements in the ϕ direction and n-doped strips on the other side, perpendicular to the beam, for measurements of the z-coordinate [57]. When charged particles traverse the detector, they create electron-hole pairs by means of the Coulomb interaction. These electrons and holes drift to the strips, which act as charge collecting electrodes. By connecting the strips to a charge-sensitive amplifier, the position of the electron-hole pair can be deduced.

During data acquisition, Belle used two configurations of the SVD. In an early stage the SVD1 consisted of three concentric cylindrical layers of DSSD, with radii of 30 mm, 45.5 mm and 60.5 mm, and covered an angular range of $23^\circ < \theta < 139^\circ$. The z-vertex resolution was 100 μm. Figure 3.4 shows the geometrical configuration of the SVD1. Due to the massive radiation damage, the SVD1 was replaced in

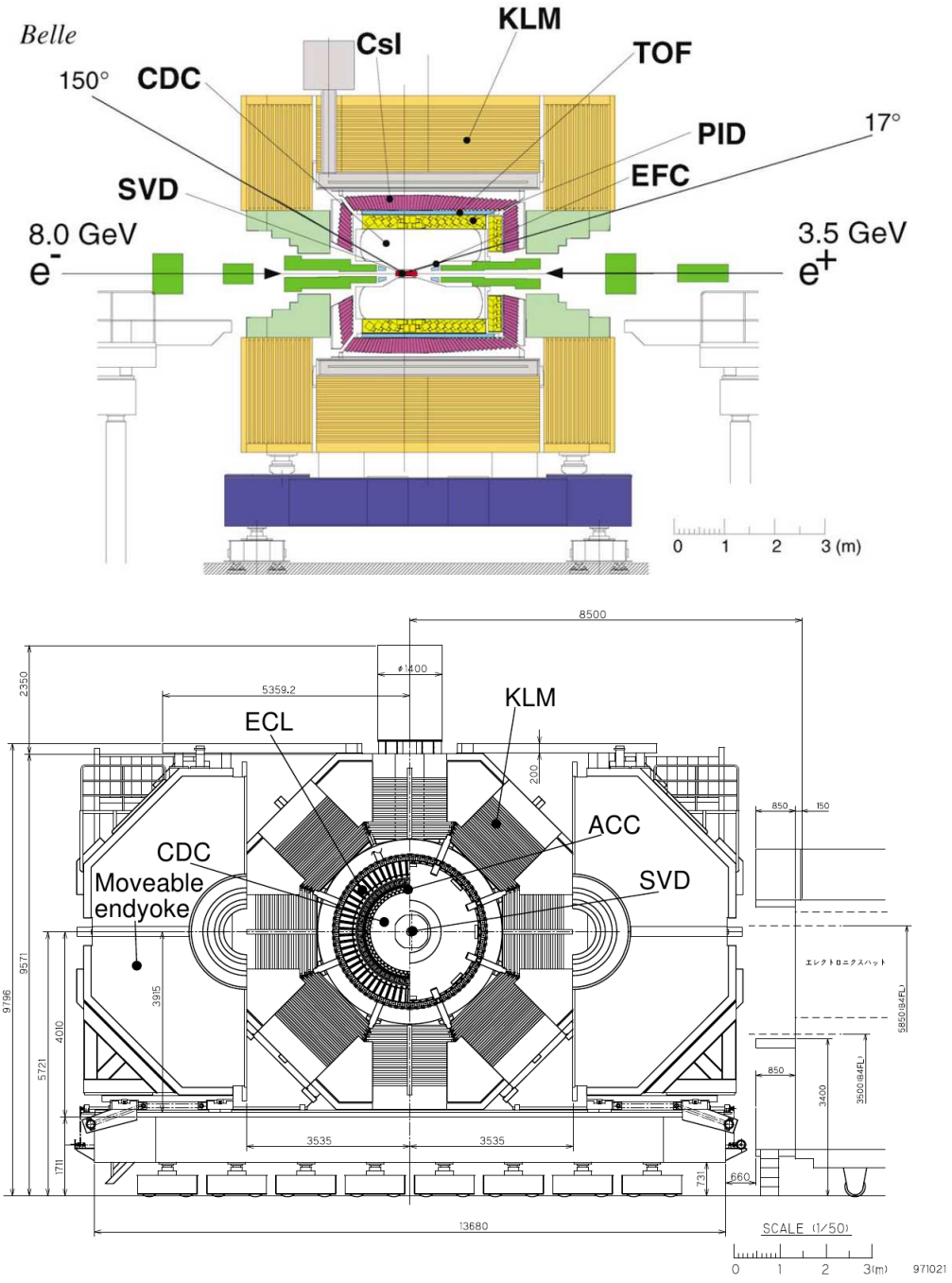


Figure 3.2: Longitudinal [56] (top) and transverse [35] (bottom) cross sections of the Belle detector with its components. Starting from the innermost to the outermost components, the detector consists of the silicon vertex detector (SVD), central drift chamber (CDC), aerogel Cherenkov counter (ACC), time of flight counters (TOF), electromagnetic calorimeter (ECL) made out CsI crystals, extreme forward calorimeter (EFC) and the kaon long and muon detector (KLM).

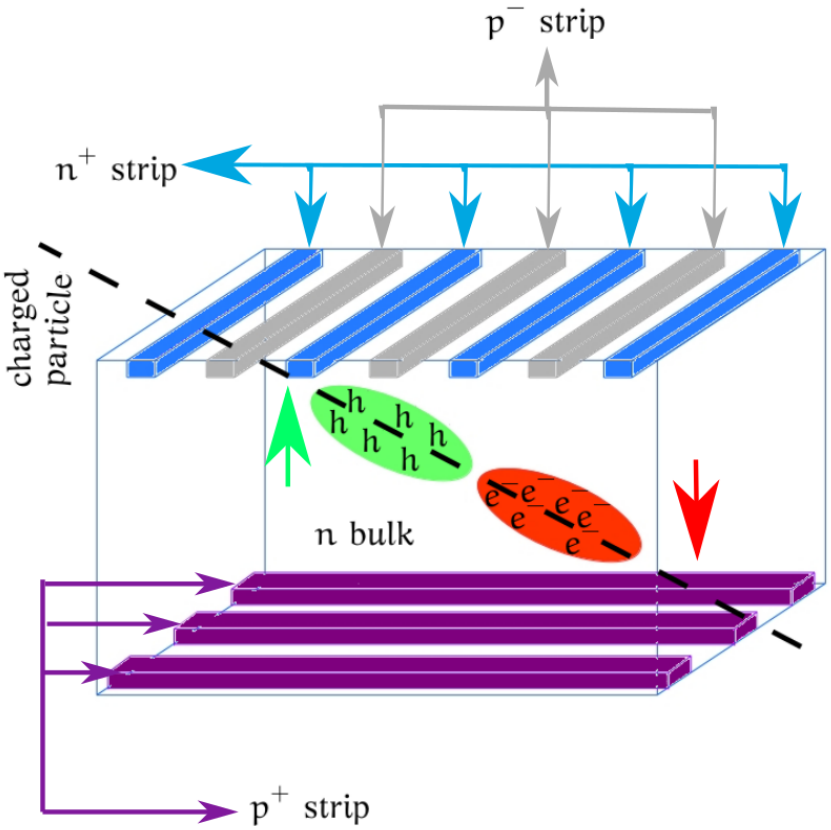


Figure 3.3: Diagram of the physics principle of a double-sided silicon strip detector, showing the creation of electron-hole pairs as a charged particle traverse the DSSD.

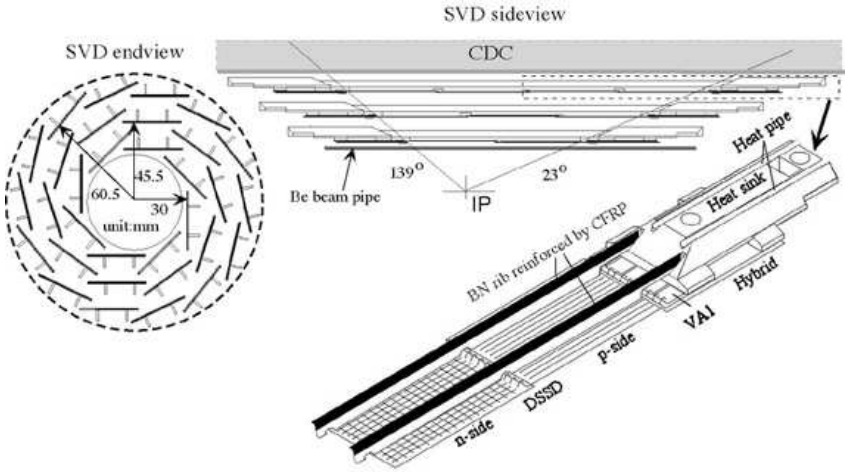


Figure 3.4: Side and end view of the SVD1 [56].

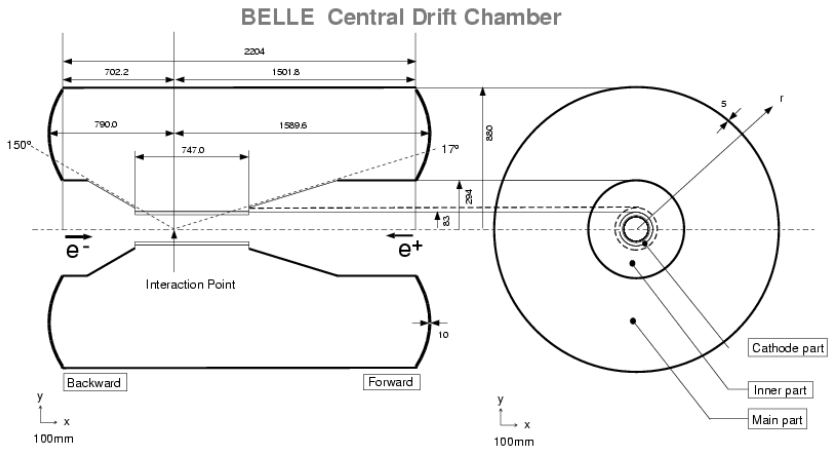


Figure 3.5: Structure of the Belle Central Drift Chamber [56], the main tracking device of the Belle detector that allows to determine the trajectory and momentum of the charged particles as well as measure the energy loss by ionization.

2003 by the SVD2, which added a fourth layer and reduced the radius of the inner layer. Hence the SVD2 has four concentric layers with radii of 20 mm, 43.5 mm, 70 mm and 88 mm. This results in an improved angular coverage ($17^\circ < \theta < 150^\circ$) and a z-vertex resolution of ~ 80 μm.

To get an indication of the performance of the SVD, the momentum and polar angle dependent resolution for the impact parameters have to be determined. The impact parameters are defined as the distance of closest approach to the IP in the $r\phi$ plane (dr) and in the z-direction (dz). They are obtained from cosmic ray events obtained simultaneously during the collision data acquisition [58]. These resolutions are given below for the two SVD configurations:

$$\text{SVD1}: \quad \sigma_{dr} = 19.2 \oplus \frac{54}{p\beta \sin^{3/2}\theta}[\mu\text{m}], \quad \sigma_{dz} = 42.2 \oplus \frac{44.3}{p\beta \sin^{5/2}\theta}[\mu\text{m}],$$
$$\text{SVD2}: \quad \sigma_{dr} = 21.9 \oplus \frac{35.5}{p\beta \sin^{3/2}\theta}[\mu\text{m}], \quad \sigma_{dz} = 27.8 \oplus \frac{31.9}{p\beta \sin^{5/2}\theta}[\mu\text{m}].$$

In the above expressions, σ_x represents the resolution of a given impact parameter, p, β and θ are the momentum (in GeV/c), velocity and polar angle of the track, respectively. The symbol $\oplus$ indicates a quadratic sum.

3.2.2 Central Drift Chamber (CDC)

The CDC is the main tracking device in the Belle detector, and performs three major functions [35]. First, it reconstructs the trajectory of charged particles, determines their momenta from their curvature in the magnetic field and measures their hit coordinates in the detector volume. Second, it provides information on the type of charged particle from measurements of specific energy loss by ionization (dE/dx). Finally, it supplies fast trigger signals for charged particles.

Figure 3.5 shows the setup of the CDC: it is a cylindrical wire chamber consisting of 50 layers of anode wires and three cathode strip layers. The anode wires form groups of three to six layers of axial wires, parallel to the beam, to measure p_T, followed by layers of three to four stereo wires slightly tilted to an angle with respect to the beam axis between -58 mrad and 72 mrad to give the polar angle θ [59].

The cathode strips are azimuthally segmented in eight sectors and divided into 64 to 80 strips for a total of 1792 channels. This segmentation provides additional information to precisely determine the z-coordinate [60]. The CDC is 2400 mm long with an inner and outer radii of 83 mm and 874 mm, respectively. It contains 8400 almost square drift cells made up of a sense wire with positive voltage, surrounded by eight field wires with negative voltage. It is filled with a low atomic number (Z) gas mixture of 50% ethane (C_2H_6) and 50% helium (He) to minimize Coulomb scattering. Low-Z gases are useful in reducing the background due to synchrotron radiation given their small photo-electric cross section [61]. They also improve the resolution of the momentum and dE/dx measurements. When a charged particle impinges the gaseous medium, it ionizes the gas molecules by knocking out electrons that are then attracted to the sense wires by the action of their electric fields. In their way towards the sense wires, the electrons create cascades through secondary ionization, a process known as gas amplification [62]. The collection of these charge carriers in the sense wires induces an electrical signal that can be measured. The drifting time of the electrons to the wires is used for tracking.

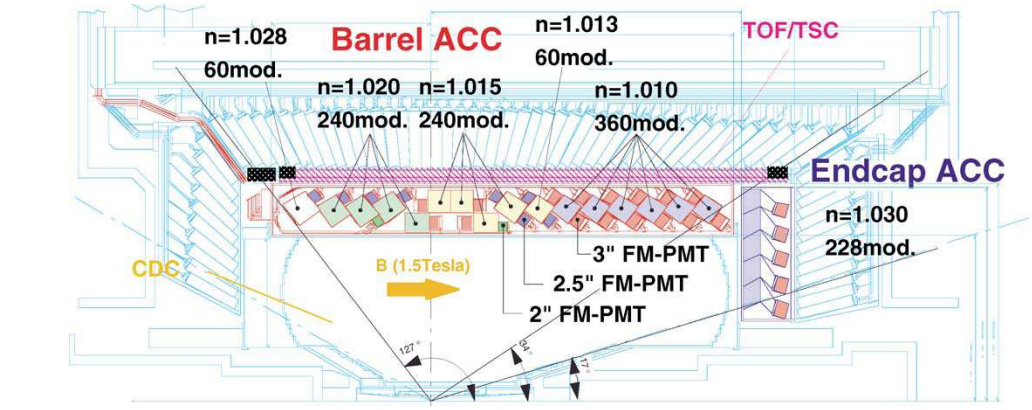
The CDC is built asymmetrically in the z-direction to account for the fact that the decay products of the $\Upsilon(4S)$ are boosted in the forward direction, hence the azimuthal angle acceptance is $17^\circ < \theta < 150^\circ$. It is characterized by an overall spatial resolution of 130 μm, a dE/dx resolution of 6% for Bhabha scattering and μ-pair events and a transversal momentum resolution (when combined with the information of the SVD) given by $\sigma_{p_T}/p_T = 0.0019 p_T[GeV] \oplus 0.0030/\beta$ [35].

3.2.3 Aerogel Cherenkov Counter (ACC)

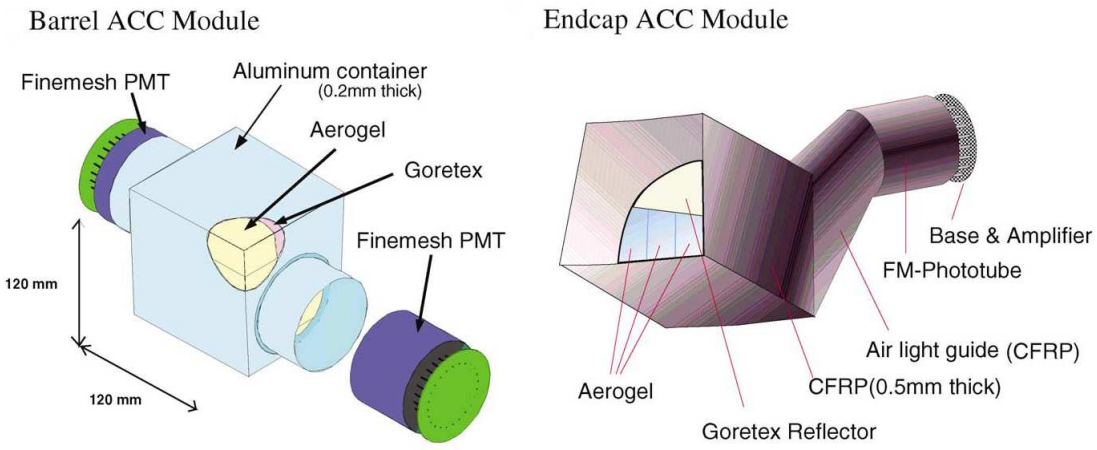
Charged particles moving through an optically transparent medium with a velocity exceeding that of light in the medium emit photons, a process known as Cherenkov radiation [63]. In a medium with index of refraction n, the velocity of light is c/n. The light propagates in a forward cone with a vertex angle $\theta_C = \cos^{-1}(c/vn)$, from which the velocity v of the particle can be inferred. The ACC exploits this radiation effect to provide information about the particle velocity and, together with a momentum measurement from other detectors (e.g. CDC), it is possible determine the mass of the particle and thus the particle identity. This is especially helpful in separating $K^\pm$ from $\pi^\pm$ in the relatively high momentum range of B meson decay products (1.2 GeV/c $\rightarrow$ 3.5 GeV/c). The ACC layout and typical modules employed in the barrel and end-cap regions are presented in Fig. 3.6. The ACC comprises 960 counter modules in the barrel part and 228 counter modules in the end-cap of the detector. Each module consists of an aluminum encased block of silica aerogel, with a refractive index ranging from 1.01 to 1.03, such that the Cherenkov light can be only emitted by particles lighter than the kaon. The aerogels are attached to one or two fine-mesh photomultiplier tubes (FM-PMT) to detect Cherenkov light. The polar angle coverage [35] is $33.3^\circ < \theta < 127.9^\circ$ in the barrel, and $13.6^\circ < \theta < 33.4^\circ$ in the forward region.

3.2.4 Time of Flight Counters (TOF)

The vast majority of the particles in B meson decays have momenta lower than 1 GeV, hence it is crucial for precision measurements to have an optimal separation between kaons and pions in the low momentum range. The TOF system measures the time a charged particle takes to travel from the IP to the system, for which a timing resolution of 100 ps is required. It contains 64 modules along the barrel region, just outside the ACC, spanning a polar angle range of $33^\circ < \theta < 121^\circ$. Each module consists of two TOF counters, made up of plastic scintillators, and one thin trigger scintillation counter (TSC), which are read out by fine-mesh PMTs. A sketch of the TOF/TSC module is displayed in Fig. 3.7.



(a)



(b)

Figure 3.6: (a) Arrangement of the ACC in the Belle detector whose main function is to measure the velocity of the particle from its collected Cherenkov radiation, (b) typical ACC module used in the barrel and end-cap parts of the detector [56]. In the figure PMT and CFRP stand for photomultiplier and carbon-fiber reinforced plastic.

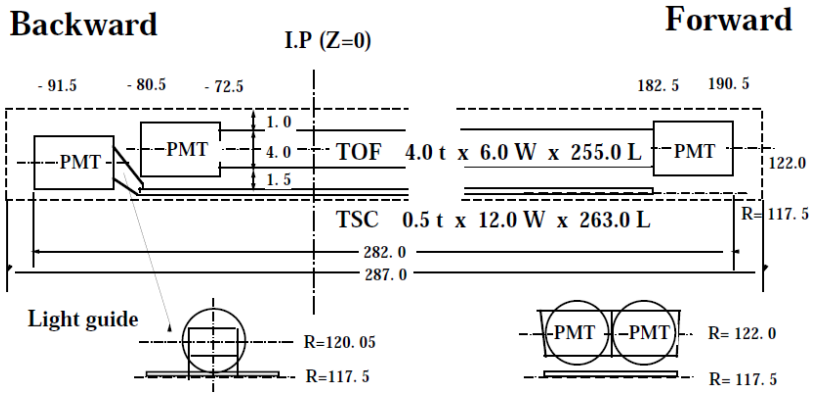


Figure 3.7: Dimensions of TOF/TSC module [56]. The TOF measures the time a charged particle takes to travel from the IP to the system.

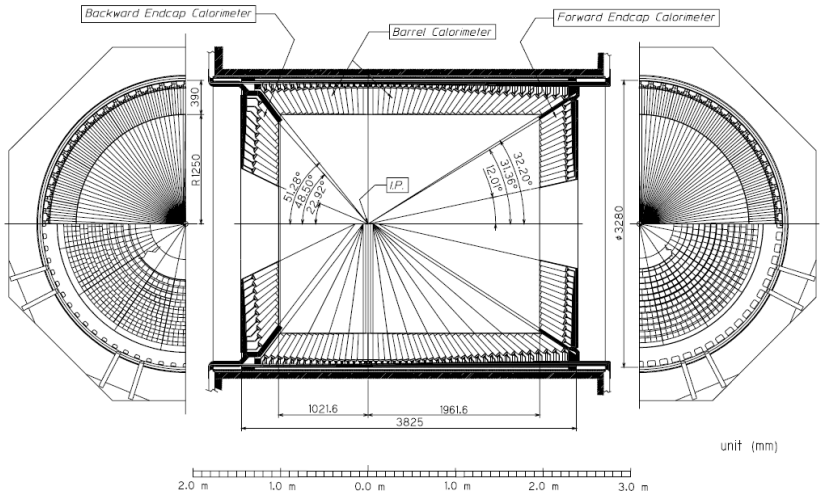


Figure 3.8: Geometrical configuration of the Belle ECL with its three distinct regions: the barrel, the backward and the forward endcaps [56]. The ECL measures the energy of photons and electrons from the scintillation light they produced in the CsI(TL) crystals.

3.2.5 Electromagnetic Calorimeter (ECL)

The ECL measures the direction and energy of photons and electrons by completely absorbing them, and also provides information for electron identification. The operating principle of the ECL relies on the development of electromagnetic showers, the result of interactions of electrons and photons with matter, in which they lose energy via Bremsstrahlung and pair production[2]. These processes generate secondary electrons and photons which in turn interact in a similar manner, producing a cascade of secondary particles whose energies eventually fall below the region dominated by radiation loss and deplete their energy mainly through ionization [64]. These energies are then absorbed by a scintillator and converted into visible light that can then be measured by a photodetector. A typical quantity, intrinsic to a material, is the radiation length (X_0) which describes the longitudinal development of an electromagnetic shower[3]. This quantity is important in the construction of the ECL since it determines the thickness of the detector to completely absorb an electromagnetic shower.

Photons in B decays are mostly products of cascade decays and thus they carry a relatively low energy of the order of 500 MeV. Nonetheless, there are also decays with energetic photons such as $B^0 \to K^{*0}\gamma$ and $B^0 \to \pi^0\pi^0$, reaching photon energies of up to 4 GeV [65]. Moreover, one third of hadronic decays involve π^0 mesons and consequently their reconstruction demands an efficient selection of two nearby photons. These requirements impose a good energy resolution in the range from 20 MeV to 4 GeV. However, there is a 3% loss in acceptance due to the gap between the barrel and end-cap components. The ECL consists of 8736 Thallium doped Cesium iodide, CsI(Tl), crystals, each having a tower shape with about 6 cm × 6 cm cross section and 30 cm length (16.2 X_0). The scintillation light produced by these crystals is read out by a pair of silicon PIN photodiodes.

The Belle ECL includes a 3 m long barrel section with an inner radius of 1.25 m and annular end-caps at $z = 2.0$ m and $z = -1.0$ m from the IP as displayed in Fig. 3.8. It spans the polar angle range of

[2] Bremsstrahlung is the radiation given off by a charged particle due to its acceleration by an electric field or another charged particle. Pair production refers to the transformation of energy into mass, particularly to the generation of a fermion and its anti-fermion from a neutral boson.

[3] It is also defined as the distance in the material at which an electron will radiate a factor of $1/e$ its energy.

$17^\circ < \theta < 150^\circ$, representing a solid angle coverage of 91%. The energy and position resolution are parameterized as [56]:

$$\frac{\sigma_E}{E}[\%] = \left(1.34 \oplus \frac{0.066}{E} \oplus \frac{0.81}{E^{1/4}}\right),$$

$$\sigma_{pos}[\%] = \left(0.27 + \frac{3.4}{E^{1/2}} + \frac{1.8}{E^{1/4}}\ \text{mm}\right),$$

respectively, with E measured in GeV. This detector also measures the shower energy to track momentum ratio (E/p) that is used to distinguish hadrons from electrons.

3.2.6 Extreme Forward Calorimeter (EFC)

The main purpose of the EFC is to extend the angular acceptance of the ECL and thus cover the polar angle regions from 6.4° to 11.5° in the forward direction and 163.3° to 171.2° in the backward direction. Due to its proximity to the IP, the scintillating material used for the EFC is a bismuth germanate crystal (BGO or $Bi_4Ge_3O_2$), which has a high radiation tolerance. Information from the EFC is not employed in event reconstruction. Instead, it is used for determining the integrated luminosity collected by the Belle detector. It also serves as a beam mask to reduce background in the CDC.

3.2.7 $K_L^0\mu$ (KLM)

The KLM is the outermost component of the Belle detector and is placed outside the solenoid magnet. It consists of an octagonal barrel region and two end-cap regions, which have an alternating structure of 47 mm thick iron plates and 44 mm thick active detector element. There are 15 detector layers and 14 iron layers in the barrel region, and 14 detector layers and 14 iron layers in each end-cap. This detector component is based on glass-electrode resistive plate chambers (RPC), which comprises two parallel plate electrodes with a high bulk resistivity separated by a gas-filled gap. Figure 3.9 illustrates the structure of an RPC module. The gas composition has a non-combustible mixture of 62% freon[4] , 30% argon (Ar) and 8% butane-silver[5]. Ionizing particles traversing the gas-filled gap generate a discharge in it that is read out by external pickup strips, and hence supplying the location and time of the ionization. The barrel region covers an angular range from 45° to 125° in the polar angle, while the end-caps extend this range to 20° and 155°.

Hadrons interacting with matter produce a succession of particles as result of decay of excited nuclei. The secondary particles are mostly pions and nucleons, which then interact further developing a shower analogous to the electromagnetic case. Since neutral pions are frequent in hadronic decays, there is also an electromagnetic component present in hadronic showers. The longitudinal profile of these showers in a given material is characterized by the interaction length[6] (λ_I). Thus for completely absorbing a hadron and measure its energy, a dense and thick material must be used. The iron layers in the KLM correspond to 3.9 interactions lengths of material while the ECL adds another 0.8 interaction lengths. Long-lived neutral particles like K_L^0 pass through the inner components of the detector without being observed and produce a shower in the KLM. The location of this shower determines the direction of the K_L^0, but due to fluctuations in the size of the shower, an energy measurement of the K_L^0 is not possible. The multiple layers of active detectors and iron allow the distinction between muons and charged hadrons based upon

[4] CH_2FCF_3 also referred to as HFG-134a.

[5] The butane silver is a mixture of approximately 70% n-butane and 30% iso-butane.

[6] The hadronic interaction length is much larger than the radiation length for a given material.

their range and transverse scattering. On that account, a charged particle is identified as a muon if its trajectory is barely deflected in the KLM and leaves a corresponding track in the CDC.

3.3 Particle Identification

The final state particles in B meson decays, stable enough to be observed by the Belle detector, are the $e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$, protons, neutrons and photons. Hence we require a sufficient separation among these particles to correctly measure a particular decay channel. For each charged particle, a probability density function (PDF) is calculated for every component of the Belle detector (i.e., CDC, ACC, TOF, ECL, SVD and KLM). The PDFs are then combined to produce a likelihood to account for the particle hypothesis (i.e., $e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$), for which this output peaks at values close to unity. A description of the variables employed for particle identification is discussed below.

3.3.1 Electron identification

To associate a certain track and calorimeter cluster with an electron, the following discriminating observables are used:

- The ratio of the energy deposited in the ECL and the momentum measured by the CDC, E/p. As the mass of the electrons is negligible compared to the other final state particles, this value is close

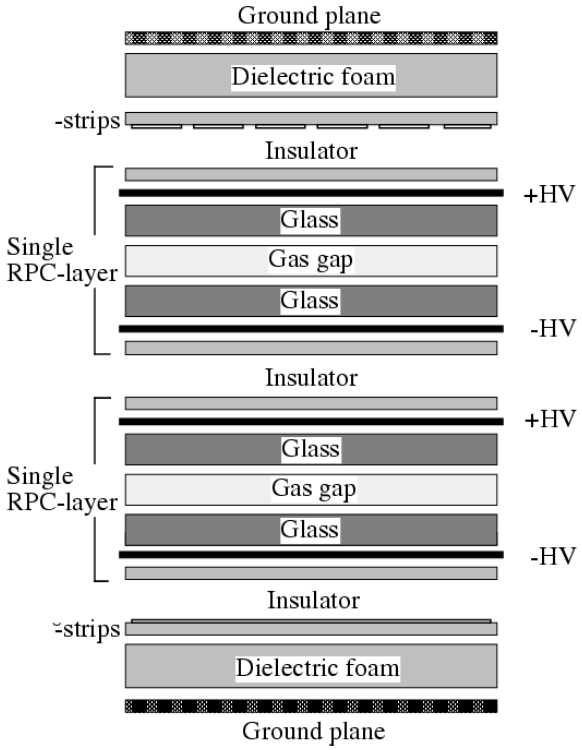


Figure 3.9: Cross section of a double-gap RPC module [56].

to unity for electrons.

- The transverse shower shape in the ECL, which is given by the ratio of the energy summed in a 3×3 array of crystals (`E9`), surrounding the center of the shower to that of a 5×5 array of crystals around the same center. This variable peaks at 0.95 for electrons, while for pions tends to be at lower values.
- The specific energy loss dE/dx measured in the CDC
- The matching between a cluster in the ECL and the charged track position extrapolated to the ECL.
- The light yield in the ACC.
- The time of flight measured by TOF

A PDF is assigned to each discriminant to form an electron ($\mathcal{L}_e$) and non-electron ($\mathcal{L}_{\bar{e}}$) likelihoods. They are subsequently combined into a likelihood for identifying an electron given by

$$\mathcal{L}_{eid} = \frac{\prod_i \mathcal{L}_e^i}{\prod_i \mathcal{L}_e^i + \prod_i \mathcal{L}_{\bar{e}}^i} \tag{3.1}$$

where i runs over the electron input parameters [66].

3.3.2 Muon identification

For muon identification, information from the SVD, CDC and KLM are used to distinguish between muons and hadrons. In this case, two quantities are defined to produce the PDFs: ΔR, the difference between the measured and expected range of the track, and χ^2_r, the goodness of fit of the transverse deviations of all hits associated with the track (normalized by the number of hits) from the re-extrapolated crossings. Since the these two quantities are uncorrelated, the joint PDF is a product of these separate PDFs. The muon likelihood is given by

$$\mathcal{L}_{\mu id} = \frac{\mathcal{L}_\mu}{\mathcal{L}_\mu + \mathcal{L}_\pi + \mathcal{L}_K}, \tag{3.2}$$

where $\mathcal{L}_\mu$, $\mathcal{L}_\pi$ and $\mathcal{L}_K$ are the PDFs for muons, pions and kaons, respectively [67].

3.3.3 Identification of pions and kaons

To separate pions from kaons the combined information of the following measurements are needed:

- dE/dx by the CDC
- the number of photoelectrons in the ACC
- time of flight from the TOF

The product of the three likelihood functions yields the global probability for a track to correspond to either a kaon (P_K) or a pion (P_π). Thence, a particle is identified as a kaon or a pion by selecting a value on the likelihood ratio (PID) [56]:

$$\texttt{Prob}(K) = \frac{P_K}{P_K + P_\pi}, \tag{3.3}$$

for example tracks with values of $\texttt{Prob}(K)$ close to unity are regarded as kaons, while tracks with values close to zero are identified as pions. This is the standard likelihood used at Belle for which values above or below a certain probability X are accepted as either kaons or pions. Another convention is to calculate a likelihood for which values close to unity corresponds to pions given by

$$\texttt{Prob}(\pi) = 1 - \texttt{Prob}(K). \tag{3.4}$$

3.4 Trigger

In particle collisions different processes occur simultaneously, leading to the generation of a signal in the detector, albeit most of these events are of no interest. Furthermore, there is a limiting factor on the storage capacity and the data taking rate for which these additional events must be discarded. To select a particular reaction certain criteria must be imposed that once satisfied, activate other operations in the system, such as the recording instruments. The electronic logic required to carry out this procedure is referred to as trigger [68]. The cross-sections and trigger rates for different physical processes of interest at the energy scale of the $\Upsilon(4S)$ resonance are listed in table 3.2. The Belle trigger system retains some of the Bhabha and $\gamma\gamma$ events for luminosity measurement and detector calibration. It consists of three stages: Level-1 (L1) hardware trigger, Level-3 (L3) software trigger and the offline Level-4 trigger.

In the L1 trigger, signals from the different subdetectors are provided to the central trigger system called the *global decision logic* (GDL), which decides to trigger on hadronic ($B\bar{B}$ and continuum), Bhabha and $\mu^+\mu^-$ events. For instance, to trigger a hadronic event, at least one of the following conditions has to be met: three or more charged track candidates, high levels of deposited energy in the ECL and four isolated neutral clusters in the ECL. The efficiency of the trigger for hadronic events is greater than 99.5%. A schematic view of the L1 trigger with the information supplied by every subdetector is displayed in Fig. 3.10; at the end of this stage, the GDL takes about 2.2 μs to pass an event to the *data acquisition system* (DAQ).

The next stage is the L3 software trigger that runs online on a computer farm, filters the events using a fast tracking algorithm and subsequently saves them for further analysis. The offline L4 trigger applies additional requirements on the energy measured in the ECL by accepting events with a minimum energy of 4 GeV and at least one track stemming from the IP. These events are then stored in files called *data summary tapes* (DST) that are available to all Belle users. To perform specific analysis, the data is skimmed to guarantee an easy access to a given channel in the form of minimal sets of DST (MDST).

Physics process	Cross section [nb]	Rate [Hz]
$\Upsilon(4S) \to B\bar{B}$	1.2	12
Hadron production from continuum	2.8	28
$\mu^+\mu^-$ and $\tau^+\tau^-$	1.6	16
Bhabha $\theta_{\text{lab}} > 17^\circ$	44	4.4
$\gamma\gamma$ ($\theta_{\text{lab}} > 17^\circ$)	2.4	0.24
2γ process ($\theta_{\text{lab}} > 17^\circ$, $p_t \geq 0.1$ GeV)	15	35

Table 3.2: Cross sections of e^+e^- collisions and trigger rates for different processes of interest for physics analysis at the $\Upsilon(4S)$ resonance [56].

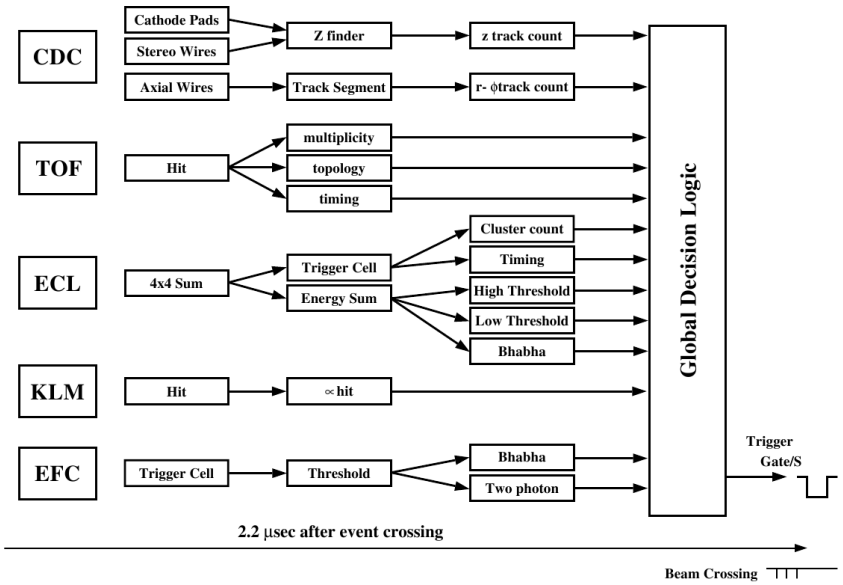


Figure 3.10: The Level-1 trigger system for the Belle detector [56].

CHAPTER 4

ANALYSIS TECHNIQUES

In high energy physics analyses, the goal is to disentangle processes that contain the physics of interest, referred to as *signal*, from other processes known as *background*. As the signal typically represents a tiny fraction of the total events, the focus of the analysis at first is the reduction of the background by using a binary categorization tool whose aim is to increase the probability that a given signal candidate lies within the selected sample. The selection of signal candidates brings with it a trade-off between *efficiency* and *purity*. The former refers to the amount of signal candidates left in the sample after the selection compared to the original number of signal events, whereas the latter shows the fraction of signal candidates with respect to the total events after the selection. If S and B denote the signal and background, respectively, the efficiency ϵ and the purity p can be written as:

$$\epsilon = \frac{S_{\text{selection}}}{S_{\text{total}}}, \quad p = \frac{S_{\text{selection}}}{S_{\text{selection}} + B_{\text{selection}}}. \tag{4.1}$$

Hence, if the signal process is a rare decay, for example, and the selection procedure results in a sample with high purity, meaning with little background, at the moment of measuring the signal in the data one ends up with a sample with little statistics and the inability of extracting meaningful information about the signal. On the other hand, if the selection is highly efficient (which incurs in the presence of more background), depending on how well modeled and similar the background process is to the signal, an observation of the latter in the data may not be possible. Consequently, one must assure a selection process with a balance between purity and efficiency. This chapter presents an overview of the methods employed directly or indirectly in this thesis for that aim. The first part of this chapter, section 4.1, describes some of the typical event classification techniques widely used in particle physics. Section 4.2 introduces the main approaches of B meson reconstruction methods used at the B-factories, with special emphasis in the full reconstruction.

4.1 Event Classification in High Energy Physics

The main task of a binary classification algorithm is to optimize the discrimination between signal and background in the data sample, which is crucial to obtain a signal-enhanced sample to perform precision measurements of physical parameters. The first step is to form a set of observables or variables that characterize the signal, and exhibit a distinct behavior from the background. The next step is to apply a

requirement on the set of variables to decide the likelihood for an event to be identified as a signal or background. Typically the choice of this requirement follows the optimization of a figure of merit (FOM) or physics-motivated reasons over the distribution of the variables. In particle physics jargon, one talks about *cuts* when referring to a selection such that events above (or below) a reference value are more likely to be signal than events below that value. In the following, three classification methods that are used to some extent in this thesis are presented.

4.1.1 Cut based methods

A common practice to select signal events is by applying sequential cuts on a set of input variables, a procedure that can be pictured as cutting away a region of the data sample. In most cases, cuts are applied at the pre-selection stage. For example, the reconstruction of a decay starts by imposing conditions on the event topology, the allowed distance traveled by the decay products from the collision point and the identification of the final state particles such as electrons, muons, pions, etc. Afterwards, different cuts can be applied to kinematic variables to reject background and refine the selection. Depending on the variable, there can be physics motivated cuts where an optimization is not required, for instance typical cuts for the decay $B^+ \to D^0 \ell^+ \nu$ with $D^0 \to K^- \pi^+$ are a 3σ mass window of the D^0 meson usually modeled with a Gaussian, and the number of charged particles in the signal side, in this case three, one lepton and two charged hadrons. In other cases the cut follows an FOM optimization for the signal. For example, the boundaries of the extra energy of photons in the event. Although simple to implement, a selection based solely on cuts can result in a sample with low efficiency and/or purity, inasmuch as it does not take into account the correlations between the input variables. Cutting on two highly correlated variables brings no additional information to the analysis, since the effect of the requirements on these variables can be reduced to a single cut on one of the variables. In practice, most of the variables used in high energy physics analyses are correlated at some degree, implying that cutting on a single variable may throw away events within the phase space of the signal. A more efficient way of selecting these events involve a multidimensional space where the signal can be separated from the background. A simplistic example can be pictured as follows: imagine that the signal and background can be represented as two distinct tilted ellipses in a two dimensional space parameter, but the projection of the events on one variable shows an apparent overlap in the distributions of both the background and the signal. Therefore, when cutting on a single variable, signal events get lost in the selection. To circumvent this issue, one can select the region of the two dimensional space containing the signal ellipsis to maximize the number of signal events. This idea is explored in multivariate analyses to improve the selection.

4.1.2 Multivariate methods

A more sophisticated approach of a classification technique is a multivariate analysis (MVA) that uses the multi-dimensional observable space rather than each observable separately [69]. Its main task is to optimize a mapping function from a set of input variables $\vec{x}$ while taking into account correlations among them and produce a single output discriminator. The categorization of the data as signal or background is then performed as a simple cut on this powerful discriminator. The idea of multivariate techniques is tightly related to the concept of machine learning, which has its origins in the pursuit of artificial intelligence. With the advent of vast amounts of data, the development of automated computer algorithms that allowed learning special features from data was necessary. The goal of the learning process is thus to be able to respond and provide predictions for future data [70]. In supervised machine learning the estimation of the mapping function $f(\vec{x}, \vec{w})$, which is an approximation to the unknown function $f(\vec{x})$, is achieved using a training data set to determine the adjustable parameters $\vec{w}$, also known as weights.

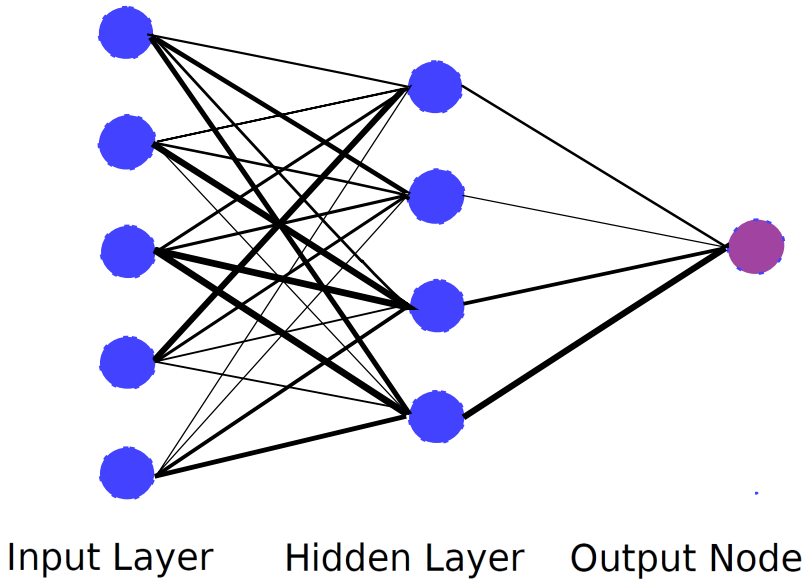


Figure 4.1: Structure of the multilayer perceptron as implemented in the NeuroBayes package. The width of the lines represent their weights.

The training sample encodes the known information between input-output relationships for signal and background. In high energy physics, Monte Carlo simulations generally serve as training samples.

After the training process, the usual procedure is to measure the quality of the learned model by computing the *loss function* using a test sample, which is statistically independent from the training sample [71]. The use of a different sample is necessary to avoid biases and problems derived from tuning on statistical fluctuations. A frequent problem during the implementation of this method is overtraining, in which a model extremely tuned to the training sample is provided, such that the predictive power on an independent sample is completely lost. As the model built during the training must be generalized on an independent sample, another dataset must be introduced to correct this erratic tendency (prevent overtraining), validate the selection and stop the optimization procedure when no further optimization is obtained. This sample is referred to as *validation sample* and is the one on which the measurement is carried out [35]. The overall performance of an MVA algorithm can be visualized in the receiver operating characteristic (ROC) curve, which mostly shows the background rejection as a function of the signal efficiency. When comparing two or more MVA algorithms, the one with the largest area underneath the ROC curve offers on average the best performance. A description of two of the most popular machine learning algorithms in particle physics is presented: neural networks, used in the reconstruction of B mesons (described in section 4.2) and boosted decision trees, employed in the event categorization for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay.

Neural Networks

This algorithm was developed in an attempt to reproduce the decision process that takes place in the neurons of animal brains. When the brain perceives an external stimulus, it generates a response that tells the body whether or not to react to that particular stimulus. In a much smaller scale, only when a certain threshold is reached within the neurons, can the stimuli be transmitted and processed by the

neural network by sending electrochemical signals known as spikes or action potentials. The equivalent of neurons in the machine learning argot is the node. The strength of the connection between nodes is given by the weights, in analogy with the synapse process in the nervous system. The set of input variables correspond to the stimuli, and the activation threshold is encompassed in the transfer function. In general, neural networks perform a nonlinear transformation of the input variables through a transfer function, which is usually a sigmoid function. The sigmoid function has the advantage of mapping any real number in the interval $(-\infty, \infty)$ to the interval $(-1, 1)$ and is given by

$$S(x) = \frac{2}{1 + e^{-\beta x}} - 1, \tag{4.2}$$

where β represents the activation rate of each node in the neural network. The most widely used neural network topology is the multilayer perceptron (MLP), also known as feed forward neural networks. It consists of an interconnected group of nodes arranged in layers; each node processes the received information according to the transfer function, and passes on the result to the next layer of nodes [70]. The input node is followed by one or more hidden layers, and the output node provides the final response of the network; this topology is depicted in Fig. 4.1. The connections between nodes are characterized by distinct weights, which are determined together with the thresholds during the training phase. The target t, i.e. the expected output value of the network, for the neural network is known for each event in the training sample and can be either $t = 1$ for signal or $t = -1$ for background. To measure the quality of the training the entropic loss function is implemented, which is defined as:

$$E_D = w \sum_i \log\left(\frac{1}{2}(1 + t_i o_i + \varepsilon)\right), \tag{4.3}$$

where t_i are the known targets, o_i are the corresponding network outputs and ε is a regularization constant introduced to avoid numerical problems for untrained networks. The last quantity is reduced in each training iteration and reaches zero after a few iterations. The weight parameter, w, is inserted to reduce overfitting and to allow a stable generalization ability.

The Belle full reconstruction module of B mesons uses a Bayesian MLP based on the NeuroBayes package [72], which applies bayesian statistics to incorporate a priori knowledge. The conditional probability to observe B when A has already been observed is given by $P(B|A) = P(B \cap A)/P(A)$. Since $P(B \cap A) = P(A \cap B)$, this leads to the Bayes′ theorem

$$P(B|A) = \frac{P(A|B)P(A)}{P(B)}. \tag{4.4}$$

The relevance of this theorem lies in the interpretation that A and B are defined to be theory and data, respectively. Hence $P(\text{data}|\text{theory})$ is the likelihood, $P(\text{theory})$ is the Bayesian prior and $P(\text{theory}|\text{data})$ represents the a posteriori probability. $P(\text{data})$ is called the evidence. Given that the output of the MLP lies in the interval $(-1, 1)$, to be interpreted as a probability it needs to be rescaled such that the value belongs to the interval $(0, 1)$. This is achieved through the transformation

$$o_t = \frac{o + 1}{2}, \tag{4.5}$$

where o is the output of the neural network. If, on the other hand, the signal to background ratio in the training differs from that of the data as a result of an optimization of the signal component for the training,

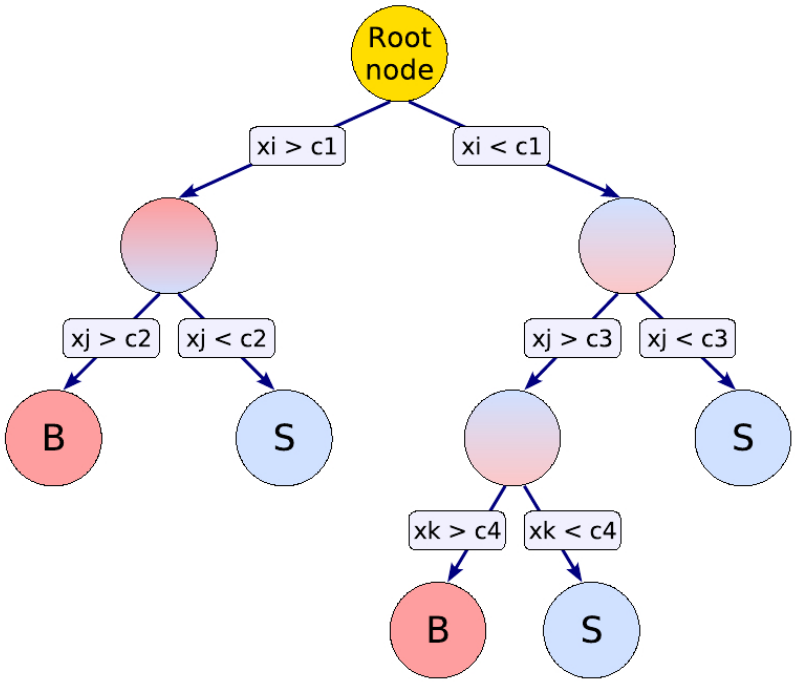


Figure 4.2: Visualization of yes/no decisions (expressed as cuts on variables x_i) that split the data in regions dominated by signal (S) or background (B) events [74].

the output needs to be corrected as:

$$o_p = \frac{1}{1 + \left(\frac{1}{o_t} - 1\right) \frac{P_p(B)}{P_p(S)} \frac{P_t(S)}{P_t(B)}}, \tag{4.6}$$

where $P_t(S)$ and $P_t(B)$ are the a priori probabilities of being signal and background, respectively, in the training dataset, and $P_p(S)$ and $P_p(B)$ are the same in the prediction dataset (real data). This formula is used in the full reconstruction algorithm to calculate the signal probability for modes with low purity so that the signal fraction has to be increased for the network training.

The NeuroBayes package offers a robust algorithm that is less likely to be affected by overtraining than the average MLP and enhances the generalization abilities of the network [73]. This feature is in part due to the preprocessing of the input variables carried out in three distinctive stages. The first stage corresponds to the flattening of the distributions of the input variables within the interval $(0, 1)$, which has the goal of reducing the effect of outliers in the measurement. This is accomplished by drawing a histogram for each variable with a varying bin width, such that each bin contains the same amount of events. Subsequently, the underlying PDF for signal and background are inferred and used to calculate the purity for each input variable. The purity distribution is transformed again to a Gaussian with a mean zero and unit width. The final stage corresponds to the decorrelation of the input variables by calculating their correlation matrix and then diagonalizing it. The diagonalization is executed through iterative Jacobian rotations. The preprocessed variables can afterwards enter the MLP and follow the training process.

Boosted Decision Trees

A *decision tree* (DT) is a sequential application of cuts over the set of input variables devised to separate the data into two regions dominated either by signal or background events. It starts by searching for the most discriminating variable and a cut value that provides the best FOM, i.e. the best separation between signal and background following a given function such as the precision given by $FOM = S/\sqrt{S+B}$, thus segmenting the data into two regions known as nodes. Each node receives a similar treatment, starting an iterative process that finishes once a given criterion is satisfied. This criterion can be a minimum number of events present in the node or no further improvements in the FOM. Since the graphical representation of the series of conditions used to stratify the data resembles an upside-down tree, this algorithm is called a decision tree [75]. Figure 4.2 gives a generic view of a DT. Following this analogy, the first node that starts the selection is called the root node, nodes with a split receives the name of branches, and the terminal nodes where no division occurs are known as leaves. The depth of the tree is given by the number of layers of the trees, that is, the number of consecutively applied cuts. A frequently used FOM is the *Gini index*, defined as:

$$G(p) = p(1-p), \tag{4.7}$$

where p is the purity of the node [74]. Nodes dominated by signal events have a higher purity than those populated by background events.

The major advantage of DTs is their simplicity, as they allow people to follow the progress along the tree and see how decisions are being made as well as see the relative importance of the input variables. As they can be displayed graphically, they are easily interpreted even by a non-expert. Nonetheless, DTs lack the level of predictive accuracy of other machine learning algorithms like neural networks and are sensitive to statistical fluctuations in the training sample. This instability is reflected in the fact that a small change in the configuration results in a substantially different classifier response; thus a DT is prone to overtraining. To overcome these drawbacks, different ensemble techniques have been developed resulting in methods with a significant improvement in their predictive power.

The basic element of the ensemble techniques lies in the concept of weak learners. An individual DT is a weak learner in the sense that it has a larger probability than a random choice of making the correct classification but with a poor accuracy. In other words, a weak learner is a classifier that does not always label efficiently an event as signal. An example would be the determination of the sex of a person based on their height. Saying that a person is male if they are above 1.70 m will incur in an error, but the chance that it is true is bigger than 50%. The ensemble techniques, such as boosting and bagging, take a set of weak learners, construct a forest of DTs and obtain a strong learner or discriminator based on the majority vote of each individual DT in the forest. Hereafter boosted decision trees (BDT) denote the set of modified DTs according to the ensemble algorithm [74]. As stated in the introduction to multivariate methods, the goal of the ensemble algorithm is to minimize a loss function such that it provides the best predictive performance and incorporates the correct compromise between bias and variance. The bias refers to the difference between the learned function and the true function, whereas variance is a measure of the sensitivity of the learned function to inputs [70]. Some of the most widely used ensemble algorithms are outlined below.

- *Adaptive Boost (AdaBoost)*, like in all boosting algorithms, the final discriminator of the forest of DTs is a weighted sum of their individual outputs. The particularity of AdaBoost resides in the way the weights are assigned: here a misclassified event will have a bigger weight than correctly identified events [76]. The initial DT is trained using the same weights, the following DTs are trained using an altered sample where the weights of previously misclassified events are multiplied by a boost weight α [74]. The boost weight α is related to the misclassification rate err of the

previous tree by:

$$\alpha = \frac{1 - err}{err}. \tag{4.8}$$

The weights of the complete event sample are then renormalized to keep the sum of weights constant. If $h(\vec{x})$ encodes the individual classifier such that $h(\vec{x}) = +1$ for signal and $h(\vec{x}) = -1$ for background, the boosted event classification $y_{\text{boost}}(\vec{x})$ is thus given by:

$$y_{\text{boost}}(\vec{x}) = \frac{1}{N_{\text{collection}}} \sum_{i}^{\text{collection}} \ln(\alpha_i) h_i(\vec{x}), \tag{4.9}$$

where the sum is over all classifiers in the collection. Large (small) values for $y_{\text{boost}}(\vec{x})$ indicate a signal-like (background-like) event. AdaBoost can then be understood as a procedure to minimize an exponential loss function.

- *Gradient Boost*. Since the exponential loss lacks robustness in the presence of outliers or mislabeled data points, the effectiveness of the AdaBoost algorithm is expected to degrade for noisy settings. An attempt to remove this flaw is by modifying the loss function. The GradientBoost algorithm [77] implements the binomial log-likelihood loss for classification, which is given by

$$L(F, y) = \ln(1 + e^{-2F(\vec{X})y}), \tag{4.10}$$

where $F(\vec{X})$ represents the model response and y is the true value. Because a straightforward boosting algorithm for this loss function cannot be derived, one has to resort to a steepest-descent approach to do the minimization. To do so, one has to calculate the gradient of the loss function and then grow a regression tree whose leaf values are adjusted to match the mean value of the gradient in each region defined by the tree structure. The repetition of this procedure yields the desired collection of DTs which minimizes the loss function.

- *Bagging* is an acronym for Bootstrap AGGregatING. This algorithm involves the construction of many DTs on independently drawn bootstrapped replicas (randomly selected subset) of the training sample, such that the final discriminator is a simple average of the individual classifiers [78]. Bagging is not a boosting algorithm per se, instead it smears over statistical representations of the training data and is hence suited to stabilize the response of a classifier.

A BDT can implement a conjunction of ensemble algorithms to improve the performance of the classification process. This thesis adopts a stochastic gradient boosting algorithm [79] to carry out the event classification for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay. This algorithm introduces a bagging-like resampling procedure to generate a random subsample on which a gradient boost is then applied.

4.2 Reconstruction of B mesons

Due to their short lifetime, B mesons decay inside the beam pipe of the detector, close to the interaction point, and therefore cannot be directly measured. Instead, the information that can be extracted from the detector comes from the final state particles (FSP), namely particles with a long enough lifetime or stable particles that deposit signals in subdetectors such that their four-momenta can be measured. These particles are: $e^\pm$, $\mu^\pm$, $\pi^\pm$, $K^\pm$, p, $\bar{p}$, γ and K_L^0. The momenta of these FSP are summed up to form intermediate particles and different consistency checks are then performed to associate them with a

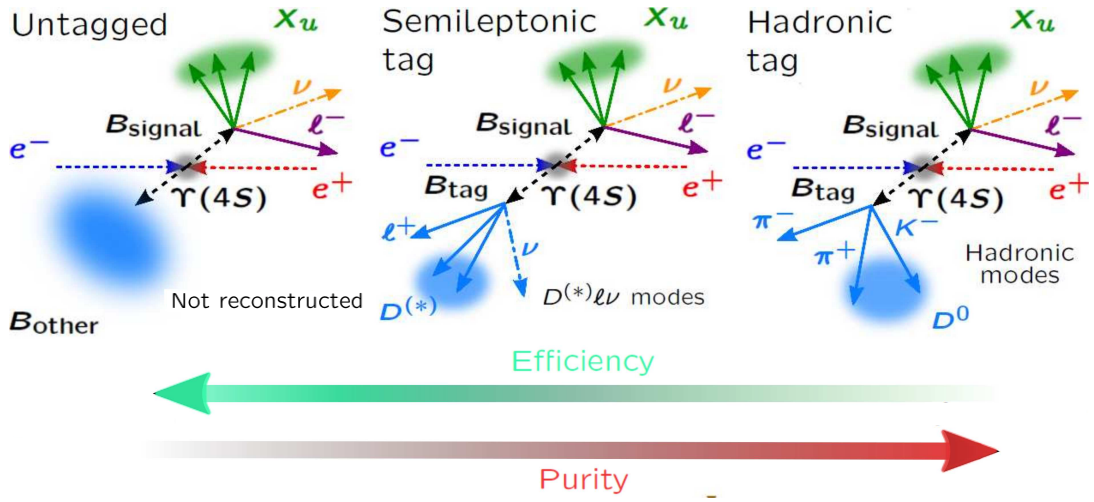


Figure 4.3: Reconstruction techniques for the signal B meson, comparing their efficiency and purity. In the untagged method (left) only the B_{sig} is reconstructed resulting in a large sample of signal and background events. The tagged methods, semileptonic (center) and hadronic (right), reconstruct the other B-meson (B_{tag}) either in charm semileptonic or hadronic B decays, improving the background rejection but also incurring in a smaller signal sample than in the untagged method. Analyses utilizing the hadronic tag technique require a large data sample to provide precise results as the efficiencies achieved with this method are smaller compared with the other two methods.

specific B meson decay. To pinpoint candidates one can make use of discriminating variables sensitive to B meson properties. However, not all B meson decays can be fully reconstructed. Those decays involve one or more neutrinos that are invisible to the detector so their momenta cannot be measured. To compensate this drawback, one can exploit the almost perfect hermeticity of the Belle detector to impose kinematic constraints on the decay to infer the neutrino and consequently determine in an indirect manner the B meson momentum. These kinematic constraints also benefit from the fact that the $B\bar{B}$ pairs are produced without any additional particles and the initial state energy is precisely known. For instance, one can partially reconstruct semileptonic B decays by assigning the four-momentum of the neutrino to the total missing four-momentum of the event,

$$P_\nu = P_{\text{miss}} = P_{\text{beam}} - \sum_i P_i, \tag{4.11}$$

that is the difference between the beam four-momentum (P_{beam}) and the sum of all four-momenta of the reconstructed particles (P_i).

Traditionally, the $B-$factories have employed three approaches to reconstruct the B meson, namely untagged, semileptonic tag and hadronic tag methods. In the untagged method the reconstruction only focuses on the B decay of interest, which is referred to as signal B or B_{sig}. The untagged method offers a large sample of signal candidates but also incurs in a large background, which translates in a high signal efficiency but a poor purity. Typical values for signal reconstruction efficiency are of the order of 5%. The other two methods involve the reconstruction of the accompanying B meson in exclusive channels. This meson is called the tag B (B_{tag}). In the semileptonic tag technique, the B_{tag} is reconstructed in semileptonic decays to charm mesons such as $D^{\pm}$, D^0, $D^{*\pm}$, D^{*0} among others, which are subsequently reconstructed in hadronic decay modes. Despite the large branching ratio of semileptonic decays of B_{tag} to charm mesons (around 11%), the presence of an additional neutrino restricts the kinematic constraints that can be established, thus the properties of the B_{sig} cannot be determined from the B_{tag}.

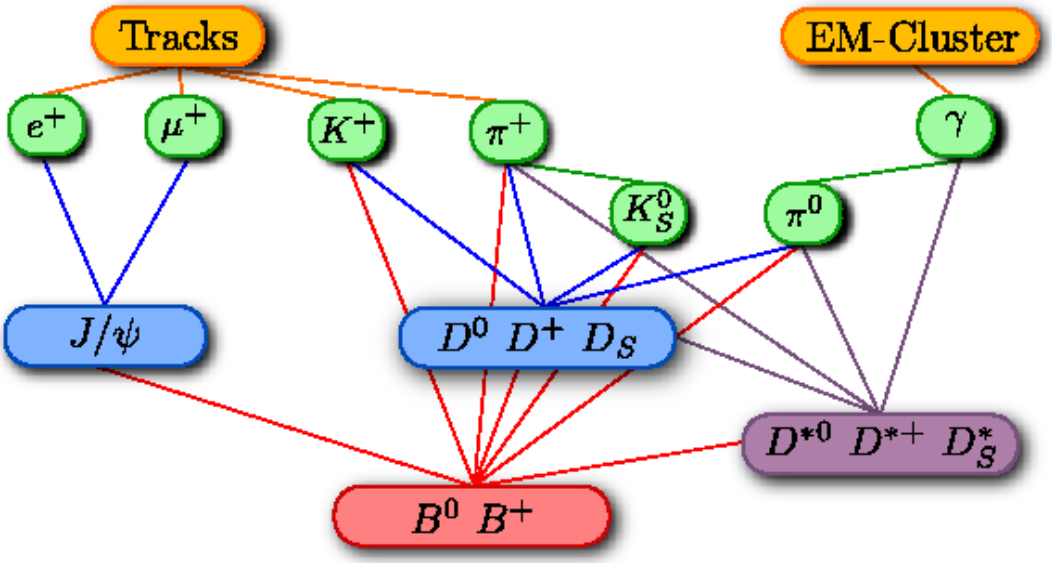


Figure 4.4: The four stages of the full reconstruction represented by different colors. The connections symbolize the particles used in the reconstruction of intermediate particles [73].

The efficiencies obtained with this method are around 0.6%. Finally, the hadronic tag method, also known as full reconstruction, reconstructs the B_{tag} in hadronic modes containing charm mesons of the form $B \to \bar{D}^{(*)} n\pi$, $B \to \bar{D}^{(*)} D_S^{(*)}$ or $B \to J\psi K m\pi$, where n and m indicate any number ($n, m < 10$) of charged or neutral pions, respectively. This method allows the determination of the direction, momentum, flavor and charge of the B_{sig}. Although this method provides a considerable signal purity, the obtained signal efficiencies are relatively poor, below 0.1%. A descriptive representation of the aforementioned methods is depicted in Fig. 4.3. The analysis presented in this thesis makes use of the hadronic tag method, which is described in further detail below.

4.2.1 NeuroBayes Full Reconstruction of B mesons

At Belle, the B_{tag} hadronic reconstruction[1] proceeds in a hierarchical order starting from the FSP to the B meson candidates in four stages, as illustrated in Fig. 4.4. At each stage a NeuroBayes neural network combines all the available information on a given candidate to generate a single scalar variable referred to as network output. This variable can be interpreted as the probability for a given particle to be correctly reconstructed. The network output for each reconstructed particle becomes an input for the neural network in the next stage. In the first stage $\pi^\pm$, $K^\pm$, K_S^0, γ and π^0 candidates are reconstructed and classified. The inputs for the neural network at this point include measurements of time-of-flight, the energy loss in the CDC, Cherenkov light in the ACC for the charged particles, and shower shape variables for photons. The second stage goes on with the reconstruction of D^0, $D_{(S)}^\pm$ and J/ψ candidates. Some of the most discriminating variables used at this stage in the training of the neural network are the network outputs of the daughters, the invariant mass of the candidate, the angle between the momentum of the $D_{(S)}$ meson and the vector joining the $D_{(S)}$ decay vertex and the interaction point, and the significance of the distance between the decay vertex and the interaction point. Excited charmed mesons are reconstructed in the third stage using similar input variables as the one used in the second stage, except for the invariant mass

[1] The internal name of this module is `ekpfullrecon`.

Table 4.1: Reconstructed decay modes for stage two of the reconstruction of B_{tag} and their branching fractions $\mathcal{B}$ [73].

Mode	$\mathcal{B}$ (%)	Mode	$\mathcal{B}$ (%)	Mode	$\mathcal{B}$ (%)
$D^0 \to K^-\pi^+$	3.89	$D_S^+ \to K^+K_S^0$	1.49	$D^+ \to K^-\pi^+\pi^+$	9.40
$D^0 \to K^-\pi^+\pi^+\pi^-$	8.09	$D_S^+ \to K^+\pi^+\pi^-$	0.69	$D^+ \to K_S^0\pi^+$	1,49
$D^0 \to K^-\pi^+\pi^0$	13.9	$D_S^+ \to K^+K^-\pi^+$	5.50	$D^+ \to K_S^0\pi^+\pi^0$	6.90
$D^0 \to \pi^-\pi^+$	0.14	$D_S^+ \to K^+K^-\pi^+\pi^0$	5.60	$D^+ \to K^-\pi^+\pi^+\pi^0$	6.08
$D^0 \to \pi^+\pi^-\pi^0$	1.44	$D_S^+ \to K^+K_S^0\pi^+\pi^-$	0.96	$D^+ \to K_S^0\pi^+\pi^+\pi^-$	3.10
$D^0 \to K_S^0\pi^0$	1.22	$D_S^+ \to K^-K_S^0\pi^+\pi^+$	1.64	$D^+ \to K^+K^-\pi^+$	0.98
$D^0 \to K_S^+\pi^+\pi^-$	2.94	$D_S^+ \to K^+K^-\pi^+\pi^+\pi^-$	0.88	$D^+ \to K^+K^-\pi^+\pi^0$	1.50
$D^0 \to K_S^+\pi^+\pi^-\pi^0$	5.40	$D_S^+ \to \pi^+\pi^+\pi^-$	1.10		
$D^0 \to K^-K^+$	0.39	$J/\Psi \to e^+e^-$	5.94		
$D^0 \to K^-K^-K_S^0$	0.47	$J/\Psi \to \mu^+\mu^-$	5.93		

of the candidate. The mass difference $\Delta m = m_{D_{(S)^*}} - m_{D_{(s)}}$ is employed instead. Finally, the B mesons are reconstructed in the fourth stage with variables including the network outputs of the daughters, the mass of the D mesons or Δm, the angle between the B meson momentum and the beam, and the energy difference[2] ΔE. The decay modes of the reconstructed intermediate particles and their branching ratio are arranged in tables 4.1 and 4.2. The combination of all the decay modes results in 1104 exclusive B decay channels. The final output available for the rest of the analysis is the network output of the B mesons o_{tag}, which takes values between 0 and 1 with larger values corresponding to a higher probability that a B meson was correctly reconstructed. In the best case scenario the efficiency of the NeuroBayes full reconstruction is about 0.18% for B^0 mesons and 0.28% for $B^\pm$ mesons. This value lowers as further cuts on o_{tag} and variables sensitive to B meson properties such as ΔE are applied.

Continuum suppression

Nonresonant events not stemming from $e^+e^- \to \Upsilon(4S)$, but rather from the production of light quarks of the form $e^+e^- \to q\bar{q}$ with $q = u, d, s, c$ quarks are called continuum events. In comparison with B mesons that decay almost at rest in the lab frame and whose decay products are isotropically distributed in the detector, continuum events are characterized by a preferred direction showing a jet-like structure. At the production energy of the $\Upsilon(4S)$, the cross section of continuum events is about three times greater than that for $B\bar{B}$ meson pairs[3]. Different variables that take advantage of the event topology, such as Fox-Wolfram-Moments [80] (FWM), are then used to separate B meson decay from continuum events. The NeuroBayes full reconstruction[4] also allows to include information of event shape variables in the training of the neural network to obtain an additional variable o_{tag}^{cs} that discriminates between these two event types. The module also includes as input variables the super Fox-Wolfram-Moments [81] (SFWM) in the computation of the network output. The SFWM compute the FWM not for the whole event, but

[2] This variable corresponds to the difference between the energy of the B meson and the energy of the beam, both evaluated in the CMS frame

[3] For continuum events the cross section is 3.7 nb and for $B\bar{B}$ is 1.1 nb.

[4] This module is called `ekpcontsuppress`.

Table 4.2: Reconstructed modes for D^* and B mesons with their branching fractions $\mathcal{B}$ [73].

Mode	$\mathcal{B}$ (%)	Mode	$\mathcal{B}$ (%)	Mode	$\mathcal{B}$ (%)
$D^{*+} \to D^0\pi^+$	67.7	$B^+ \to D_S^{+*}\bar{D}^0$	0.760	$B^0 \to D_S^+ D^-$	0.720
$D^{*+} \to D^+\pi^0$	30.7	$B^+ \to D_S^+\bar{D}^{0*}$	0.820	$B^0 \to D^{*-}\pi^+$	0.276
$D^{*0} \to D^0\pi^0$	61.9	$B^+ \to D_S^{+*}\bar{D}^{0*}$	1.710	$B^0 \to D^{*-}\pi^+\pi^0$	1.500
$D^{*0} \to D^0\gamma$	38.1	$B^+ \to \bar{D}^0 K^+$	0.037	$B^0 \to D^{*-}\pi^+\pi^+\pi^-$	0.700
$D_S^{*+} \to D_S^+\gamma$	94.2	$B^+ \to D^-\pi^+\pi^-$	0.107	$B^0 \to D^{*-}\pi^+\pi^+\pi^-\pi^0$	1.760
$B^+ \to \bar{D}^0\pi^+$	0.484	$B^+ \to J/\Psi K^+$	0.101	$B^0 \to D_S^{*+} D^-$	0.740
$B^+ \to \bar{D}^0\pi^+\pi^0$	1.340	$B^+ \to J/\Psi K^+\pi^+\pi^-$	0.107	$B^0 \to D_S^* D^{*-}$	0.800
$B^+ \to \bar{D}^0\pi^+\pi^+\pi^-$	1.100	$B^+ \to J/\Psi K^+\pi^0$	0.047	$B^0 \to D_S^{*+} D^{*-}$	1.770
$B^+ \to D_S^+\bar{D}^0$	1.000	$B^+ \to J/\Psi K_S^0\pi^+$	0.094	$B^0 \to J/\Psi K_S^0$	0.087
$B^+ \to \bar{D}^{0*}\pi^+$	0.519	$B^0 \to D^-\pi^+$	0.268	$B^0 \to J/\Psi K^+\pi^-$	0.120
$B^+ \to \bar{D}^{0*}\pi^+\pi^0$	0.980	$B^0 \to D^-\pi^+\pi^0$	0.760	$B^0 \to J/\Psi K_S^0\pi^+\pi^-$	0.100
$B^+ \to \bar{D}^{0*}\pi^+\pi^+\pi^-$	1.030	$B^0 \to D^-\pi^+\pi^+\pi^-$	0.800		
$B^+ \to \bar{D}^{0*}\pi^+\pi^+\pi^-\pi^0$	1.800	$B^0 \to \bar{D}^0\pi^0$	0.026		

separately for the B_{tag} candidate and the rest of the event.

CHAPTER 5

DATASETS AND EVENT SELECTION

Experimentally, semileptonic decays are reconstructed with simple objects in the final state: a lepton that could be either an electron or a muon, and a hadron that could result from the combination of one or more particles. Afterwards, these two objects are merged to partially reconstruct the B meson, and the neutrino information is then inferred from the missing momentum of the whole event. This analysis implements full reconstruction to build up the B_{tag} meson, allowing to know further information about the B_{sig}, e.g. direction, charge, flavor and momentum, hence introducing kinematic constraints to reduce substantial contributions from the background. To extract physical parameters of interest from this decay, one needs to rely on a model for the signal and possible sources of background. This is achieved with the use of Monte Carlo (MC) simulations. Based on the MC, different requirements can be found to enhance the signal purity in the event and later on to extract the parameters of interest from experimental data. The data remains untouched in the first stages of the analysis, which ensures the selection, optimization and validation of the analysis procedure is entirely carried out using the MC samples. This course of action is referred to as *blind analysis*, which is intended to circumvent bias in the measurement.

Throughout this thesis the inclusion of the charge conjugate mode is implied, unless otherwise stated, meaning that the notation used for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay also covers the $B^- \to \pi^+\pi^-\ell^-\bar{\nu}_\ell$ decay. Henceforward the analysis of the datasets and the generation of figures were carried out using the ROOT software framework [82]. This chapter is organized as follows: section 5.1 describes the MC samples and the corrections implemented for the decay channel used for this analysis, and Section 5.2 presents the strategy for reconstructing the signal side and suppressing the background with the utilization of a BDT.

5.1 Data and Monte Carlo Samples

This analysis uses the entire Belle $\Upsilon(4S)$ data sample of 711 fb^{-1}, containing an estimated $(772 \pm 11) \times 10^6$ $B\bar{B}$ pairs. A 79.4 fb^{-1} sample of off-resonance data were collected 60 MeV below the resonance peak energy of the $\Upsilon(4S)$. After the application of the continuum suppression module from the full reconstruction technique, this sample is left with a poor statistics and is not directly used in the analysis, and its usage is limited to the assessment of model uncertainties in the simulated processes of continuum events. The MC samples are produced in sets with a total integrated luminosity equivalent to that of the on-resonance data. These sets are referred to as streams. The MC samples are simulated using the EvtGen [83] package and the response of the detector is modeled using GEANT3 [84]. The simulated

physics processes are listed below:

1. **Generic B decays:** groups all the processes that happen in the reaction $e^+e^- \to \Upsilon(4S) \to B\bar{B}$, where the B meson decays entirely via the dominating quark-level transition $b \to cW$. It is composed of two samples: decays of neutral B mesons (B^0) called `mixed`, and decays of charged mesons ($B^\pm$) known as `charged`. Each sample is generated with ten streams of MC. A list of the branching ratios used in the MC modeling for semileptonic B decays to charm mesons compared to the current world average[1] is presented in table A.1 of appendix A.1.

2. **Generic continuum MC:** simulates the nonresonant processes in the reaction $e^+e^- \to q\bar{q}$ with $q = u, d, s, c$ quarks. It consist of two samples: `charm` which contains $e^+e^- \to c\bar{c}$ processes, and `uds` that simulates the events with the light quarks u, d and s. Both samples are generated with six streams of MC.

3. **Rare *B* decays:** this sample includes processes involving radiative decays, decays that proceed via loop transitions such as $b \to s$ quark transitions or electroweak penguin diagrams, leptonic decays, hadronic decays to charmless meson, and decays to baryons among others. Most of these decays have a branching ratio lower than 1×10^{-5}. This sample is also split in two samples depending on the charge of the B mesons: `rare_charged` and `rare_mixed`. Each sample is generated with 50 streams of MC.

4. **Signal MC:** at the moment of MC production for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay there was no form factor calculation available for this channel in the entire phase space. The semileptonic generators available such as ISGW2 [39] only admit three input particles (one meson or intermediate resonance, one charged lepton and one neutrino) and not four like it is required in the current case. Consequently, the PHase SPace (PHSP) generator was selected to simulate the signal decay, as it treats the channel as a four-body decay with events having equal probability in all points of the phase space. In addition to the PHSP generator, the PHOTOS [85, 86] package was also used to account for radiative effects. This sample is assumed to contain the physics of the resonant[2] and nonresonant contributions to the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ channel. The previous samples are available to all Belle collaborators, from them the "generic" samples are common to all analyses modules. The MC signal for $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay was only produced for this thesis. Calculations of the branching ratio for this decay are available from LQCD and LCSR in limited regions of phase space, and hence are not used as input for generating the signal. Estimations of branching fractions in the whole phase space for semileptonic decays of B mesons with multi-pion final states are given in reference [87] in terms of the ratio $|V_{ub}/V_{cb}|^2$. A measurement of this ratio is provided by the LHCb collaboration using the exclusive b-flavored baryon decays $\Lambda_b^0 \to p\mu^-\bar{\nu}_\mu$ and $\Lambda_b^0 \to \Lambda_c^+\mu^-\bar{\nu}_\mu$ resulting in $|V_{ub}/V_{cb}| = 0.083 \pm 0.006$ [88]. With this value, the branching ratios for B semileptonic decays to multi-pion final states are presented in table 5.1.

   Hence, the assumed branching ratio for the decay $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ is 31.7×10^{-5}. Another calculation [89] gives the partial branching ratio for this decays in two regions of the di-pion mass; below 1 GeV the branching ratio is 15.6×10^{-5} and above that value is 28.6×10^{-5}. However, the values for this calculation are discarded as they assume that masses above 1 GeV are dominated

[1] The world average refers to the average of all the measurements reported in the particle data group (PDG) [1].

[2] This is mainly motivated by the fact that the missing mass squared, which is defined as the squared of the missing momentum, does not distinguish between resonant and nonresonant states, since it coincides with the mass squared of the neutrino for true events in $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decays.

X_u	BR [10^{-5}]	X_u	BR [10^{-5}]
π^0	7.6	$4\pi^0$	1.4
$\pi^+\pi^-$	31.7	$\pi^+\pi^-2\pi^0$	35.9
$\pi^0\pi^0$	2.8	$2\pi^+2\pi^-$	20.7
$\pi^+\pi^-\pi^0$	70.4	$2\pi^+2\pi^-\pi^0$	7.6
$3\pi^0$	3.5	$\pi^+\pi^-3\pi^0$	6.2

Table 5.1: Branching ratio (BR) of semileptonic $B^\pm$ decays to multi-pion final states according to calculations from reference [87] using $|V_{ub}/V_{cb}| = 0.083 \pm 0.006$ [88]. The hadronic component is denoted by X_u.

by nonresonant contributions, which is not supported by a recent Belle measurement [47] where there was not a significant nonresonant contribution for the high mass region[3]. Approximately 188 streams of MC are produced to study the signal channel.

5. $\boldsymbol{B \to X_u \ell \nu}$**:** contains events occurring in the reaction $e^+e^- \to \Upsilon(4S) \to B\bar{B}$, such that one B meson decays through the cabibbo suppressed transition $b \to u\ell\nu$ at quark level. This sample groups a series of exclusive semileptonic decays of B mesons to one light charmless meson (X_u). As in the Generic B decays, these processes are arranged in two samples, dependent on the charge of the B meson, `bulnu_charged` and `bulnu_mixed`. A total of 20 streams are generated for each MC sample. A list of the branching ratios of these exclusive channels, compared with the current world average, is given in table 5.2. This sample does not include an inclusive[4] component since the V_{ub} generator, used to model this contribution, incorrectly describes nonresonant contributions in the entire phase space. High multiplicity mass terms that can contribute to the nonresonant component come from decays such as $B^+ \to \pi^+\pi^-\pi^0\ell^+\nu_\ell$ and $B^+ \to \pi^+\pi^-\pi^0\pi^0\ell^+\nu_\ell$. However, after simulating these processes with the PHSP generator and examining their contributions in the signal region defined in section 5.2.4, it was found to be negligible and not taken into account for the rest of the analysis.

5.1.1 Corrections to the MC samples

Precision measurements of physical parameters demand an excellent understanding of the background, however the use of MC simulations cannot perfectly describe the background due to several assumptions in its generation. These assumptions may involve a lack of precise knowledge in some of the parameters used to simulate the decay, differences in efficiency in the data and MC or the theory on which the MC relies does not correctly model the processes, among other reasons. To fix this problem, different calibrations are performed on well known decays, known as control samples, where an event-wise weight is inferred from the comparison of the data and the MC samples. This weight can then be used in the particular channel under study. In case an analysis needs corrections from different sources, the total event is given by the product of all weights that apply to it. The required corrections vary from one analysis module to another depending on the FSP in the channel, whether or not the B_{tag} is reconstructed, among

[3] This measurement focuses in the determination of the branching ratio of the $B \to \rho\ell\nu$ decays and the extraction of $|V_{ub}|$ from this decay; there is no detailed study of the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ channel.

[4] The inclusive $B \to X_u\ell\nu$ component contains all possible decays involving a $b \to u$ quark transition, including resonant and nonresonant contributions. As many exclusive (resonant) channels are included in the sample, the term inclusive is used here to be regarded as nonresonant component.

B^+ Mode	$\mathcal{B}(10^{-3})$(MC)	$\mathcal{B}(\times 10^{-3})$(PDG)	Form factor model
$\pi^0\ell\nu$	0.073	0.077±0.012	LCSR
$\rho^0\ell\nu$	0.149	0.107±0.013	LCSR
$\eta\ell\nu$	0.084	0.038±0.009	ISGW2
$\eta'\ell\nu$	0.033	0.023±0.008	ISGW2
$\omega\ell\nu$	0.115	0.115±0.016	LCSR
$a_0(980)^0\ell\nu$	0.003		ISGW2
$a_1(1260)^0\ell\nu$	0.005		ISGW2
$a_2(1320)^0\ell\nu$	0.024		ISGW2
$b_1(1235)^0\ell\nu$	0.082		ISGW2
$f_1(1285)\ell\nu$	0.062		ISGW2
$f_2'(1525)\ell\nu$	0.087		ISGW2
B^0 Mode	$\mathcal{B}(\times 10^{-3})$(MC)	$\mathcal{B}(\times 10^{-3})$(PDG)	Form factor model
$\pi^+\ell\nu$	0.136	0.136±0.007	LCSR
$\rho^+\ell\nu$	0.277	0.277±0.024	LCSR
$a_0(980)^+\ell\nu$	0.007		ISGW2
$a_1(1260)^+\ell\nu$	0.123		ISGW2
$a_2(1320)^+\ell\nu$	0.047		ISGW2
$b_1(1235)^+\ell\nu$	0.154		ISGW2

Table 5.2: Exclusive channels used in the Monte Carlo simulation to model the $B \to X_u\ell\nu$ component.

other factors. An outline of the corrections applicable to the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ channel are presented in the following.

Charged pion identification efficiency correction

Charged pion identification is based on the probability of classifying a track as a pion as described in section 3.3.3. To examine the efficiency of this classification, the $D^{*+} \to D^0(K^-\pi^+)\pi^+_{\text{slow}}$ decay is used, which contains a clean sample of charged kaons and pions. The decay has a low available phase space which results in the π^+_{slow} having a small momentum. This decay exhibits a clear distinction between correctly reconstructed events and the background, which is visible in the distribution of the mass difference ($\Delta m = m_{D^{*+}} - m_{D^0}$). From the charge correlation between the hadrons in the D^0 and the $\pi^\pm_{\text{slow}}$, one can easily identify the particle without requiring a cut on the `Prob(π)` variable. To derive the correction factors, different tables [90] are available for each cut on `Prob(π)` in bins of the pion momentum in the lab frame and the pion polar angle. The corrections for each bin are given by the efficiency ratio $R_i = \epsilon_i^{\text{data}}/\epsilon_i^{\text{MC}}$, where ϵ_i^{MC} and ϵ_i^{data} are the MC and data efficiencies, respectively. The total efficiency correction is then:

$$R_\pi = \frac{1}{N}\sum_i n_i R_i, \tag{5.1}$$

where n_i is the number of events in the bin such that $N = \sum_i n_i$.

The uncertainty associated with the efficiency correction, ΔR, receives contributions from the statistical uncertainties (ΔR_i^{stat}) of the MC and data samples and systematic uncertainties (ΔR_i^{syst}) due to the

determination of the correctly reconstructed D^{*+} yields. The statistical uncertainties in different bins are assumed to be independent, while systematic uncertainties are fully correlated. The total error becomes:

$$\Delta R_\pi = \frac{1}{N}\left(\sqrt{\sum_i (n_i \Delta R_i^{\text{stat}})^2 + \left(\sum_i n_i \Delta R_i^{\text{syst}}\right)^2}\right) + R_{\text{const}}. \tag{5.2}$$

Here, R_{const} is an additional constant error for the possible dependence on the data-acquisition period estimated to be 0.003 for pion identification efficiencies.

Lepton identification efficiency correction

The charged lepton[5] identification efficiency correction proceeds in a similar fashion to the previous case in that it depends on the cut on the likelihood value for accepting tracks as electrons or muons (see sections 3.3.1 and 3.3.2). The efficiency ratios are given in bins of the polar angle and momentum of the lepton evaluated in the lab frame, as described in references [91, 92]. The calibration studies in this case are carried out in a sample of $\gamma\gamma \to \ell^+\ell^-$ decays, which is a clean channel with low multiplicity tracks, and has associated a systematic uncertainty $\Delta R_i^{\text{syst1}}$. To assess the impact of a hadronic environment in the efficiency calculation, a sample of inclusive $B \to J/\psi(\ell^+\ell^-)X$ was investigated, where X refers to any particle. From the comparison of the samples $\gamma\gamma \to \ell^+\ell^-$ and $J/\psi \to \ell^+\ell^-$, an additional systematic uncertainty is calculated ($\Delta R_i^{\text{syst2}}$). For the calculation of the total error, the statistical errors are treated as uncorrelated whereas the systematic uncertainties as correlated, thence the total efficiency correction can be written as:

$$R_\ell = \frac{1}{N}\sum_i n_i \frac{\epsilon_i^{\ell,\text{data}}}{\epsilon_i^{\ell,\text{MC}}}, \tag{5.3}$$

and its uncertainty as:

$$\Delta R_\ell = \frac{1}{N}\sqrt{\sum_i (n_i \Delta R_i^{\text{stat}})^2 + \left(\sum_i n_i \Delta R_i^{\text{syst1}}\right)^2 + \left(\sum_i n_i \Delta R_i^{\text{syst2}}\right)^2}. \tag{5.4}$$

In the above expressions n_i is the number of events in the bin, N is the total number of events and $\epsilon_i^{\ell,\text{data}}$ and $\epsilon_i^{\ell,\text{MC}}$ are the lepton efficiencies in data and MC, respectively.

Hadronic tag efficiency correction

Due to the large amount of exclusive hadronic channels in which the B_{tag} is reconstructed (1104), incorrect assumptions in their branching ratios can propagate and lead to an unbalance between the expected number of events in data and MC. To determine the extent of this effect, one can use a double-tagging approach in which one can study the B_{tag} by reconstructing the B_{sig} in decay channels with a sufficiently large branching ratio (so they can serve as control samples) and assuming that it is well modeled by MC. The last condition implies that the correction factors derived from this study depends only on the tag side, so it can be applied to other searches with a different signal decay channel. This calibration was carried out in reference [93] using B semileptonic decays to charm mesons as the signal with the number

[5] In the rest of the thesis, the term lepton is used exclusively for charged leptons and not for neutrinos.

D^*	Generic	HFLAV
ρ^2	1.30	$1.207 \pm 0.015 \pm 0.021$
R_1	1.18	1.403 ± 0.033
R_2	0.71	0.854 ± 0.020
D	Generic	HFLAV
ρ_D^2	1.16	$1.186 \pm 0.036 \pm 0.041$

Table 5.3: Comparison of HQET2 parameters for the $B \to D^{(*)}\ell\nu$ decays in the generic MC and HFLAV.

of double-tagged events given by:

$$N(B_{\text{tag}}B_{\text{sl}}) = N_{B\bar{B}} \times \mathcal{B}(B_{\text{tag}} \to f)\epsilon_{B_{\text{tag}}\to f} \times \mathcal{B}(B_{\text{sl}} \to D^{(*)}\ell\nu)\epsilon_{B_{\text{sl}}} \tag{5.5}$$

where $\mathcal{B}(B_{\text{tag}} \to f)\epsilon_{B_{\text{tag}}\to f}$ is the product of branching ratio and the reconstruction efficiency of the specific decay $B_{\text{tag}} \to f$ and $\mathcal{B}(B_{\text{sl}} \to D^{(*)}\ell\nu)\epsilon_{B_{\text{sl}}}$ is the corresponding product for the semileptonically decaying B meson [35]. The correction factor is then calculated as the ratio of the numbers of reconstructed tagged events in data and MC:

$$\begin{aligned} c_{\text{tag}}^{f} &= \frac{\mathcal{B}^{\text{data}}(B_{\text{tag}} \to f)\epsilon^{\text{data}}_{B_{\text{tag}}\to f}}{\mathcal{B}^{\text{MC}}(B_{\text{tag}} \to f)\epsilon^{\text{MC}}_{B_{\text{tag}}\to f}} \\ &= \frac{N^{\text{data}}(B_{\text{tag}}B_{\text{sl}})}{N^{\text{MC}}(B_{\text{tag}}B_{\text{sl}})} \frac{N^{\text{MC}}_{B\bar{B}} \times \mathcal{B}^{\text{MC}}(B_{\text{sl}} \to D^{(*)}\ell\nu)}{N^{\text{data}}_{B\bar{B}} \times \mathcal{B}^{\text{data}}(B_{\text{sl}} \to D^{(*)}\ell\nu)}. \end{aligned} \tag{5.6}$$

In the last step, $\epsilon^{\text{data}}_{B_{\text{sl}}} = \epsilon^{\text{MC}}_{B_{\text{sl}}}$ as the $B_{\text{sl}} \to D^{(*)}\ell\nu$ decay modes were assumed to be well modeled in the MC samples to generalize the correction to all analysis modules using the NeuroBayes hadronic reconstruction tagging method. These corrections are calculated for each reconstructed B_{tag} mode as a function of the NeuroBayes neural network output o_{tag}. The average correction factor is about 0.7.

Form factor corrections for the $B \to D^{(*,**)}\ell\nu$ decays

The shapes of the distributions of the kinematic variables are dictated by the form factors that enter the model, thus their precise knowledge will impact on the model to describe a given decay. The study of the form factors for the $B \to D^{(*,**)}\ell\nu$ decays was performed in reference [94], in which the weights are provided in bins of two variables, either the momentum transfer q^2 and the lepton momentum in the CMS p_ℓ^* as is the case for the $B \to D^{(*)}\ell\nu$ decays, or the recoil variable w and $\cos\theta_\ell$ for $B \to D^{**}\ell\nu$ decays[6]. The decays $B \to D\ell\nu$ and $B \to D^*\ell\nu$ are simulated in the generic MC using the heavy quark effective theory model (HQET2). Their form factors parameters are updated using recent values from HFLAV [41][7], as shown in table 5.3.

The decay $B \to D^{**}\ell\nu$ was produced using the ISGW2 quark model that is proven to be insufficient to describe the decay. As a consequence the decay is re-weighted to the model Leibovich-Ligeti-Stewart-Wise

[6] The recoil variable is defined as $w = \frac{M_B - M_D - q^2}{2M_B M_D}$.

[7] Heavy Flavor Averaging Group, previously shortened as HFAG.

(LLSW) [95] using recent results from HFLAV.

Branching ratio correction

Because of the copious decay channels simulated in the generic MC sample, it is not feasible to update all the branching ratios to the world average provided by the Particle Data Group (PDG) [1]. Instead, one can compare data and MC samples in regions dominated only by the background, which are usually called sidebands, and check whether or not the MC describes the data in those regions. From this comparison, one can then correct the channels that have a major impact in the modeling of the background for the particular signal channel. Decays such as $B \to D^{(*,**)}\ell\nu$, or exclusive charmless semileptonic decays with meson masses below 1 GeV, are corrected according to the measured values provided either by the PDG or the HFLAV. The correction factors for these decay channels are arranged in table 5.4.

5.2 Event Selection

The first requirement of the analysis lies in filtering hadronic events to reject physical processes with a larger cross section than the $B\bar{B}$ pair production, as shown in table 5.5. This is accomplished by applying the "HadronB" skim [96] to the data, which is a pre-selection tool common to all Belle analyses, and has a good performance in reducing non-hadronic background while retaining most of the B decays. The requirements in this skim are based on the fact that $B\bar{B}$ events have several particles in the final state, and hence have a larger track and cluster multiplicity than the other physical processes, subsequently the energy of the decay products of B decays exhibit a low energy spectrum. Among these requirements it is noteworthy to mention that with the presence of at least three tracks, the total energy in the ECL should be above 18% and below 80% of the CMS energy, and a minimum for the visible energy of 20% the CMS energy. The latter term is defined as the sum of "good tracks[8]" momenta and "good photons energy[9]". After the application of the HadronB skim, the dominant background corresponds to continuum events ($e^+e^- \to q\bar{q}$), which are roughly three times larger than $B\bar{B}$ processes. Under the assumption that all tracks and clusters in the reconstruction of events not belonging to the B_{tag} must be associated with the B_{sig}, one can establish additional selections to reject background and build the signal decay. This section presents the definition of the objects used to reconstruct the signal decay as well as different variables used across this analysis, describing a set of conditions known as pre-selection, and finally discusses the application of a BDT for further rejection of background. Some cuts in the pre-selection stage or in the selection of objects are not optimized in the sense of choosing the best cut according to a figure of merit (FOM), but rather standard selections used for semileptonic B decays that take into account the acceptance of the different detector components, or well-established cuts within the Belle collaboration.

5.2.1 Definition of objects

The visible objects on the signal side consist of charged leptons and charged pions, however some variables require the establishment of neutral particles such as neutral pions or photons. The prerequisite for these particles are given below.

[8] Tracks with a minimum transversal momentum of 100 MeV and impact parameters satisfying $dr < 2$ cm and $dz < 4$ cm.
[9] Clusters with energies above 100 MeV that cannot be associated with tracks in the CDC.

Decay Mode	MC	PDG/HFLAV	Weight
$B^- \to D^0 \ell^- \bar{\nu}_\ell$	2.31×10^{-2}	$(2.33 \pm 0.10) \times 10^{-2}$	1.009
$B^- \to D^{*0} \ell^- \bar{\nu}_\ell$	5.79×10^{-2}	$(5.59 \pm 0.19) \times 10^{-2}$	0.965
$B^- \to D_1^0 \ell^- \bar{\nu}_\ell$, $D_1^0 \to D^{*+}\pi^-$	5.4×10^{-3}	$(2.8 \pm 0.1 \pm 1.5) \times 10^{-3}$	0.52
$B^- \to D_2^{*0} \ell^- \bar{\nu}_\ell$, $D_2^{*0} \to D^{*+}\pi^-$	8.2×10^{-4}	$(7.7 \pm 0.6 \pm 0.4) \times 10^{-4}$	0.94
$B^- \to D_1^{'0} \ell^- \bar{\nu}_\ell$, $D_1^{'0} \to D^{*+}\pi^-$	5.4×10^{-3}	$(1.3 \pm 0.3 \pm 0.2) \times 10^{-3}$	0.24
$B^- \to D_0^{*0} \ell^- \bar{\nu}_\ell$, $D_0^{*0} \to D^+\pi^-$	6.1×10^{-3}	$(2.8 \pm 0.3 \pm 0.4) \times 10^{-3}$	0.46
$\bar{B}^0 \to D^0 \ell^- \bar{\nu}_\ell$	2.13×10^{-3}	$(2.20 \pm 0.10) \times 10^{-2}$	1.033
$\bar{B}^0 \to D^{*+} \ell^- \bar{\nu}_\ell$	5.33×10^{-3}	$(4.88 \pm 0.10) \times 10^{-2}$	0.916
$D^0 \to K^0 f_0(980)$, $f_0(980) \to \pi^+\pi^-$	6.24×10^{-3}	$(2.46^{+0.80}_{-0.48}) \times 10^{-3}$	0.38
$D^0 \to K^0 f_0(1370)$, $f_0(1370) \to \pi^+\pi^-$	1.44×10^{-3}	$(5.6^{+1.8}_{-2.6}) \times 10^{-3}$	3.8
$B^- \to \pi^0 \ell^- \bar{\nu}_\ell$	7.8×10^{-5}	$(7.80 \pm 0.27) \times 10^{-5}$	1.068
$B^- \to \rho^0 \ell^- \bar{\nu}_\ell$	1.49×10^{-4}	$(1.58 \pm 0.11) \times 10^{-4}$	1.06
$B^- \to \omega \ell^- \bar{\nu}_\ell$	1.15×10^{-4}	$(1.19 \pm 0.09) \times 10^{-4}$	1.035
$B^- \to \eta \ell^- \bar{\nu}_\ell$	8.4×10^{-5}	$(3.8 \pm 0.6) \times 10^{-5}$	0.45
$B^- \to \eta' \ell^- \bar{\nu}_\ell$	3.3×10^{-5}	$(2.3 \pm 0.8) \times 10^{-5}$	0.70
$\bar{B}^0 \to \pi^+ \ell^- \bar{\nu}_\ell$	1.36×10^{-4}	$(1.45 \pm 0.05) \times 10^{-4}$	1.066
$\bar{B}^0 \to \rho^+ \ell^- \bar{\nu}_\ell$	2.77×10^{-4}	$(2.94 \pm 0.21) \times 10^{-4}$	1.061

Table 5.4: Decay channels that were corrected in the MC, with their branching ratios in the MC, world average and their respective correction (weight).

Process	$\sigma_{\text{no skim}}$ [nb]	σ_{skim} [nb]	ϵ (%)
$e^+e^- \to b\bar{b}$	1.1	1.09	99.1
$e^+e^- \to q\bar{q}$	3.7	2.62	70.8
$e^+e^- \to \tau^+\tau^-$	0.93	0.05	5.38
Bhabha	37.8	0.001	0.003
$\gamma\gamma$	11.1	0.04	0.36

Table 5.5: Cross sections before and after the application of the HadronB skim and its efficiencies [96]. Hadronic states containing b quarks remain intact after the application of this skim, while other processes, except continuum events, are almost eliminated.

Selection of tracks

Due to the magnetic field inside the detector, charged particles with low momenta spiral inside the CDC and may lead to multiple reconstructions of the same particle. A pair of tracks can be regarded as duplicated tracks if they have momenta transverse to the beam direction below 275 MeV, with a small momentum difference (below 100 MeV) and an opening angle either below 15° (same charges) or above 165° (opposite charges). Once such a pair is found, the track passing closer to the IP is chosen by minimizing a χ^2-like variable based on the impact parameters (defined as $\chi^2 = 5|dr|^2 + |dz|^2$). All remaining tracks are required to satisfy $|dr| < 0.4$ cm and $dz < 2.0$ cm.

To identify tracks as charged leptons, different conditions are imposed on the particle likelihood (see

Electron	Muon
$p_e^{lab} > 0.3$ GeV	$p_\mu^{lab} > 0.6$ GeV
$17^\circ < \theta_e^{lab} < 150^\circ$	$25^\circ < \theta_\mu^{lab} < 145^\circ$
$\mathcal{L}_{eid} > 0.9$	$\mathcal{L}_{muid} > 0.9$

Table 5.6: Selection requirements in momentum, polar angle and particle identification likelihood for leptons.

eq. 3.1 for electrons and eq. 3.2 for muons), the polar angle and the momenta of the tracks evaluated in the lab frame. These conditions are summarized in Tab 5.6 for both electrons and muons. To reject leptons originating from J/ψ or ψ' decays, the invariant mass of a pair of oppositely charged leptons ($m_{\ell^+\ell^-}$) is inspected and the particles are discarded if they satisfy 3.00 GeV $< m_{\ell^+\ell^-} <$ 3.15 GeV or 3.60 GeV $< m_{\ell^+\ell^-} <$ 3.75 GeV. Only events with a single charged lepton are considered in this study. In the case of the electron channel, photons not belonging to the tag side with a maximum energy in the lab frame of 1 GeV and lying inside a cone around the electron direction smaller than 5° ($\theta_{e\gamma} < 5^\circ$) are associated to Bremsstrahlung processes. If such a photon is found, it is merged with the electron and the sum of the momenta is assumed to be the lepton four-momentum. If there is more than one photon candidate only the nearest photon is merged with the electron.

After excluding tracks from the tag side, a track not used as a lepton is accepted as a charged pion if the discriminating probability between kaons and pions $\texttt{Prob}(K)$ (see eq. 3.3) is below 0.4.

Selection of neutral particles

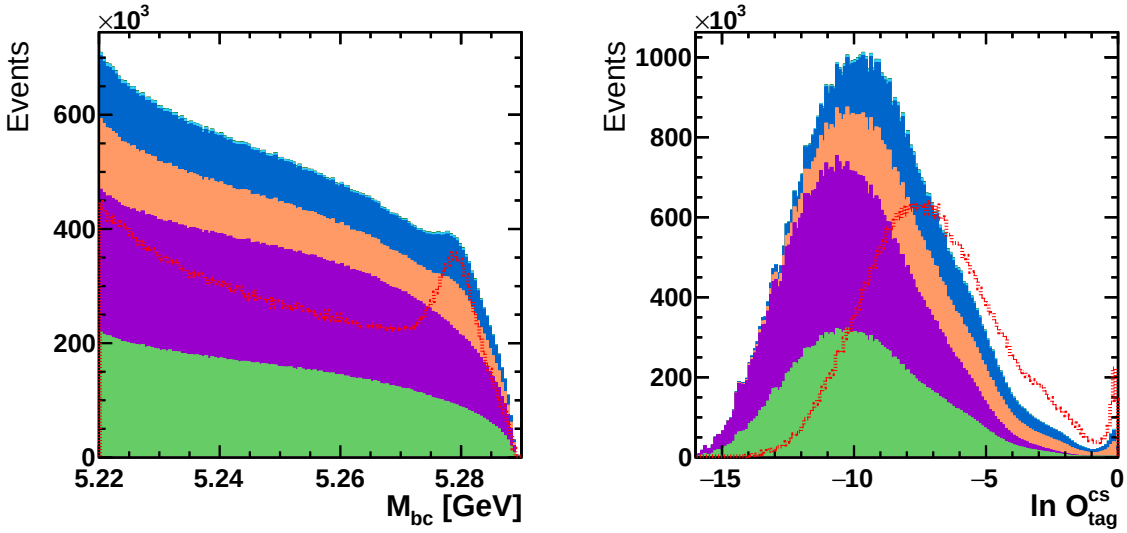
Despite the absence of an explicit neutral particle in the signal channel, some of the variables used in the analysis for background rejection request a clear definition of these objects. Photons are identified as clusters in the ECL not linked to a track in the CDC. As the ECL can be separated in regions according to the photon polar angle in the lab frame θ_γ, a minimum energy requirement of 50 MeV, 150 MeV or 100 MeV are imposed in the barrel ($32^\circ < \theta_\gamma < 130^\circ$), the forward endcap ($130^\circ < \theta_\gamma < 150^\circ$) or the backward endcap ($17^\circ < \theta_\gamma < 32^\circ$), respectively. Neutral pions are reconstructed from a pair of photons such that their invariant mass lies between 120 MeV and 150 MeV.

5.2.2 Pre-selection for $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decays

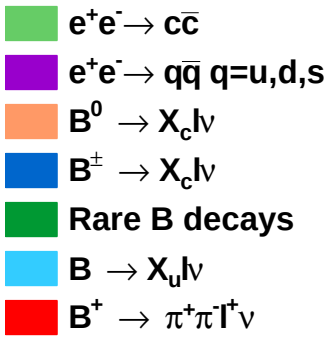
A powerful variable to remove combinatorial events in B decays, and hence guarantee well reconstructed candidates, is the beam constrained mass, which is defined as:

$$M_{bc} = \sqrt{E_{\text{beam}}^{*\,2} - p_B^{*2}}, \tag{5.7}$$

where E^*_{beam} and p^*_B are the beam energy and the momentum of the B meson, respectively, and both are evaluated in the $\Upsilon(4S)$ rest frame. In this definition, the beam energy replaces the energy of the B meson since it is known better than the latter quantity and therefore results in an improvement in resolution and separation power. This variable is calculated for the B_{tag} meson and its distribution is shown in Fig. 5.1(a). For the signal and B decays in general, this distribution has a peak structure around the nominal mass of the B meson, and a long tail due to random combinations; for continuum events this peaking structure is non-existent. Following the selection from previous analyses in B decays, this variable is required to be above 5.27 GeV, for which the resonant (correctly reconstructed) events exhibit a Gaussian shape.



(a) M_{bc} (b) $\ln o^{cs}_{tag}$



(c) Physics processes

Figure 5.1: Distributions at the start of the selection of the beam constrained mass (a) and the logarithm of the NeuroBayes output with continuum suppression (b) showing the contributions from different processes (c). The shape of the signal contribution (red dashed line) is scaled with an arbitrary normalization for comparison purposes. The Gaussian peak in the signal corresponds to correctly reconstructed events.

Another variable especially helpful in reducing continuum background is the output of the NeuroBayes neural network with continuum suppression $o^{\mathrm{cs}}_{\mathrm{tag}}$ that arises from the reconstruction of the B_{tag} meson, the logarithm of this variable is presented in Fig. 5.1(b). To determine on which value to cut to select the signal, the FOM precision is optimized. This FOM is defined as:

$$FOM = \frac{S}{\sqrt{S+B}}, \tag{5.8}$$

where S and B represent the number of events in signal and background, respectively. This FOM can be modified such that it can be expressed in terms of familiar quantities. Assume, for instance, that N_S is the total number of generated signal events, then $S = \epsilon N_S$ where ϵ is the signal reconstruction efficiency

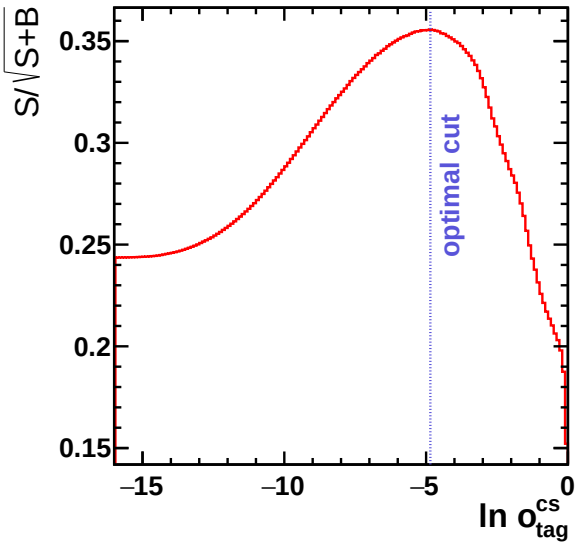


Figure 5.2: Figure of merit and optimal cut for $\ln o^{\text{cs}}_{\text{tag}}$.

resulting from the selection. Plugging in this factor into the numerator of eq. 5.8, the FOM becomes $\sqrt{N_S}\sqrt{\epsilon S/(S+B)} = \sqrt{N_S}\sqrt{\epsilon p}$, with p the purity. This relation makes explicit the compromise between efficiency and purity in the optimization. The optimal cut value on this variable is given by $\ln o^{\text{cs}}_{\text{tag}} > -4.9$ as shown in Fig. 5.2.

In semileptonic decays, leptons not stemming from a B meson are also a source of background; they generally result from cascade decays of intermediate particles such as D or D^* mesons. To veto these events, one can exploit one of the features of the hadronic tag method involving the knowledge of the charge of the B meson, and the fact that leptons from secondary decays carry an opposite charge to that of the prompt lepton. Accordingly, a negative charge correlation between the signal prompt lepton and the B_{tag} is requested. Furthermore, only charged B mesons are accepted in the analysis, which removes background from neutral B mesons. The previous four requirements complete the selection on the tag side. Hereafter, the selection is focused on the signal side. The topology of the signal channel involves three prongs; the two charged pions and the lepton. As a result, only three tracks are required on the signal side.

Since the neutrino is not actually reconstructed in the event due to its invisibility to the detector, the neutrino information is inferred from the missing momentum of the event. For a good reconstructed semileptonic decay, where the neutrino is the only lost particle, the previous quantities should be unchanged. The relation between the four-momenta of all the particles in the event and the missing four-momentum is defined as

$$P_{\text{miss}} = P_{\Upsilon(4S)} - P_{B_{\text{tag}}} - P_\ell - P_{\pi^+} - P_{\pi^-}, \tag{5.9}$$

and hence the missing mass squared M^2_{miss} is given by

$$M^2_{\text{miss}} = |P_{\text{miss}}|^2. \tag{5.10}$$

This variable is the most powerful discriminator in semileptonic decays with hadronic tags. For the

signal events the M^2_{miss} peaks around the mass of the neutrino, $M^2_{\text{miss}} \approx 0$, while for the background the tendency is towards high values as can be noticed in Fig. 5.3(d). Given the large range of this variable, and the fact that the signal is concentrated in the $|M^2_{\text{miss}}| < 1$ GeV^2 region, the extent of this variable is reduced to the interval $[-2.0\ \text{GeV}^2, 6.0\ \text{GeV}^2]$. Fig. 5.3 illustrates the effect of the previous requirements in the pre-selection variables, taking out the cut on the plotted distribution. The subsequent improvement of the signal selection for each requirement can be seen in the cut-flow of table 5.7, where the efficiencies for signal and background are shown at each stage in the pre-selection. The number of events for the signal also gets reduced as many random combinations of tracks and clusters contribute to the reconstruction of the decay channel. The idea of the pre-selection is then to correctly define the signal.

Criteria	Signal	$B \to X_u \ell \nu$	B^+ Generic	B^0 Generic	Continuum	Rare
$M_{\text{bc}} > 5.27$ GeV	21.8	20.5	19.9	19.3	16.5	23.1
$\ln(o^{\text{cs}}_{\text{tag}}) > -4.9$	8.61	7.81	5.90	6.03	0.546	4.88
$Q_\ell \times Q_{B^{\text{tag}}} < 0$	7.84	6.88	3.96	3.23	0.299	2.84
$B^{\text{tag}} = B^\pm$	7.20	4.79	3.20	0.352	0.134	1.62
$N_{\text{signal-tracks}} = 3$	6.11	3.76	0.752	0.045	0.007	0.263
$M^2_{\text{miss}} \leq 6$ GeV^2	6.01	3.29	0.381	0.018	0.002	0.142

Table 5.7: Relative efficiency [%] for signal and background processes respect to the merging of objects in the reconstruction of $B^+ \to \pi^+\pi^-\ell^+\nu$ decays at the pre-selection stage.

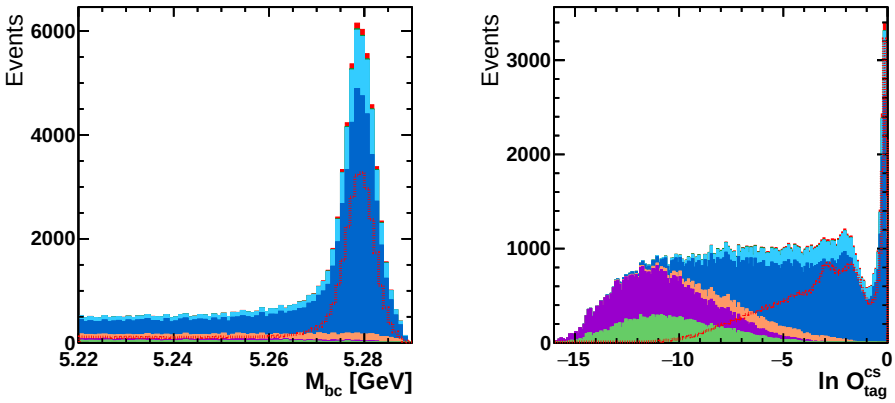
5.2.3 Background Suppression with Boosted Decision Trees

Although the pre-selection disposes of a vast amount of continuum background, the remaining events are dominated by semileptonic decays to charm mesons. Variables that distinguish between this background and the signal are difficult to find, since most of the kinematic distributions for the background resemble the signal. In addition, any potential variable for background suppression must have as little correlation as possible with the five variables that describe the signal decay and with M^2_{miss}, to ensure the access to the whole phase space of the signal decay. Since the available variables do not allow a clear separation between signal and background, a cut-based analysis is not sufficient for the rest of the analysis. Instead, an stochastic gradient Boosted Decision Trees (BDTG) is trained to learn the subtle patterns of the signal in the simulated data and hence improve the separation between signal and background. This method is available in the TMVA [74] toolkit of the ROOT [82] framework. To carry out the BDTG algorithm, the signal and generic B decays MC samples are split into two samples, each containing half of their initial number of events. One sample corresponds to the training sample used for the background studies and selection in this chapter, and the other to the validation sample on which the measurement is performed.

Input variables of the BDTG

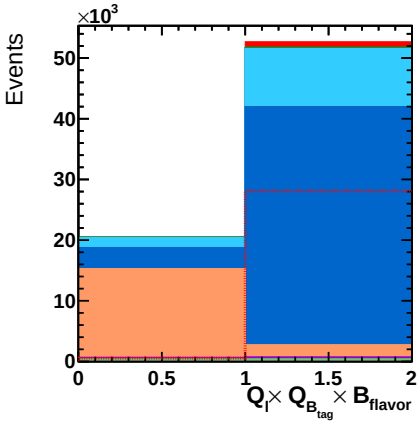
The definitions of the variables implemented in the training of the BDTG are as follows.

1. $\Delta\mathbf{E}_{\text{sig}}$: the difference between the beam and the B_{sig} meson energies in the CMS.
2. θ_{miss}: the polar angle of the missing momentum in the lab frame.
3. $\mathbf{N}_{\pi^0}$: the multiplicity of π^0 candidates on the signal side.

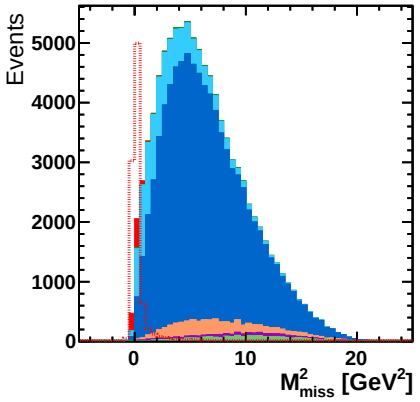


(a) M_{bc}

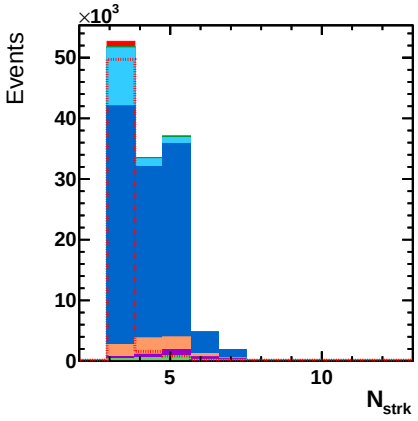
(b) $\ln o^{\mathrm{cs}}_{\mathrm{tag}}$



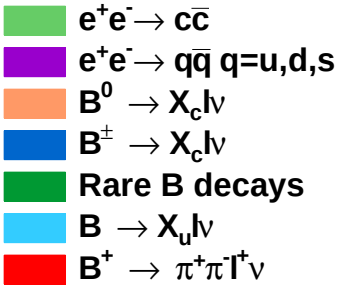
(c) Charge correlation and B flavor



(d) M^2_{miss}



(e) Track multiplicity on signal side



(f) Physics processes

Figure 5.3: Distributions of the pre-selection variables when all cuts are applied except the one on the variable plotted: (a) M_{bc}, (b) $\ln o^{\mathrm{cs}}_{\mathrm{tag}}$, (c) $Q_\ell \times Q_{B_{\mathrm{tag}}} \times B_{\mathrm{flavor}} = 1$ implies a negative charge correlation between the signal lepton and the B_{tag} and the selection of charged B mesons, otherwise $Q_\ell \times Q_{B_{\mathrm{tag}}} \times B_{\mathrm{flavor}} = 0$, (d) M^2_{miss}, (e) number of tracks on the signal side, (f) the physics processes. The shape of the signal contribution (red dashed line) is scaled with an arbitrary normalization for comparison purposes.

4. $\delta\phi_{\text{vtx}}^{\mathbf{p}_{\mathbf{xu}}}$: the angle between the dipion momentum in the lab frame and the vector joining the $\pi^+\pi^-$ decay vertex and the interaction point.

5. $\mathbf{E}_{\text{extra-clusters}}$: the sum of the energy in the CMS of all photons in the barrel region not associated to either the B_{tag} or the B_{sig} mesons.

6. $\mathbf{E}_{\text{ECL}}$: the sum of the clusters in the ECL from the whole event that are not matched to tracks and that pass the energy thresholds for photons [97]. This calculation also includes ECL clusters made by photons that were accidentally hit by a track and satisfy $E9/E25 > 0.94$. The latter is the transverse shower shape in the ECL, defined as the ratio of energy deposited in the central 3×3 array of crystals to that in the central 5×5 array of crystals. This variable is suitable to separate between overlapping hits in the ECL crystals caused by hadronic interaction with charged tracks and photons, for photons $E9/25$ peaks at one whereas for charged tracks tend to low values.

Parameter settings for BDTG training

The parameters used in the configuration for the BDTG are described as follows:

- A forest of 1000 DTs is used in the training.
- The number of steps, `nCuts`, is set to the default value of 20, increasing this number decreases the signal probability in the Kolmogorov-Smirnov test, which is an indication of how likely the algorithm is to perform under overtraining.
- The node depth, `maxDepth`, is optimised to 3.
- The minimum percentage of training events required in a leaf node is set to 2%.
- This method employs a random subsample of all events for growing the trees in each iteration by selecting the option `UseBaggedBoost=True` for which the relative fraction of events is set by `BaggedSampleFraction=0.37` .
- The learning rate (`shrinkage`) for the GradBoost algorithm is set to 0.1.

Results from the BDTG training

The shape comparisons between the signal and $B \to X_c \ell \nu$ background in the input variables used in the BDTG training are shown in Fig. 5.4. The linear correlations among them are small as can be seen in Fig. 5.5, and their separation power is ranked in table table 5.8. An additional check was made to verify the consistency between the training sample and the test sample by comparing their BDTG response depicted in Fig. 5.6, showing a low probability of occurrence for overtraining.

The optimal cut on the BDTG is searched for using the FOM of eq. 5.8 and is found to be 0.52. According to Fig. 5.7 this value remains constant even if the assumed branching ratio for the signal is duplicated or halved.

5.2.4 MC study of the $B \to X_c \ell \nu$ background

The contributions of all physical processes to the BDTG output is presented in Fig. 5.8, where the signal shape is shown in the red dashed line with an arbitrary normalization. Even after the application of the optimal cut on this classifier, the remaining events are still dominated by the background, as can be

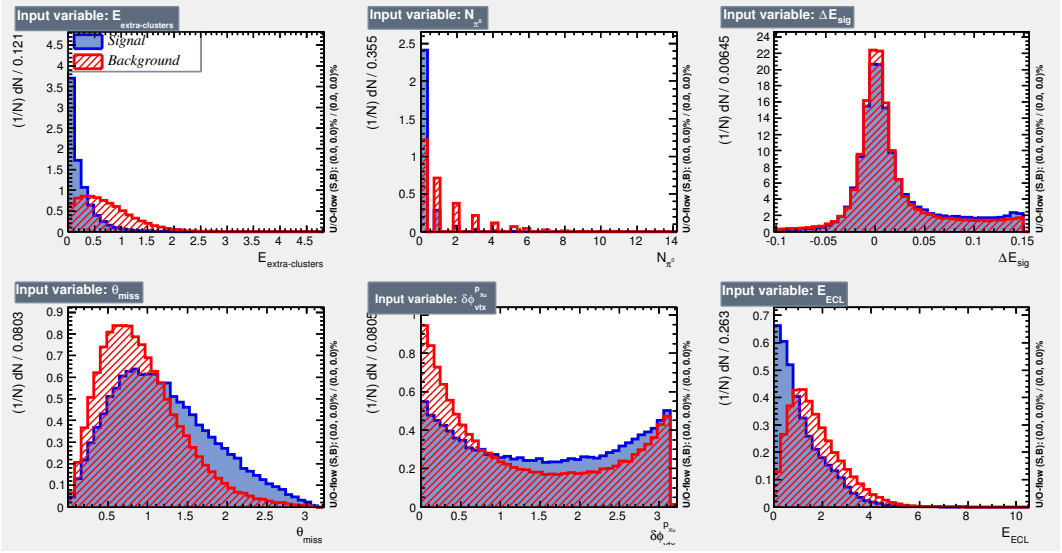
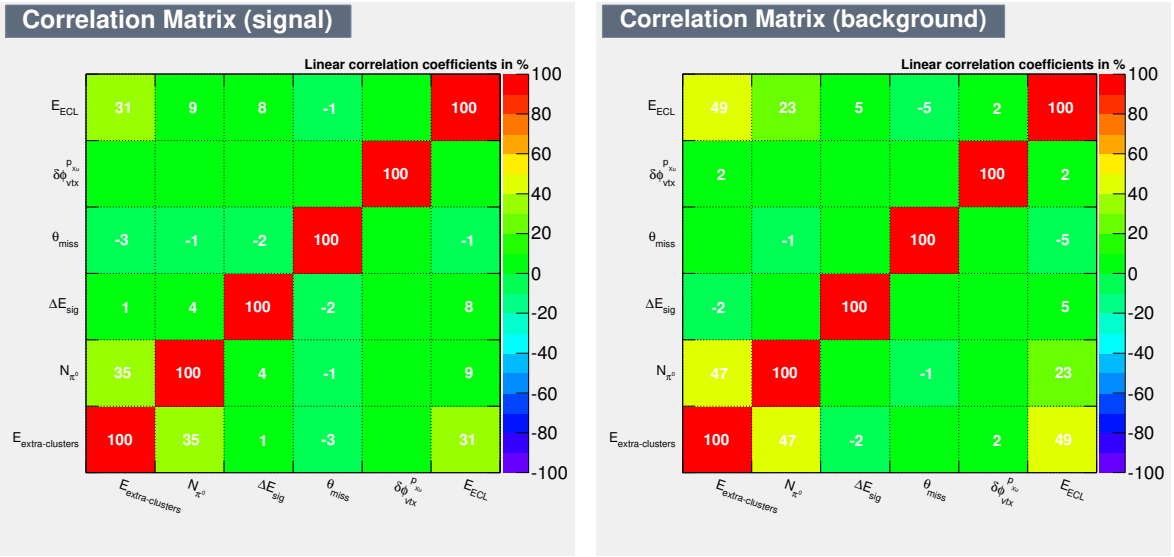


Figure 5.4: Shape comparison of input variables in the BDTG for the signal and $B \to X_c \ell \nu$ samples.



(a) Signal

(b) $B \to X_c \ell \nu$

Figure 5.5: Linear correlation matrix among input variables for signal and $B \to X_c \ell \nu$ background samples.

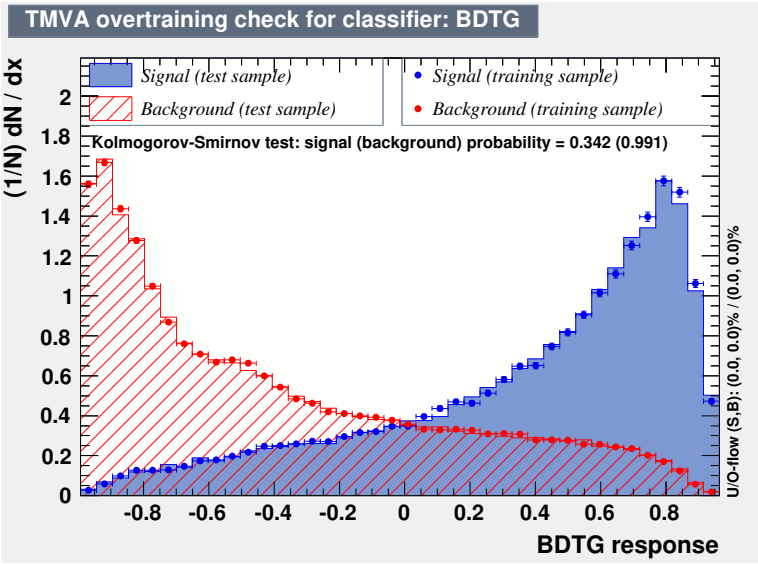


Figure 5.6: Comparison between BDTG output from the training and test samples to check overtraining.

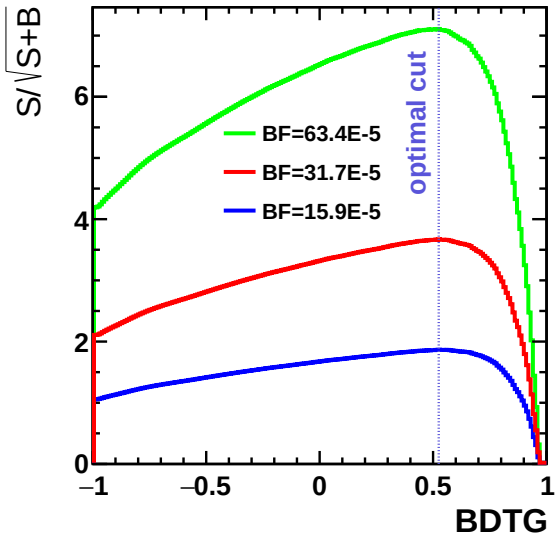


Figure 5.7: Optimal cut in the BDTG output for different assumptions of the branching ratio for the signal, this value remains constant.

Rank	Variable	Separation	Rank	Variable	Separation
1	$E_{\text{extra-clusters}}$	3.13×10^{-1}	4	θ_{miss}	5.73×10^{-2}
2	N_{π^0}	2.08×10^{-1}	5	$\delta\phi^{p_{xu}}_{vtx}$	2.52×10^{-2}
3	E_{ECL}	1.05×10^{-1}	6	ΔE_{sig}	3.56×10^{-3}

Table 5.8: Ranking of variables according to their separation power in the training of the output of the BDTG classifier.

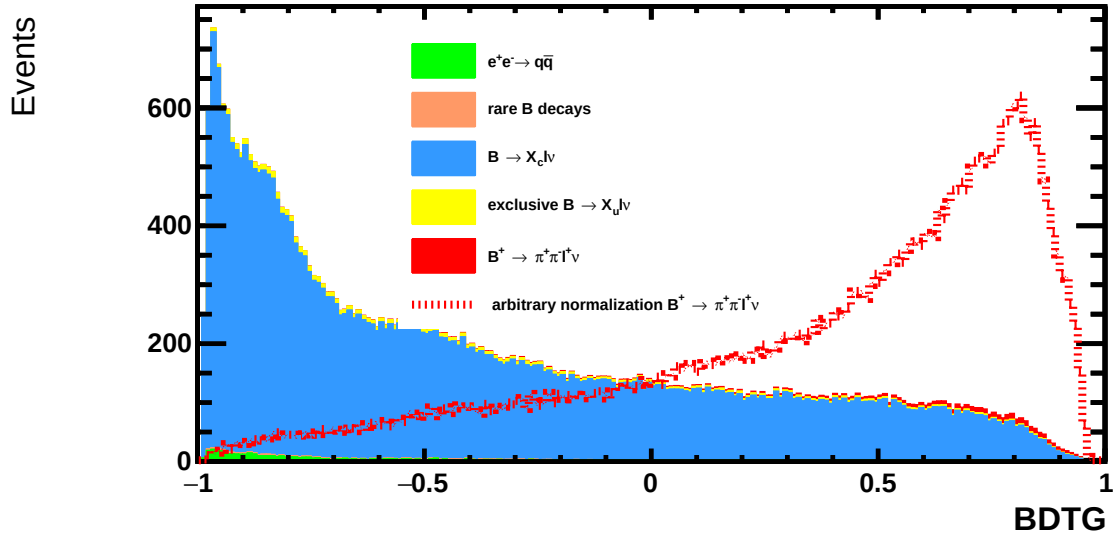


Figure 5.8: BDTG output with the expected MC contributions showing the tendency for signal-like events to have values closer to 1. The signal contribution has been scaled to an arbitrary normalization for illustrative purposes.

observed in the M^2_{miss} distribution of Fig. 5.9. This variable is used to split the MC into two regions: the signal region ($|M^2_{\text{miss}}| < 2$ GeV^2) and the sidebands (2 GeV^2 $< M^2_{\text{miss}} <$ 6 GeV^2), where the prominent background is due to $B \to X_c \ell \nu$ decays. This region will be used later in this chapter to check how well the MC describes the real data.

To study which decays make up the $B \to X_c \ell \nu$ background, one can check the mother particle identity (mother ID) for each charged pion, as visualized in Fig. 5.10. According to this, in the majority of the cases the pions come from D^0 or ρ^0 mesons, but they can also originate from different mother particles, in which case they are produced at different stages of a decay chain. To have a clear picture of the implications of this diagram, the mother ID of the pions is projected to the M^2_{miss} and invariant mass $M_{\pi^+\pi^-}$ spectra, as shown in Fig. 5.11. The peaking structures in the di-pion mass are, in general, due to cases where both pions arise from a single resonance, e.g., D^0, ρ^0, ω, $f_0(980)$. A prominent peak can be observed around $M_{\pi^+\pi^-} = 1.3$ GeV, with one pion stemming from a $K^{*\pm}$ meson and the other from a D^0 meson. The former decays into $K^0\pi^\pm$, where the missing neutral particle failed to be reconstructed. The neutral kaon long (K_L) passes through the Belle detector without depositing energy in the ECL in most of the cases. It can be partially reconstructed in the KLM but with a poor efficiency. No attempt is made to reduce this peaking background, since a K_L veto will incur in the introduction of additional systematic

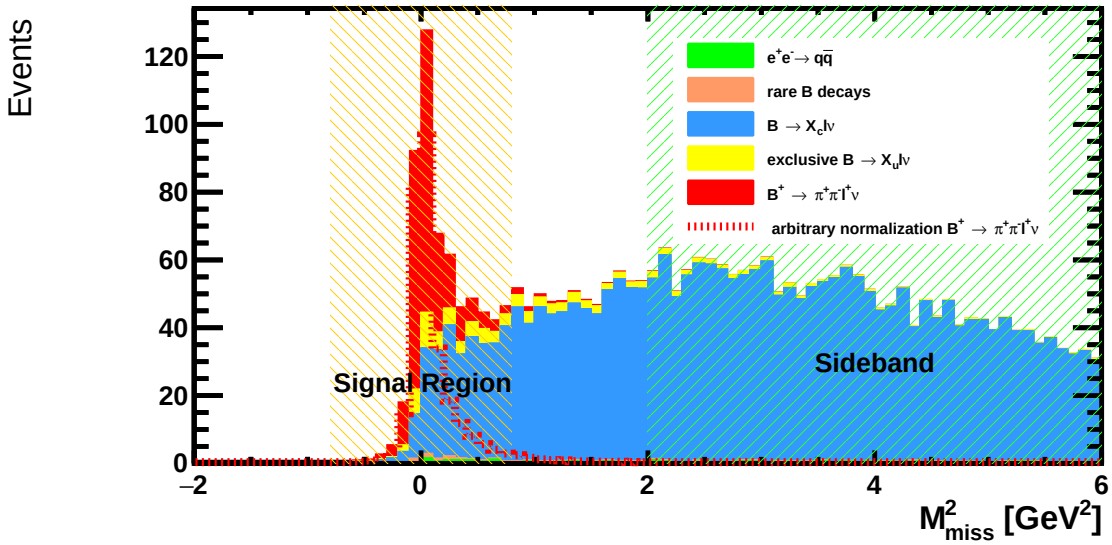


Figure 5.9: Missing mass squared distribution with the expected MC contributions after the cut on the BDTG classifier. The orange box depicts the signal region, while the green one shows the sideband.

Decay	Expected Events	Relative Efficiency [%]
$e^+e^- \to q\bar{q}$	8.34	2.9×10^{-7}
Rare B decays	6.41	8.0×10^{-3}
$B^+ \to D^0\ell^+\nu_\ell$ with $D^0 \to \pi^+\pi^-$	34.30	7.9×10^{-2}
Other $B^+ \to D^0\ell^+\nu_\ell$	114.37	5.0×10^{-4}
Rest of $B \to X_c\ell^+\nu_\ell$	146.43	1.6×10^{-5}
$B \to X_u\ell^+\nu_\ell$	27.05	1.0×10^{-3}
$B^+ \to \pi^+\pi^-\ell^+\nu_\ell$	387.10	2.9×10^{-2}

Table 5.9: Expected events after signal selection and BDTG cut in the signal region for the processes contributing to the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay channel. The relative efficiency at this stage with respect to the merging of objects in the reconstruction of $B^+ \to \pi^+\pi^-\ell^+\nu$ decays at the pre-selection stage is also provided.

uncertainties. In a much smaller scale, pions coming from D^0 and $\rho^\pm$ contribute to the region defined by 1.4 GeV $< M_{\pi^+\pi^-} <$ 1.7 GeV. All these components are primarily present in the sideband, since most of these peaking structures disappear in the signal region for $B \to X_c\ell\nu$ component (see Fig. 5.12), resulting in a flat background for di-pion masses less than 1.6 GeV. Examples of B decay chains containing these processes for the generic MC are given below:

- $B^+ \to D^0\bar{D}^0K^+\pi^0$, $\bar{D}^0 \to \pi^-K^{*+}(K^0\pi^+)$, $D^0 \to K_sK^+\pi^-$
- $B^+ \to \bar{D}^{*0}\ell^+\nu_\ell$, $\bar{D}^{*0} \to \bar{D}^0\pi^0$, $\bar{D}^0 \to K^{*+}(K^0\pi^+)\pi^-$,
- $B^- \to D^0\ell^-\bar{\nu}_\ell$, $D^0 \to f_0(980)\bar{K}^0$, $f_0(980) \to \pi^+\pi^-$

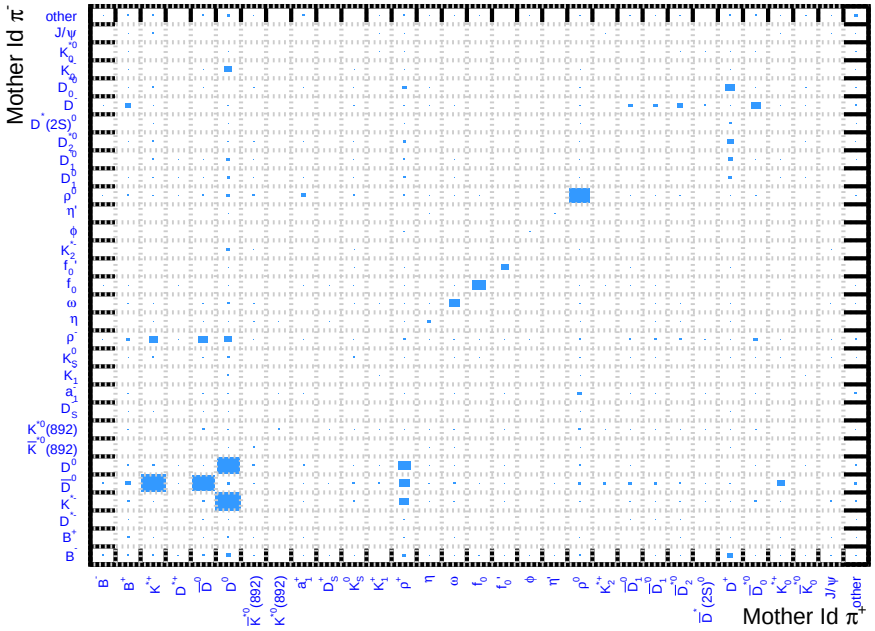
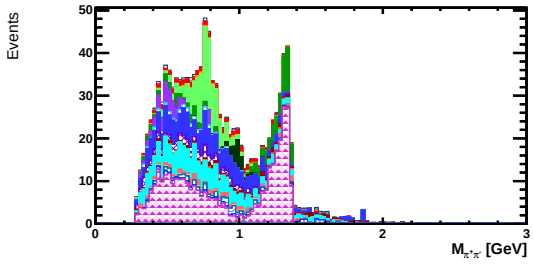
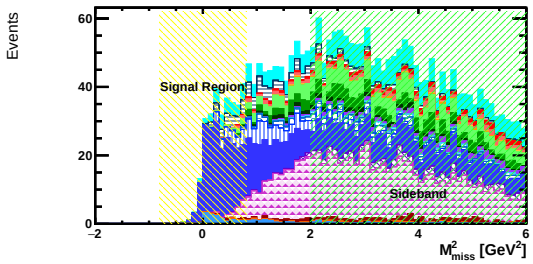


Figure 5.10: Mother ID of two charged pion combination for the $B \to X_c \ell \nu$ background. This figure shows that the pion originate from the immediate decay of another particle, that may be of another decay chain. The two pions could either share the same mother or come from different particles in the decay chain. The size of the grid indicates the frequency in which these cases occur.

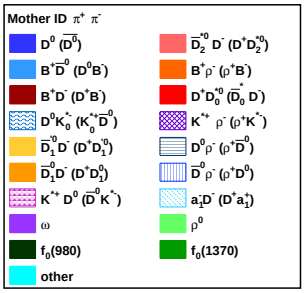
One can further study the $B \to X_c \ell \nu$ background by dividing the sample into different categories. Firstly, the particle identity of the charged pions is verified to determine if the other charged particles have been misidentified as pions and have some effect in the production of peaking structures in the hadron mass. As shown in Fig. 5.13, muons and kaons can be misidentified as pions producing a small flat contribution along the di-pion mass, from them the charged kaons are responsible for the kinematic reflection around $M_{\pi^+\pi^-} = 1.75$ GeV. Another category consists in establishing the origin of the two charged pions and the lepton, whether or not they come from a $B^\pm$ meson at some point in the decay chain. According to Fig. 5.14(a), only a small amount of events may have a lepton coming from one B and pions from another B meson. If now instead, one evaluates the origin of the lepton to corroborate that it stems from a B decay (prompt lepton), comes from an intermediate resonance like the D^0 meson (secondary lepton) or is another charged particle mislabeled as lepton (fake lepton), one can find that the contributions from fake and secondary leptons are negligible compared to prompt leptons as displayed in Fig. 5.14(b). A fourth category implies the characterization of the $B \to X_c \ell \nu$ background by the intermediate charm resonance (see Fig. 5.14(c)), in which case the D^* meson dominates the mass spectrum followed by the D meson. A small contribution from excited states D^{**} such as D_1, D_1', D_2^* and D_0^* is also expected. A final inspection of the mass spectrum in the $B \to X_c \ell \nu$ background can be done using truth matching of the final state charged pions, as shown in Fig. 5.15, which reiterates the aforementioned statements. The biggest contribution comes from decays where the two pions shared the same mother but one massive particle failed to be reconstructed as is the case for K_L mesons, the kinematic reflection at $M_{\pi^+\pi^-} = 1.75$ GeV is caused by misidentified pions, other notorious peaking



(a) Composition of $M_{\pi^+\pi^-}$ for $B \to X_c \ell \nu$



(b) Composition of M^2_{miss} for $B \to X_c \ell \nu$

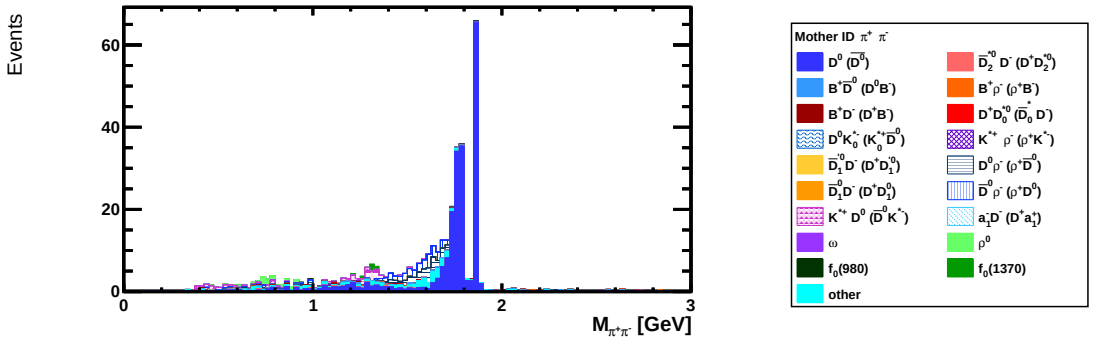


(c) Legend

Figure 5.11: Projection of the mother ID of the pions for the $B \to X_c \ell \nu$ background in the $B^- \to \pi^+\pi^-\ell^-\bar{\nu}_\ell$ channel. The first and the second particle in the legend correspond to the mother of the π^+ and π^-, respectively. In the case where there is only one particle, indicates that both pions come from the same mother. The figures reveal that peaking structures in the di-pion mass come from processes in the sidebands, for instance the light green component peaks in the di-pion mass around 0.7 GeV corresponding to events where both pions come from a ρ^0 meson, however this component is only present in the sidebands in the M^2_{miss} distribution, from which it can be inferred that there are additional particles missing in the reconstruction and hence the ρ^0 originates at some point in a cascade decay.

structure such as the one around $M_{\pi^+\pi^-} = 1.87$ GeV corresponds to correctly reconstructed particles, in this case the D^0 meson which proceeds according to the decay chain $B^- \to D^0\ell^-\bar{\nu}_\ell$ with $D^0 \to \pi^+\pi^-$.

Up to now the focus of the background has been the $B \to X_c \ell \nu_\ell$ background in the whole range of the M^2_{miss} distribution. However, at the moment of signal extraction, the background that dominates the signal region will have a major impact in the correct determination of the signal yields. An estimate of the contributions from different processes after the event selection and BDTG cut in the signal region is presented in table 5.9. This table also shows the effect of the complete selection on each physical process since the merging of objects in the reconstruction of $B^+ \to \pi^+\pi^-\ell^+\nu$ decays. This effect is denoted as relative efficiency. According to this table, the contributions from other processes but the $B \to X_c \ell \nu_\ell$ are expected to be tiny in the signal region and hence an additional selection to reject them is not necessary. The $B \to X_c \ell \nu_\ell$ component has been separated into three contributions: $B^+ \to D^0\ell^+\nu_\ell$ with $D^0 \to \pi^+\pi^-$, other $B^+ \to D^0\ell^+\nu_\ell$ and rest of $B \to X_c\ell^+\nu_\ell$. The lighter intermediate charm resonance in the $B \to X_c \ell \nu_\ell$ component is the D^0 meson, this meson has a large number of decay channels, some of them have a large contribution in the signal region, especially the $D^0 \to \pi^+\pi^-$ channel whose topology is exactly the same as the signal but does not involve a $b \to u$ quark transition. Separating it from the



(a) Composition of $M_{\pi^+\pi^-}$ for $B \to X_c \ell \nu$ (b) Legend

Figure 5.12: Projection of the dipion mass in the signal region. Here, two prominent peaks are present, one originating from the decay $B^+ \to D^0 \ell \nu$ with $D^0 \to \pi^+\pi^-$ with the same final state as the signal around $M_{\pi^+\pi^-} = 1.87$ GeV, and another one from $B^+ \to D^0 \ell \nu$ with $D^0 \to K^-\pi^+$ where one of the pions has been misidentified as a kaon producing the reflection around $M_{\pi^+\pi^-} = 1.75$ GeV.

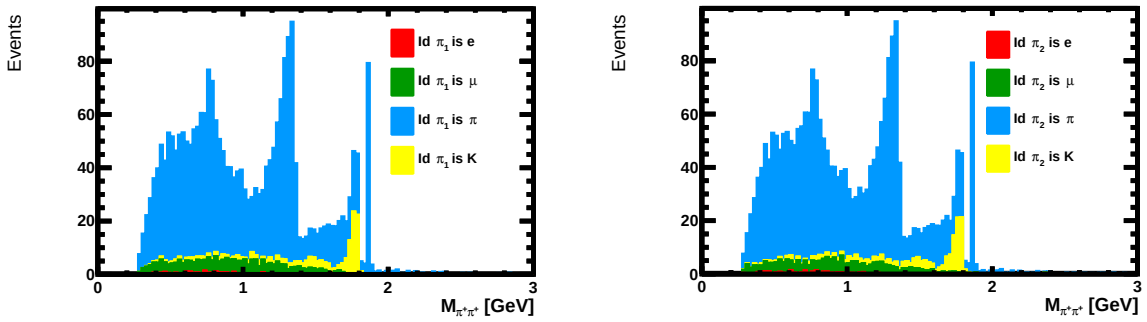
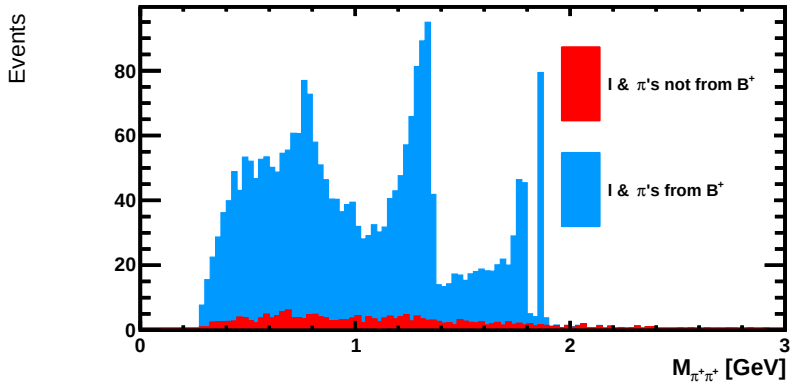


Figure 5.13: Invariant mass distribution after BDTG cut for the $B \to X_c \ell \nu$ background, where the charged pions have been identified as different particles: electron (red), muon (green), pion (blue) and kaon (yellow). The plot on the right side corresponds to the distributions of the positively charged pion, while the second to the negatively charged oney. It can be seen that is very unlikely to misreconstruct electrons as a pions; but it is possible to misidentify muons as a pions due to the similarities in their masses, however their contribution to the di-pion mass is tiny and almost constant. Finally, kaons misidentified as pions produce a peak around $M_{\pi^+\pi^-} = 1.75$ GeV, in this case one or both reconstructed pions are actually kaons.

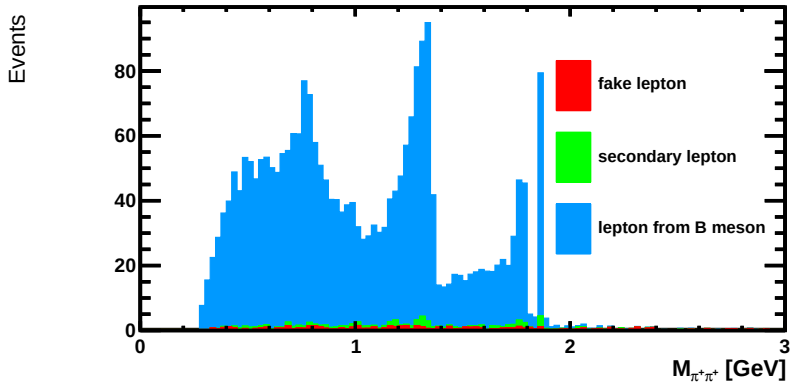
other $B^+ \to D^0 \ell^+ \nu_\ell$ channels serves as a probe to validate the signal selection, this however will be treated in the next chapter. The last component, rest of $B \to X_c \ell^+ \nu_\ell$, contains excited intermediate charm resonances, which are combinations of a D^0 or D^+ meson with additional pions or photons. They tend to populate the sideband region due to an increase in the multiplicity of final state particles.

5.2.5 Data/MC comparison of input variables in the sidebands

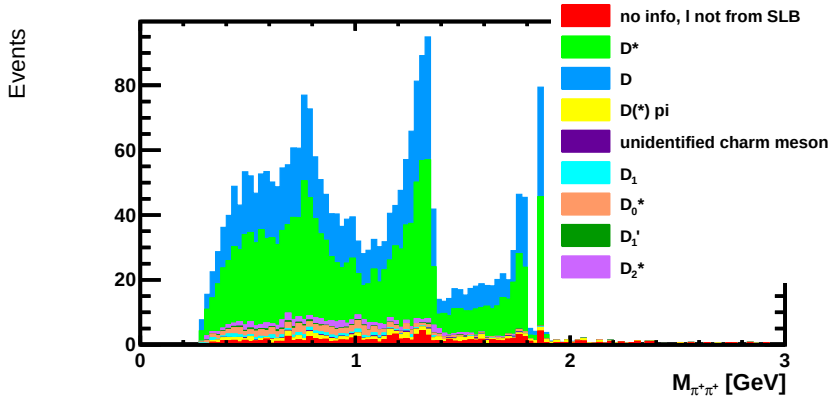
One way to verify that the MC describes the data is by comparing them in the sidebands, where little signal events are expected. This procedure also serves as a validation of the BDTG algorithm, in the sense that an alteration of the data and MC agreement after the cut on the BDTG output may be a sign of



(a) Leptons and pions in the same B decay chain

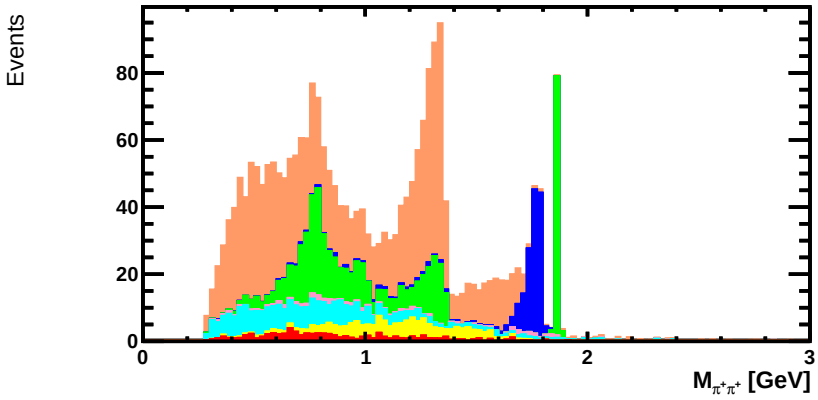


(b) Lepton origin



(c) Charm meson intermediate resonance

Figure 5.14: Contributions of different categories to the invariant mass distribution after BDTG cut. (a) Leptons and pions coming from the same B meson at some point in the decay chain dominates the di-pion mass, however cases where one of the particle may originate from different B mesons is also possible, that means that one of the three charge particles may come from the B_{tag}. (b) After the selection most of the leptons used in the analysis come directly from the B_{sig}, contributions from fake leptons (other charged particles misidentified as leptons) or secondary leptons (leptons stemming from intermediate resonances) are negligible. (c) The dominant source of background are due to the $B^+ \to D^{*0}\ell^+\nu_\ell$, followed by $B^+ \to D^0\ell^+\nu_\ell$ and in a smaller scale by semileptonic decays with an orbitally excited charm meson resonance.



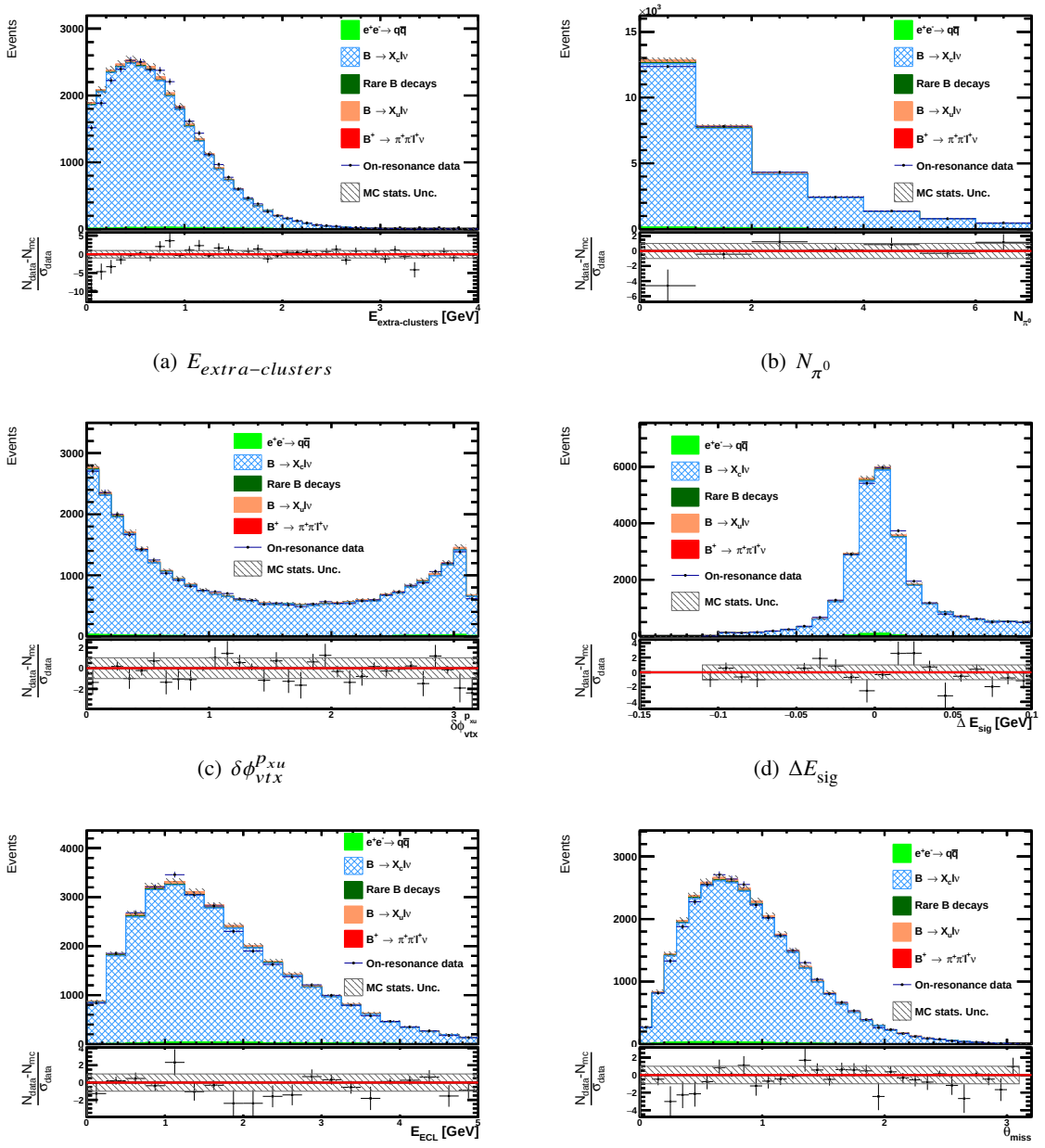
(a) $B \to X_c \ell \nu$

(b) Legend

Figure 5.15: Truth matching for the two final state pions after signal selection. The background comes mostly from decays where the two pions originate from the same mother but a third massive particle failed to be reconstructed. One peaking structure in the di-pion mass is due to two particles stemming from the same mother with at least one of them not correctly classified as pion. The majority of the peaking structures come from cases where the two pions match perfectly an intermediate resonance, but it may in turn be produced at some point in a decay chain.

overtraining. In this section the distributions of the input variables, in data and MC, before and after the cut on the BDTG response, are presented. In both cases, within the limited statistical size of the samples, the MC reproduce reasonably well the data and hence one can state that the procedure is free from overtraining.

Finally, the distributions of the selection variables before the application of the BDTG cut are shown in Fig. 5.16, and those after the cut on the BDTG output are presented in Fig. 5.17.



(a) $E_{extra-clusters}$

(b) N_{π^0}

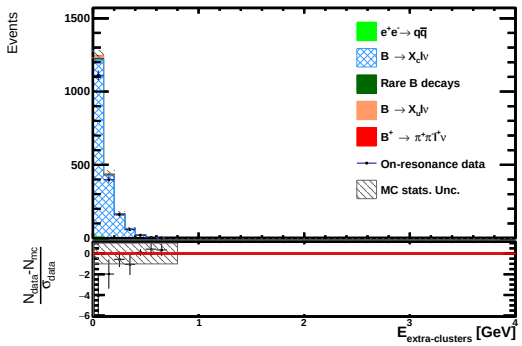
(c) $\delta\phi^{p_{xu}}_{vtx}$

(d) ΔE_{sig}

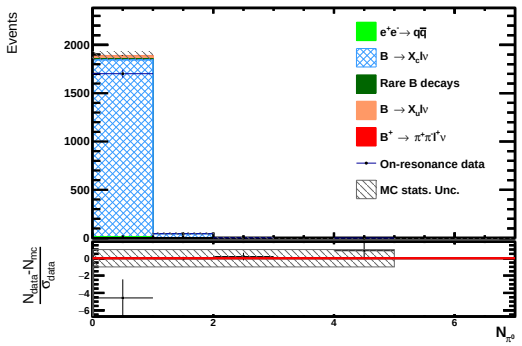
(e) θ_{miss}

(f) E_{ECL}

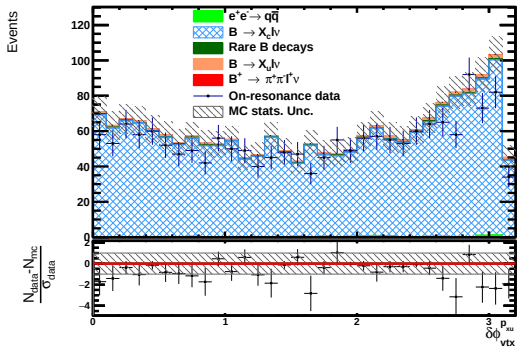
Figure 5.16: Distributions of the input variables after the pre-selection and before the cut on the BDTG response in the sidebands. The dominant background, $B \to X_c \ell \nu$, describes the data reasonably well in the sidebands as indicated in the plots underneath the figures where the majority of the pull = $\frac{N_{\text{data}} - N_{\text{MC}}}{\sigma_{\text{data}}}$ values, (N_i the bin content for either data or MC and σ_{data} the statistical error in data), lie in the range from -2 to 2.



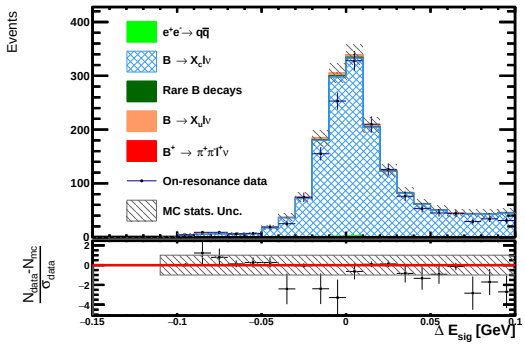
(a) $E_{extra-clusters}$



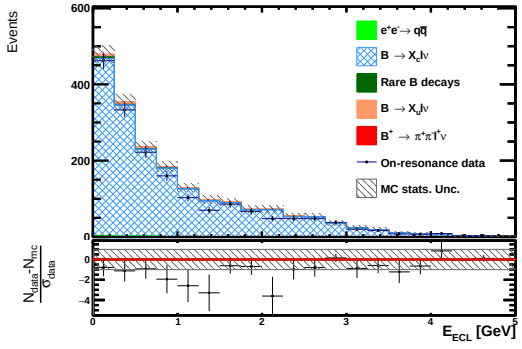
(b) N_{π^0}



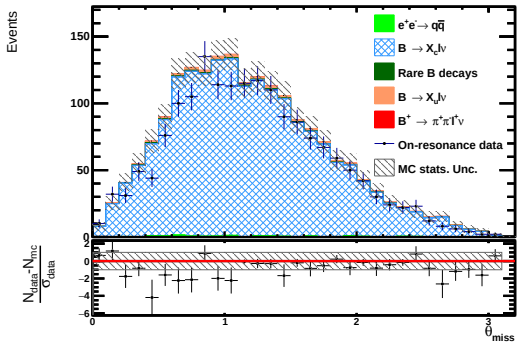
(c) $\delta\phi^{p_{xu}}_{vtx}$



(d) ΔE_{sig}



(e) θ_{miss}



(f) E_{ECL}

Figure 5.17: Distributions of the input variables after the cut on the BDTG output in the sidebands. The dominant background, $B \to X_c \ell \nu$, describes the data reasonably well in the sidebands as indicated in the plots underneath the figures where the majority of the pull = $\frac{N_{\text{data}}-N_{\text{MC}}}{\sigma_{\text{data}}}$ values, (N_i the bin content for either data or MC and σ_{data} the statistical error in data), lie in the range from -2 to 2.

CHAPTER 6

SIGNAL EXTRACTION

Up to this point, the selection and optimization have been done based solely on the MC samples. Section 6.1 discusses the use of the signal as well as other alternatives to simulate the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay to determine the signal reconstruction efficiencies and study model dependence. This section also presents a procedure to extract the event yields for the signal and background processes in bins of phase space variables, as well as the determination for an appropriate bin size, to perform the measurement. Section 6.2 describes the fitting procedure implemented in this thesis, which has been approved by the Belle Collaboration to unblind the method and carry out the measurement on data. Finally, section 6.3 outlines three methods to validate the fit procedure.

6.1 Study of the signal reconstruction efficiency and model dependence

An efficiency correction is necessary in order to compare physics spectra of kinematic variables, such as the di-pion mass $M_{\pi^+\pi^-}$ or the di-lepton mass squared (also known as momentum transfer squared) q^2, to theoretical predictions or even with other experimental results. However, such a correction may lead to a different distribution of the physical variables, showing features according to the model used to generate the signal instead of the underlying structure. In this case, one speaks of model dependence. As the decay rate for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay is expressed in terms of form factors that are calculated from theory, which in turn are functions of $M_{\pi^+\pi^-}$ and q^2. These variables can be used to determine a model dependence in the analysis. To do so, the signal reconstruction efficiency, which is defined as the ratio between the reconstructed signal after the complete selection with the MC corrections and the generated signal, is examined as a function of either q^2 or $M_{\pi^+\pi^-}$. A significant deviation from a flat efficiency would indicate a model dependence, implying as well that the analysis does not have access to the complete phase space of the decay. In practice, a perfectly flat efficiency is difficult to attain, since the selection requirements to reduce the background also affect the signal, which necessitates the use of low correlated variables in the BDTG training. As a consequence, the selection requirements should guarantee an efficiency distribution as flat as possible to avoid model dependence.

Another approach to test the impact of model dependence in the measurement is by comparing the efficiency distributions using different models to simulate the same final state and confirm that they provide a similar estimation of the efficiency. As the signal was produced with the PHSP generator in this analysis, given the lack of inclusive generators that simulate this decay, the alternative was to

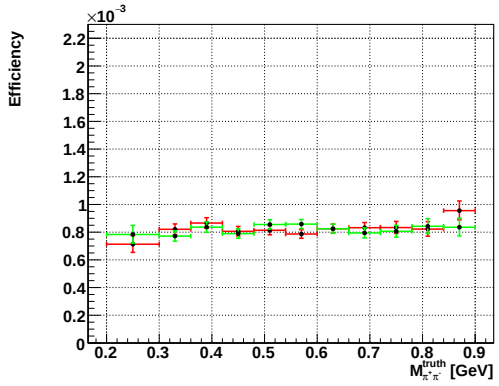
simulate semileptonic B decays with an intermediate resonance reconstructed in a $\pi^+\pi^-$ final state. Nonetheless, resonances do not span the whole range in the phase space as an inclusive simulation since they have a finite width and hence their coverage in the dipion mass is limited to the interval $[M_R - 2\Gamma_R, M_R + 2\Gamma_R]$, with M_R the nominal mass of the resonance and Γ_R its decay width. The range of momentum transfer squared q^2 varies with the resonant state, as the maximum value depends on the mass of the resonance through $q^2_{\text{max}} = (M_B - M_R)^2$, where M_B is the mass of the B meson. To compare the signal reconstruction efficiency with other models, four resonant channels are produced with different generators such as PHSP, ISGW2 and SLPOLE of the EvtGen [83] package, which are based on different theoretical calculations. These resonant channels are $B^+ \to f_0(500)\ell^+\nu_\ell$, $B^+ \to \rho^0\ell^+\nu_\ell$, $B^+ \to f_2(1270)\ell^+\nu_\ell$ and $B^+ \to \rho^0(1450)\ell^+\nu_\ell$. Form factors from the LCSR calculations for the $B^+ \to \rho^0\ell^+\nu_\ell$ and $B^+ \to f_2(1270)\ell^+\nu_\ell$ decays are taken from references [98]and [99], respectively, and enter the SLPOLE generator as input values. The comparison of the signal reconstruction efficiency of the inclusive model with resonant channels are displayed in Fig. 6.1 in bins of $M_{\pi^+\pi^-}$ and in Fig.6.2 in bins of q^2. In both cases, taking into consideration the statistical uncertainties, the estimated value of the efficiency is comparable in the different cases while its distribution is approximately flat implying that the selection provides a model independent measurement. Since the generator used in the production of the signal MC is not based on a QCD calculation, (form factors for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay), a systematic uncertainty due to the modeling of the signal is assigned, which will be discussed in section 7.1.1.

6.1.1 Strategy for a model independent measurement

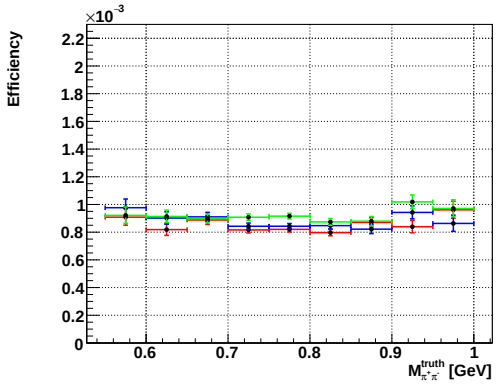
Ideally a model independent measurement of the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ channel would involve the five phase space variables that describe the decay, which can be accomplished through a five dimensional unbinned fit. However, the low statistics after the selection for this channel makes the implementation of this strategy difficult and the search for an alternative solution necessary. One possible approach for extracting the signal is to carry out a fit to M^2_{miss} in bins of phase space variables such as $M_{\pi^+\pi^-}$ or q^2. The use of the M^2_{miss} distribution is driven by its separation power between signal and background, an advantage of the implementation of the full reconstruction technique. The minimum bin size for the phase space variables is determined by their detector resolution, which is given as the difference between the reconstructed and true values of the variable. The true values depend only on generator-level information. At this level, for example, the momentum of the neutrino is fully determined, contrary to the case at the reconstructed level. The distributions of the resolution for the $M_{\pi^+\pi^-}$ and q^2 variables are presented in Fig. 6.3(a) and Fig.6.3(b), respectively. The $M_{\pi^+\pi^-}$ resolution is modeled as the sum of two Gaussian functions, whereas the q^2 resolution is model as the sum of a Crystal Ball [100–102] and two Gaussian functions, resulting in $\sigma_{M_{\pi^+\pi^-}} = (4.35 \pm 0.07) \times 10^{-3}$ GeV and $\sigma_{q^2} = (4.5 \pm 0.2) \times 10^{-2}$ GeV^2. These small resolutions are good enough to discard corrections such as unfolding. To avoid large statistical fluctuations and systematic uncertainties, the bin size of these variables must be wide enough to extract a significant number of events for the signal. Two configurations are explored in this thesis to perform a measurement of the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay, one in bins of $M_{\pi^+\pi^-}$ and another one in bins of $M_{\pi^+\pi^-}$ and q^2.

To determine the bin width in $M_{\pi^+\pi^-}$ and q^2 to be used in the measurement, a toy MC technique[1] is employed. For that aim, a fit to fake data in the M^2_{miss} distribution is performed, for which four histogram templates are used. These templates correspond to the signal, $B \to D^0\ell\nu$, the rest of $B \to X_c\ell\nu$ and a fixed background. The latter contains small amounts of events from the $B \to X_u\ell\nu$, $e^+e^- \to q\bar{q}$ and rare B decay processes, as well as events from $B^+ \to D^0\ell^+\nu_\ell$ with $D^0 \to \pi^+\pi^-$; they are then normalized to their MC expectation. A total of 1000 toy MC are generated to study the mean relative error on the signal

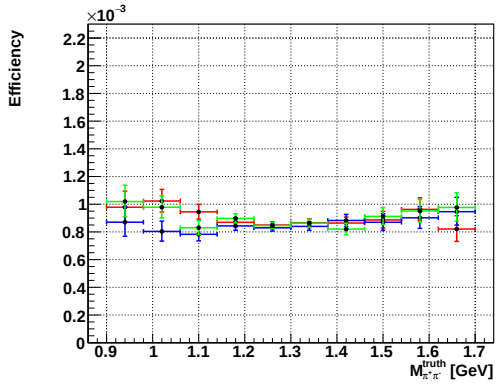
[1] A toy MC is generated as a set of fake data following a Poisson distribution based on the sum of the MC processes.



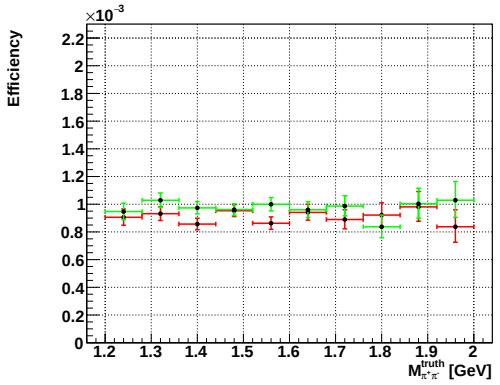
(a) $B^+ \to f_0(500)\ell^+\nu_\ell$



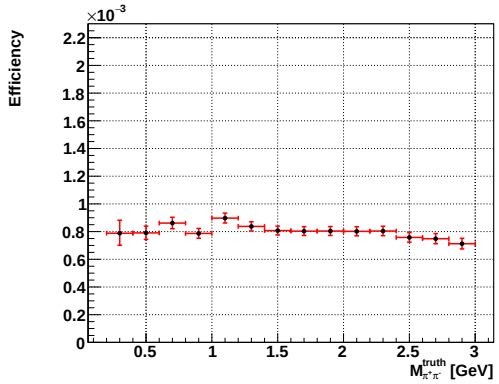
(b) $B^+ \to \rho^0\ell^+\nu_\ell$



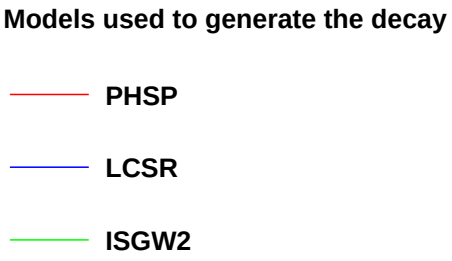
(c) $B^+ \to f_2(1270)\ell^+\nu_\ell$



(d) $B^+ \to \rho^0(1450)\ell^+\nu_\ell$

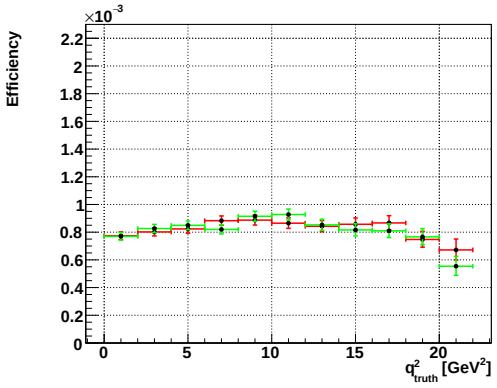


(e) PHSP $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$

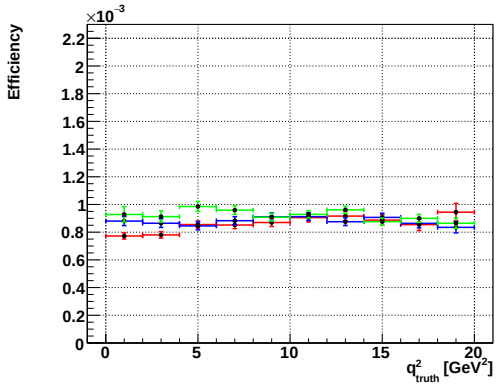


(f) Decay generators

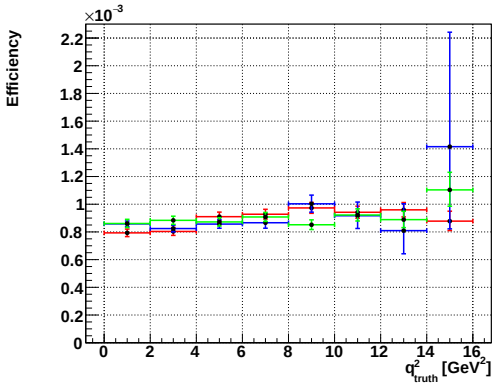
Figure 6.1: Signal reconstruction efficiency as a function of the di-pion mass for five channels simulating the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay using different generators. The distinct ranges in the di-pion mass for each channel is a consequence of the finite width of the resonances. Within the statistical uncertainties for each bin in $M_{\pi^+\pi^-}$, the efficiency is approximately flat and the value for the efficiency is similar for the inclusive and resonant decays indicating that the selection provides a model independent approach to the measurement. However, as the model is not based on a QCD calculation a systematic uncertainty due to signal modeling is assigned.



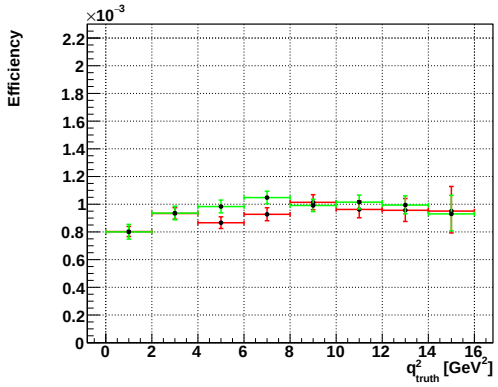
(a) $B^+ \to f_0(500)\ell^+\nu_\ell$



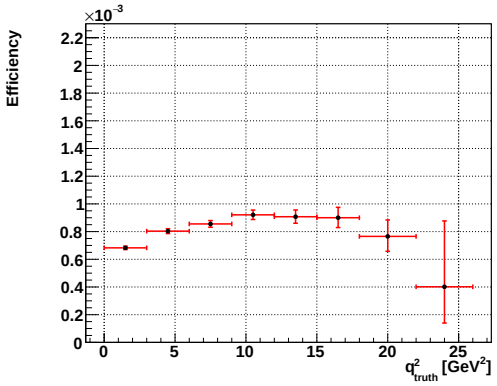
(b) $B^+ \to \rho^0\ell^+\nu_\ell$



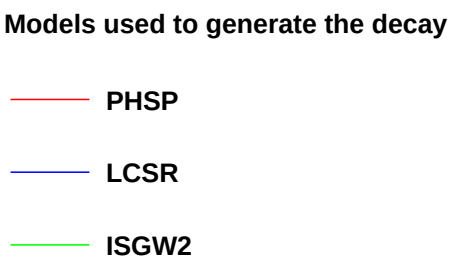
(c) $B^+ \to f_2(1270)\ell^+\nu_\ell$



(d) $B^+ \to \rho^0(1450)\ell^+\nu_\ell$

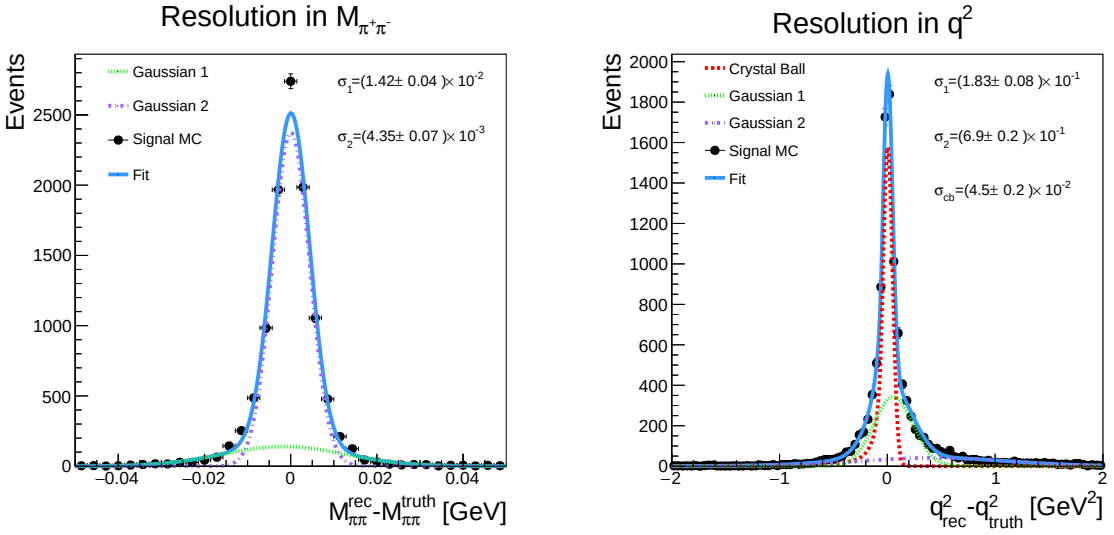


(e) PHSP $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$



(f) Decay generators

Figure 6.2: Signal reconstruction efficiency as a function of the momentum transfer squared for five channels simulating the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay using different generators. The upper bound for the q^2 interval varies for each resonant channel as $q^2_{\text{max}} = (M_B - M_R)^2$, with M_R and M_B the masses of the resonance and the B meson, respectively. Considering the statistical uncertainties for each bin in q^2, the efficiency is approximately flat and the estimated efficiency is similar for the different decay channels indicating indicating that the selection provides a model independent approach to the measurement. However, as the model is not based on a QCD calculation a systematic uncertainty due to signal modeling is assigned.



(a) (b)

Figure 6.3: Detector resolution of the phase space variables for: (a) $M_{\pi^+\pi^-}$ and (b) q^2, after signal selection for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ with the PHSP generator. The resolution is modeled as the sum of either two Gaussian functions in (a) or as a Crystal Ball function with two Gaussian functions in (b). The standard deviation gives and indication of the minimal bin width that can be used on these variables to extract the signal.

yield. The values of the signal yields and their errors are taken as the mean of a Gaussian fit to these distributions.

Choice for the bin size in $M_{\pi^+\pi^-}$

As the $\pi^+\pi^-$ mass threshold is approximately 0.3 GeV, different histogram templates for the M^2_{miss} distribution are created with di-pion mass bin sizes, varying from 40 to 90 MeV from this value. Although a small bin width is preferred to perform a model independent measurement, it is limited by the sample size. The goal is to look for a uniform bin width that allows to extract a significant amount of signal events, without having large statistical fluctuations from one bin to the next. The percent error on the signal yield as a function of the di-pion mass for different bin sizes is presented in Fig. 6.4, where it can be noticed that above $M_{\pi^+\pi^-} = 1.6$ GeV and below $M_{\pi^+\pi^-} = 0.5$ GeV, the error is dominated by statistical fluctuations. To reduce the error, the first two bins can be merged into one and a large bin can be formed for $M_{\pi^+\pi^-} > 1.6$ GeV. In the intermediate values of the di-pion mass the statistical errors obtained from bin sizes of 70 MeV, 80 MeV and 90 MeV are similar, but for the first two bins the error is greater for 70 MeV than for the other cases. A bin width of 80 MeV is thus selected to carry out the measurement. When performing the measurement on real data, bins above $M_{\pi^+\pi^-} = 1.42$ GeV yielded results consistent with zero signal events, and consequently all bins above this value were merged into a large bin. The final configuration contains 14 bins in the di-pion mass. To study migration effects bin by bin in the analysis, the percentage of events in the reconstructed di-pion mass that lie outside the same range in the true di-pion mass for each of the 14 bins of the fit configuration is calculated. The results are presented in Table 6.1 where the maximum value is 6.3%, and therefore a correction due to migration effects is not necessary.

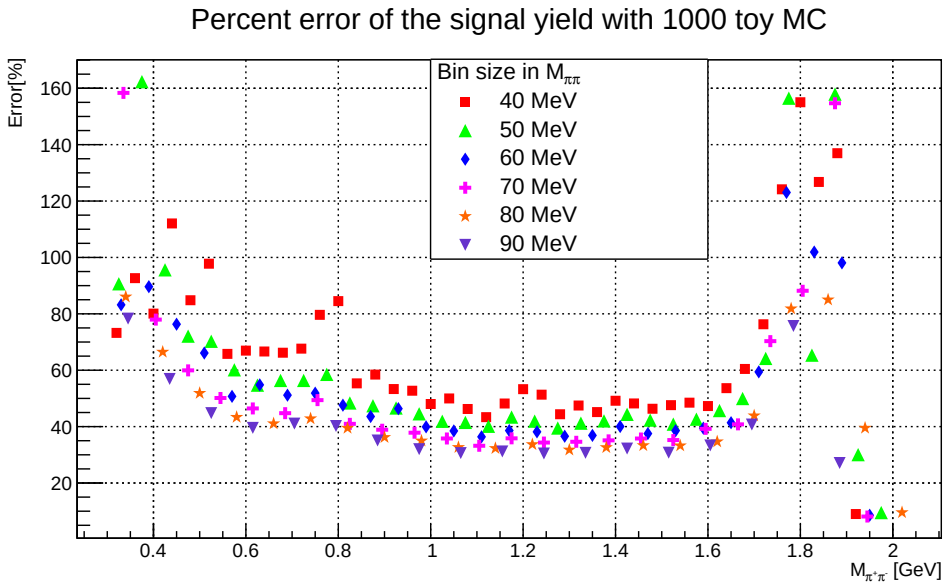


Figure 6.4: Percent error on the signal yield as a function of the di-pion mass for different bin sizes, after performing 1000 toy MC fits.

$M_{\pi\pi}^{\text{truth}}$ Bin [GeV]	Events [%]	$M_{\pi\pi}^{\text{truth}}$ Bin [GeV]	Events[%]
$M_{\pi\pi}^{\text{truth}} \leq 0.46$	1.4	$0.94 < M_{\pi\pi}^{\text{truth}} \leq 1.02$	3.8
$0.46 < M_{\pi\pi}^{\text{truth}} \leq 0.54$	4.2	$1.02 < M_{\pi\pi}^{\text{truth}} \leq 1.10$	4.6
$0.54 < M_{\pi\pi}^{\text{truth}} \leq 0.62$	3.8	$1.10 < M_{\pi\pi}^{\text{truth}} \leq 1.18$	4.8
$0.62 < M_{\pi\pi}^{\text{truth}} \leq 0.70$	2.8	$1.18 < M_{\pi\pi}^{\text{truth}} \leq 1.26$	6.3
$0.70 < M_{\pi\pi}^{\text{truth}} \leq 0.78$	4.3	$1.26 < M_{\pi\pi}^{\text{truth}} \leq 1.34$	5.9
$0.78 < M_{\pi\pi}^{\text{truth}} \leq 0.86$	4.6	$1.34 < M_{\pi\pi}^{\text{truth}} \leq 1.42$	5.7
$0.86 < M_{\pi\pi}^{\text{truth}} \leq 0.94$	3.8	$1.42 < M_{\pi\pi}^{\text{truth}}$	0.1

Table 6.1: Percentage of reconstructed mass events outside the bin in the true hadronic mass.

Choice for a two dimensional binning in the ($M_{\pi^+\pi^-}$,q^2) space.

The maximum q^2 value decreases with an increment in the di-pion mass value, as can be seen in Fig. 6.5. This fact must be taken into account when deciding on the length of the interval for q^2 and $M_{\pi^+\pi^-}$. The range of q^2 extends to values above 20 GeV^2 for low di-pion masses, but does not surpass 12 GeV^2 for high di-pion masses. To search for suitable binning sizes on these variables, different configurations are tested by combining different bin widths in q^2 and $M_{\pi^+\pi^-}$. For the di-pion mass three configurations are evaluated: seven 200 MeV width bins, six 250 MeV width bins and five 300 MeV width bins. On the other hand, five different configurations for q^2 are tried out, with bin widths of 2 GeV^2, 3 GeV^2, 4 GeV^2, 5 GeV^2 and 6 GeV^2. The last bin contains events above either $q^2 = 10$ GeV^2 or $q^2 = 12$ GeV^2. The percent error on the signal yield for the discussed combinations of bins in the (q^2, $M_{\pi^+\pi^-}$) arrangement is illustrated in Fig. 6.6. Note that the larger the bin number in a configuration, the larger the statistical fluctuation, especially the cases with a bin widths in q^2 of 2 GeV^2 and 3 GeV^2. To further reduce the

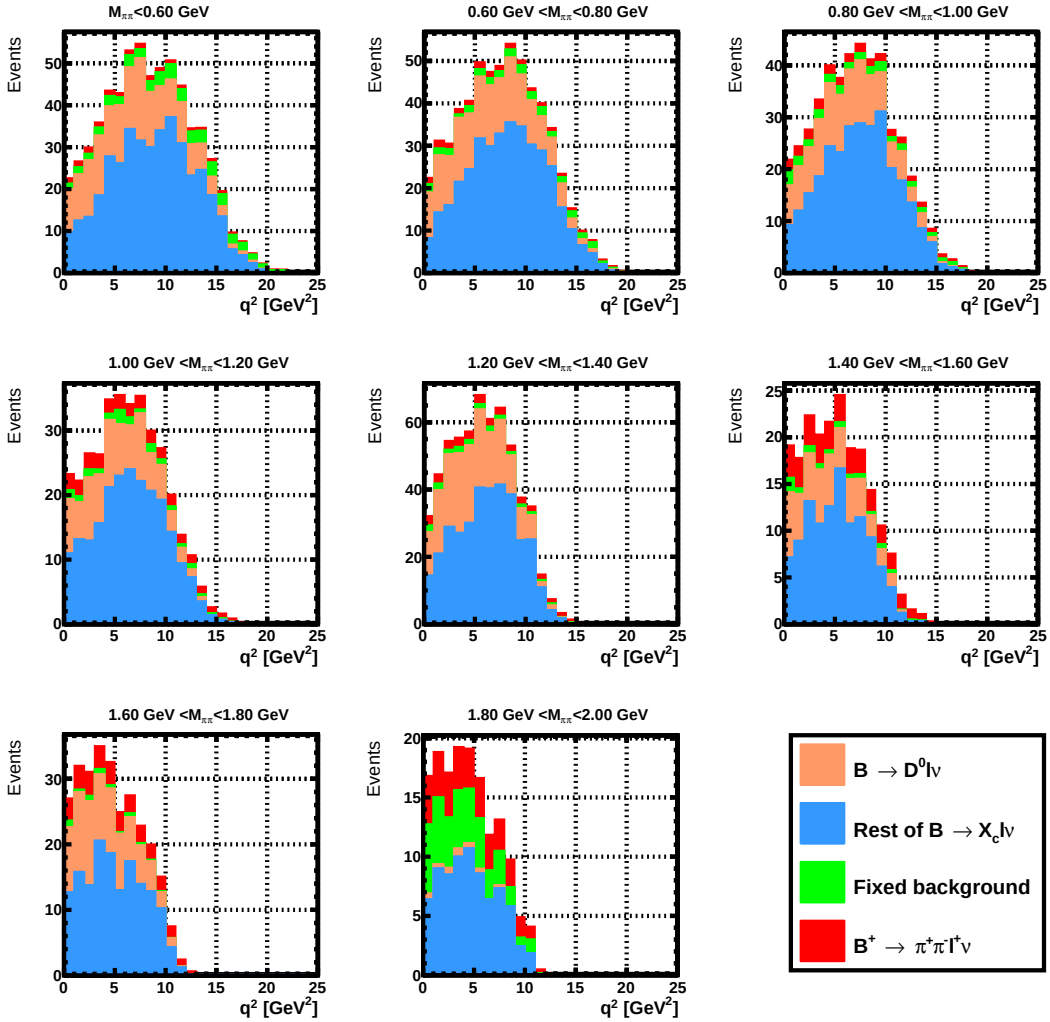
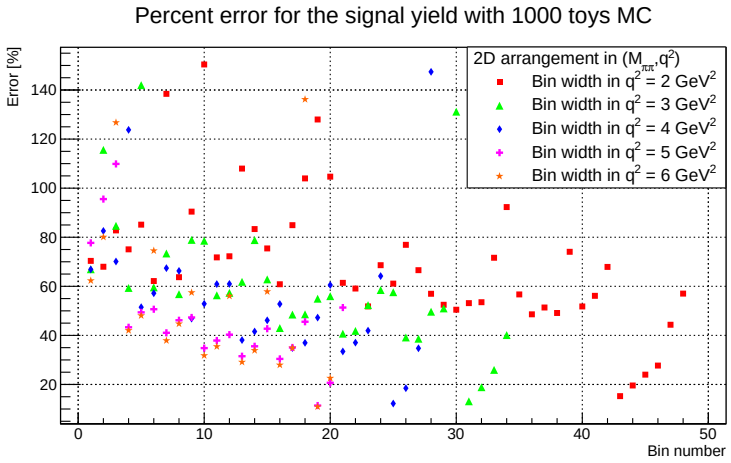
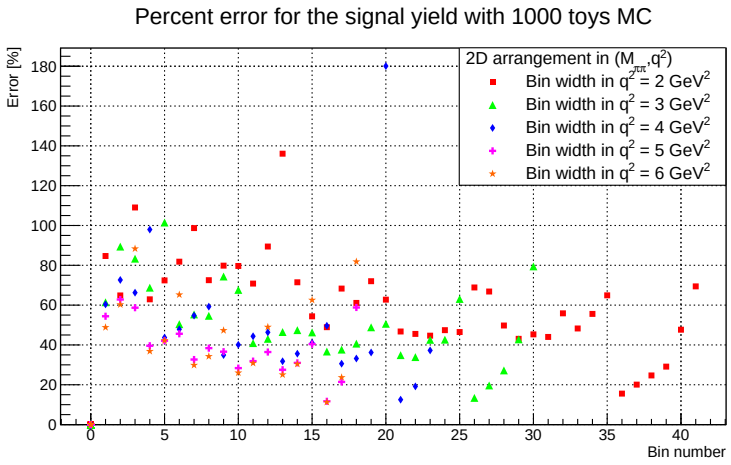


Figure 6.5: Projection of the q^2 distribution in different regions of the invariant mass $M_{\pi^+\pi^-}$. Note that the length of the interval q^2 decreases as the $M_{\pi^+\pi^-}$ increases.

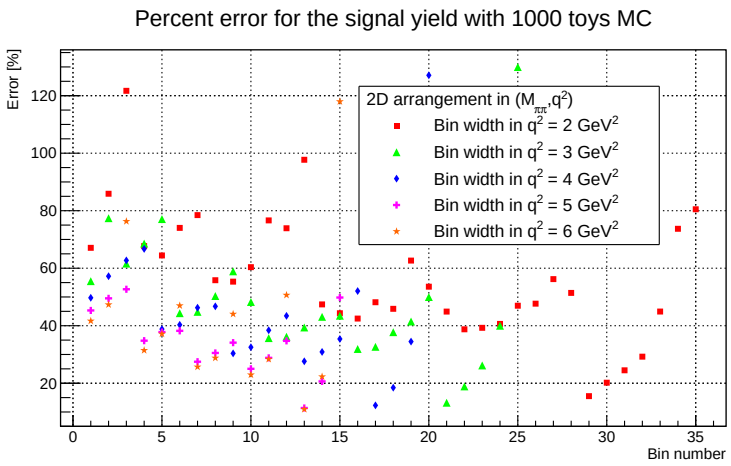
statistical error on the measurement, the last two bins in q^2 are merged. The final configuration consists of five bins in $M_{\pi^+\pi^-}$ and three bins in q^2 as presented in table 6.2 for a total of 15 bins.



(a) Bin width in $M_{\pi^+\pi^-}$ = 200 MeV



(b) Bin width in $M_{\pi^+\pi^-}$ = 250 MeV



(c) Bin width in $M_{\pi^+\pi^-}$ = 300 MeV

Figure 6.6: Percent error on the signal yield as a function of the bin number in the (q^2, $M_{\pi^+\pi^-}$) two dimensional space after 1000 toy MC fits. The code number for the bin number starts with the lower bin in $M_{\pi^+\pi^-}$ segmented in the q^2 intervals and continues this sequence for the next bin in $M_{\pi^+\pi^-}$.

6.2 Fit procedure and measurement in data

Having selected the appropriate bin size in phase space variables to extract the signal, the MC can now be used to estimate the contributions of each process in the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay from on-resonance data. This estimation is achieved by using a fit procedure, that in the present analysis is based on a binned extended maximum likelihood. This section introduces the formalism of this method and presents the event yields from each process extracted from the fit.

6.2.1 Binned extended maximum likelihood formalism

In the standard maximum likelihood approach the parameters relevant to the shape of distributions are determined assuming that the absolute normalization corresponds to the total number of events [103]. Conversely, in the extended maximum likelihood method (EML), the absolute normalization is also a free parameter in the fit that has to be estimated. The derivation of the formulas and notation used in this section follow that of reference [104].

Consider that an experiment yielded n_{tot} observations of a random variable x distributed according to a PDF $f(x;\theta)$, which in turns depends on unknown parameters $\theta = (\theta_1, \cdots, \theta_m)$. The observations are arranged in a histogram with N_{bins} containing a certain number of entries $n = (n_1, \cdots, n_N)$, with expected values $\nu = (\nu_1, \cdots, \nu_N)$ given by

$$\nu_i(\theta) = n_{\text{tot}} \int_{x_i^{\text{min}}}^{x_i^{\text{max}}} f(x;\theta)dx, \tag{6.1}$$

where x_i^{min} (x_i^{max}) is the lower (upper) edge of the bin i. Equation 6.1 divided by n_{tot} gives the probability $p_i^{n_i}$ to be in bin i. As a result, the histogram can be interpreted as a single measurement of an N-dimensional vector whose joint PDF is given by a multinomial distribution[2],

$$\begin{aligned} f_{\text{joint}}(n;\nu) &= \frac{n_{\text{tot}}!}{n_1!\cdots n_N!} p_1^{n_1\cdots n_N} \\ &= \frac{n_{\text{tot}}!}{n_1!\cdots n_N!} \left(\frac{\nu_1}{n_{\text{tot}}}\right)^{n_1} \cdots \left(\frac{\nu_N}{n_{\text{tot}}}\right)^{n_N}. \end{aligned} \tag{6.2}$$

In the case that n_{tot} is treated as a random variable distributed according a Poisson probability with mean ν_{tot}, the joint PDF can be expressed as the product of the Poisson distribution of n_{tot} and the multinomial

[2] It describes the probability that n_1 entries end up in the first bin, n_2 entries in the second one and so on.

$M_{\pi\pi}$ bin [GeV]	q^2 Bin [GeV^2]
$M_{\pi\pi} \leq 0.6$	$q^2 \leq 4$
$0.6 < M_{\pi\pi} \leq 0.9$	$4 < q^2 \leq 8$
$0.9 < M_{\pi\pi} \leq 1.2$	$8 < q^2$
$1.2 < M_{\pi\pi} \leq 1.5$	
$1.5 < M_{\pi\pi}$	

Table 6.2: Two dimensional arrangement in $(q^2, M_{\pi^+\pi^-})$ space to carry out the measurement.

distributions of the N bins, hence

$$f_{\text{joint}}(n;\nu) = \frac{\nu_{\text{tot}}^{n_{\text{tot}}} e^{-\nu_{\text{tot}}}}{n_{\text{tot}}!} \frac{n_{\text{tot}}!}{n_1! \cdots n_N!} \left(\frac{\nu_1}{n_{\text{tot}}}\right)^{n_1} \cdots \left(\frac{\nu_N}{n_{\text{tot}}}\right)^{n_N}. \tag{6.3}$$

Defining $\nu_{\text{tot}} = \sum_{i=1}^{N} \nu_i$ and $n_{\text{tot}} = \sum_{i=1}^{N} n_i$, the joint PDF becomes

$$f_{\text{joint}}(n;\nu) = \prod_{i=1}^{N} \frac{\nu_i^{n_i} e^{-\nu_i}}{n_i!}, \tag{6.4}$$

with ν_i a function of the parameters θ and ν_{tot} given by

$$\nu_i(\nu_{\text{tot}}, \theta) = \nu_{\text{tot}} \int_{x_i^{\min}}^{x_i^{\max}} f(x;\theta) dx. \tag{6.5}$$

Equation 6.4 indicates that the problem of fluctuating the overall normalization according to a Poisson distribution is equivalent to multiplying the Poisson distributions, characterized with number of entries n_i and mean value ν_i, for each bin in the histogram.

The joint PDF in Eq. 6.4 is identified as the extended likelihood, and in order to determine the best estimator for the parameter θ, it is computationally convenient to maximize the logarithm of the extended likelihood with respect to the parameter θ [105], as it is easier dealing with sums than with products, thus the estimators $\hat{\theta}_1 \cdots \hat{\theta}_m$ are the solutions of the equations

$$\frac{\partial \log L(\nu_{\text{tot}}; \theta_1 \cdots \theta_m)}{\partial \theta_j} = -\frac{\partial \nu_{\text{tot}}}{\partial \theta_j} + \sum_{i=1}^{N} \frac{\partial n_i \log \nu_i(\nu_{\text{tot}}; \theta_1 \cdots \theta_m)}{\partial \theta_j} = 0, \ \forall j. \tag{6.6}$$

6.2.2 Fit parameters and inputs

The shapes and PDFs of the processes are taken from MC simulation as histogram templates, which consist of the $B^+ \to \pi^+\pi^-\ell^+\nu$ signal, the $B^+ \to D^0\ell^+\nu_\ell$ decays, the rest of $B \to X_c\ell^+\nu_\ell$ and the fixed background. The first three components are allowed to float in the fit. The rest of $B \to X_c\ell^+\nu_\ell$ decays contain B^+ and B^0 meson decays to charm intermediate resonances different from the D^0 meson. The fixed background includes tiny contributions from continuum, $B \to X_u\ell\nu$, rare B and $B^+ \to D^0(\pi^+\pi^-)\ell^+\nu_\ell$ decays (see table 5.9), since these processes are statistically limited they are fixed to their predicted number of events to avoid unrealistic values. The $B^+ \to D^0(\pi^+\pi^-)\ell^+\nu_\ell$ is only present in the di-pion mass region from 1.85 GeV to 1.88 GeV, where contributions from other processes are almost negligible. This component is fixed to the measured value in data. The details about the event yield of the $B^+ \to D^0(\pi^+\pi^-)\ell^+\nu_\ell$ decay will be discussed in section 8.1.1. The fit is carried out on the M^2_{miss} distribution to exploit the good discriminating power between the signal and background using a bin width of 0.2 GeV^2 in the range from -1 GeV^2 to 6 GeV^2. The event yields for each component are reported in two binning configurations, one in 14 bins of the $M_{\pi^+\pi^-}$ (see table 6.1) and another in 15 two dimensional bins in $(M_{\pi^+\pi^-}, q^2)$ as shown in table 6.2.

6.2.3 Fit results

The event yields for signal and background processes after performing the fit are presented in tables 6.3 and 6.4 for the one and two dimensional binning configuration, respectively. The signal is not uniformly

distributed across the di-pion mass, being dominant in the region from 0.62 GeV to 0.94 GeV, and statistically limited in the regions above 1.42 GeV and below 0.46 GeV. From the two dimensional configuration results, it can be observed that most of the signal events in the dominant region in the di-pion mass are concentrated in the high q^2 region ($q^2 > 8\text{GeV}^2$). The goodness of fit is verified following the chi-square calculation by Baker and Cousins [106], which is applicable to fits derived from a maximum likelihood method where the data obey the Poisson statistics. The chi-square is defined as:

$$\chi^2 = -2 \sum_i \left[-y_i + n_i + n_i \ln \left(\frac{y_i}{n_i} \right) \right], \tag{6.7}$$

where y_i and n_i are the result from the fit and the value of the data for a given bin i, respectively. The distributions for the M^2_{miss} in bins of $M_{\pi\pi}$ are shown in figures 6.7, 6.8 and 6.9. The corresponding figures in the two dimensional binning configuration are presented in the appendix A.2. From these figures, one can notice that underneath the signal, for the majority of the bins, there is little contamination from the background. The pull distribution below each M^2_{miss} distribution, calculated as the difference between the data and the fit curve divided by the data error in each bin, does not exceed 2σ, indicating a good consistency between the data and the fit result.

Bin	$M_{\pi\pi}$ bin [GeV]	Signal	$B^+ \to D^0 \ell\nu$	Rest of $B^+ \to X_c \ell\nu$	Fixed background	Data	χ^2/NDF	Prob.[%]
1	$M_{\pi\pi} \leq 0.46$	6.6±3.6	117.5±38.9	79.1±37.4	20.2	225.0	25.6/32	78.2
2	$0.46 < M_{\pi\pi} \leq 0.54$	9.8±3.9	66.6±28.0	80.5±27.9	17.1	179.0	29.7/32	58.3
3	$0.54 < M_{\pi\pi} \leq 0.62$	10.1±3.9	130.2±43.9	60.8±42.5	14.8	216.0	37.3/32	23.9
4	$0.62 < M_{\pi\pi} \leq 0.70$	23.3±5.8	61.9±37.8	123.5±38.7	9.3	220.0	27.9/32	67.2
5	$0.70 < M_{\pi\pi} \leq 0.78$	89.9±10.3	112.3±53.1	123.0±52.7	12.4	337.0	45.4/32	5.8
6	$0.78 < M_{\pi\pi} \leq 0.86$	50.9±7.7	17.9±38.4	133.2±41.3	12.3	214.0	29.6/32	58.8
7	$0.86 < M_{\pi\pi} \leq 0.94$	29.2±6.1	51.3±29.9	57.6±29.4	7.8	146.0	42.8/32	9.6
8	$0.94 < M_{\pi\pi} \leq 1.02$	10.2±3.8	20.5±30.7	81.9±32.5	6.1	119.0	15.0/32	99.5
9	$1.02 < M_{\pi\pi} \leq 1.10$	8.9±3.3	37.2±22.4	90.4±23.6	4.0	140.0	26.3/32	75.1
10	$1.10 < M_{\pi\pi} \leq 1.18$	5.7±2.7	41.8±19.8	107.4±21.2	2.9	158.0	40.0/32	15.6
11	$1.18 < M_{\pi\pi} \leq 1.26$	15.7±4.6	62.7±31.5	123.9±32.5	3.0	205.0	41.2/32	12.8
12	$1.26 < M_{\pi\pi} \leq 1.34$	11.8±3.8	118.5±47.1	102.9±47.1	3.1	236.0	29.8/32	57.9
13	$1.34 < M_{\pi\pi} \leq 1.42$	8.4±3.5	56.5±15.9	41.4±15.0	1.8	108.0	29.1/32	61.6
14	$1.42 < M_{\pi\pi}$	12.4±13.4	153.7±36.9	215.3±31.7	66.2	448.0	25.3/32	79.5

Table 6.3: Event yields for the signal and background processes in the $B^+ \to \pi^+\pi^-\ell^+\nu$ decay obtained from an extended binned maximum likelihood fit to the M_{miss^2} distribution in bins of $M_{\pi^+\pi^-}$. The χ^2 and number of degrees of freedom (NDF) as well as the probability of the fit are also provided. The χ^2 calculation is based on the Baker-Cousins method [106].

Bin	$M_{\pi\pi}$ bin [GeV]	q^2 bin [GeV^2]	Signal	$B^+ \to D^0 \ell \nu$	Rest of $B^+ \to X_c \ell \nu$	Fixed background	Data	χ^2/NDF	Prob.[%]
1	$M_{\pi\pi} \leq 0.6$	$q^2 \leq 4$	2.3±2.2	64.6±17.9	29.2±16.4	6.3	104	42.8/32	9.6
2	$M_{\pi\pi} \leq 0.6$	$4 < q^2 \leq 8$	6.2±3.3	126.8±18.9	0.1±137.7	9.3	145	28.1/32	66.2
3	$M_{\pi\pi} \leq 0.6$	$8 < q^2$	15.8±5.0	91.1±43.6	185.2±44.2	33.1	329	30.3/32	55.3
4	$0.6 < M_{\pi\pi} \leq 0.9$	$q^2 \leq 4$	28.9±5.9	119.4±33.9	16.7±30.7	9.8	175	28.3/32	65.5
5	$0.6 < M_{\pi\pi} \leq 0.9$	$4 < q^2 \leq 8$	33.9±6.5	125.8±42.1	92.3±41.0	9.9	262	22.4/32	89.7
6	$0.6 < M_{\pi\pi} \leq 0.9$	$8 < q^2$	116.7±11.9	8.1±257.0	309.9±63.4	22.5	457	37.7/32	22.4
7	$0.9 < M_{\pi\pi} \leq 1.2$	$q^2 \leq 4$	8.2±3.3	29.4±17.5	81.1±18.6	5.8	124	20.2/32	94.7
8	$0.9 < M_{\pi\pi} \leq 1.2$	$4 < q^2 \leq 8$	8.8±3.5	135.5±35.4	55.2±34.1	4.9	204	26.7/32	73.0
9	$0.9 < M_{\pi\pi} \leq 1.2$	$8 < q^2$	27.8±6.0	0.5±188.4	169.2±28.1	6.6	204	39.0/32	18.5
10	$1.2 < M_{\pi\pi} \leq 1.5$	$q^2 \leq 4$	11.2±3.9	85.0±24.3	57.4±23.7	4.1	158	35.9/32	29.2
11	$1.2 < M_{\pi\pi} \leq 1.5$	$4 < q^2 \leq 8$	9.4±3.8	105.4±27.9	122.0±27.9	2.5	239	25.6/32	78.2
12	$1.2 < M_{\pi\pi} \leq 1.5$	$8 < q^2$	13.1±4.1	40.7±19.6	91.9±20.6	2.4	148	27.1/32	71.2
13	$1.5 < M_{\pi\pi}$	$q^2 \leq 4$	8.5±9.4	72.1±17.3	63.1±13.2	36.2	180	25.4/32	79.1
14	$1.5 < M_{\pi\pi}$	$4 < q^2 \leq 8$	7.0±5.5	80.0±17.4	53.6±14.9	22.5	161	20.2/32	94.7
15	$1.5 < M_{\pi\pi}$	$8 < q^2$	1.4±4.0	6.6±11.5	45.6±11.6	5.3	61	9.4/9	39.7

Table 6.4: Event yields for the signal and background processes in the $B^+ \to \pi^+\pi^-\ell^+\nu$ decay obtained from an extended binned maximum likelihood fit to the M_{miss^2} distribution in bins of $M_{\pi^+\pi^-}$ and q^2. The χ^2 and number of degrees of freedom (NDF) as well as the probability of the fit are also provided. The χ^2 calculation is based on the Baker-Cousins method [106].

6.3 Fit validation

To study the stability and possible bias in the fit procedure, different tests are carried out and discussed in this section. They are intended to verify the reproducibility of the central values and their statistical errors obtained in the nominal fit to on-resonance data. In case a bias is found, the central values and their statistical errors are corrected and a systematic uncertainty is assigned due to the fit procedure.

6.3.1 Generation of fake date from one stream of Generic MC (GSim test)

The first test consists in generating ten fake data sets by combining one stream of generic MC with the signal MC, $B \to X_u \ell \nu$ and rare decay processes. All samples, except the signal, are scaled according to their luminosities. The signal is normalized according to the fit results. The fitting templates are formed with the remaining streams for the generic MC (nine streams for the generic $B \to X_c \ell \nu$ and five streams for the continuum MC), and the same histogram templates for the other processes as used in the formation of the fake data. Given that there are only six streams of continuum MC, four of them are reused to make up the ten fake data samples. The next step is to carry out a fit to the $M^2_{\rm miss}$ distribution for each of the ten fake data in bins of phase space variables, and hence obtain ten points for the fitted signal yields with their statistical errors to compare with the nominal results. A graphical representation of these results are presented in Fig. 6.10 and Fig. 6.11 for the one and two dimensional binning configurations, respectively. It can be noticed that for most bins the errors obtained from a fit to fake data reproduce the errors measured in on-resonance data. However, for bins with poor statistics, for example bin 14 in the one dimensional configuration and bins 13, 14, and 15 in the two dimensional configuration, the error exhibits statistical fluctuations and do not reproduce the measured errors. On the other hand, the central values are consistent with the measured signal yields within statistical uncertainties. With these ten points, one can calculate the mean and standard deviation of the pull to find out which bins show a tendency to give biased results, the standard deviation is used in this case as the error of the pull. The pull is defined as follows:

$$\text{Pull} = \frac{N_{\rm fit} - N_{\rm input}}{\sigma_{\rm fit}}, \tag{6.8}$$

where $N_{\rm fit}$ and $\sigma_{\rm fit}$ are the fitted signal yields and their statistical errors, respectively, and $N_{\rm input}$ is the expected value, in this case the measured signal yield on real data. The results of the pull study are arranged in Table 6.5, where no apparent bias is observed for the different bins. In some cases the average of the pull is quite different from zero suggesting a biased measurement, however the error on the pull is also large preventing a definitive statement on the bias in the measurement.

6.3.2 Linearity test

The next step is to produce 1000 toy MC using the same histogram templates used in the measurement to on-resonance data, which are scaled according to the nominal fit results. One way to validate the fit procedure is by evaluating the linearity of the event yield. This implies that an increment in the event yield used as input should result in a similar increment in the output from the fit. For the signal yield, the normalization is varied between 0.6 and 1.5 times the nominal values with an increment of 0.1, resulting in ten points that are subsequently used to fit a linear function. The signal yield distribution for each variation is fitted with a Gaussian function, and the signal output and its error are taken as the mean and the standard deviation of the Gaussian. In some bins, due to the lack of statistics, the Gaussian peaks at negative values and as the signal yield is a positive defined quantity, the negative values are ignored in the linearity test. The linear regression for each bin is presented in Fig. 6.12 and Fig. 6.13 for the one and

	1D binning		2D binning	
Bin	pull average	pull SD	pull average	pull SD
1	0.01	0.22	-0.01	0.24
2	0.01	0.18	-0.01	0.21
3	-0.01	0.20	0.00	0.18
4	0.00	0.16	-0.01	0.27
5	0.00	0.26	0.01	0.17
6	0.00	0.19	0.00	0.21
7	0.02	0.19	-0.03	0.34
8	0.01	0.22	-0.02	0.32
9	-0.02	0.28	0.00	0.17
10	-0.02	0.36	-0.01	0.35
11	-0.02	0.36	-0.01	0.31
12	0.00	0.28	0.00	0.40
13	-0.01	0.36	-0.72	2.55
14	-6.77	21.3	-0.42	1.60
15	–	–	-0.08	0.76

Table 6.5: Average and standard deviation of the pull for the ten points in gsim test for each bin in the two configurations. If the standard deviation is taken as the error of the pull, the pull is consistent with zero and thus there is not a conclusive determination of the bias from this approach.

two dimensional binning configurations, respectively. In the majority of the bins, the slope is consistently one indicating a stable and bias free fit. Nevertheless, for bins especially in regions with the highest $M_{\pi\pi}$ and q^2, the slope is far off one, but the sizable error does not allow to conclude if there is a bias in the measurement. The method to study the linearity does not work on the last bin in the two dimensional configuration due to the small data sample, resulting in a sizable error on the signal yield comparable to, or larger than, the central value.

6.3.3 Pulls from Toy MC

Given that for some bins the signal yields are small with sizable statistical uncertainties, a large number of pseudo-experiments (toy MC) are necessary to correctly determine a bias in the fit procedure. This bias is introduced by imperfections in the fit variables due, for example, to poorly modeled observables in the MC and correlations between them. To ascertain a bias in the fit procedure, the pull distribution is drawn from 20000 toy MC, which are executed using histogram templates normalized according to the nominal fit results. For an unbiased result, the pull distribution is described by a unit normal distribution, i.e. a Gaussian with mean zero and standard deviation one. Deviations of the mean from zero or the standard deviation from one indicate that the event yields from the fit are not correctly estimated. This could either be because the central values or the statistical errors are underestimated, overestimated, or both. The pull distributions for the signal component are presented in Figures A.4 and A.5 in Appendix A.2 for the one and two dimensional binning configuration, respectively. It can be inferred from these figures that the mean and width of the Gaussian are not consistent with zero and one, respectively, within their statistical uncertainties. Except for the bin 14 in the one dimensional configuration, and bins 1, 13, 14, and 15 in the two dimensional configuration, the pull distribution exhibits a Gaussian shape and consequently the

fit results are corrected according to [107]:

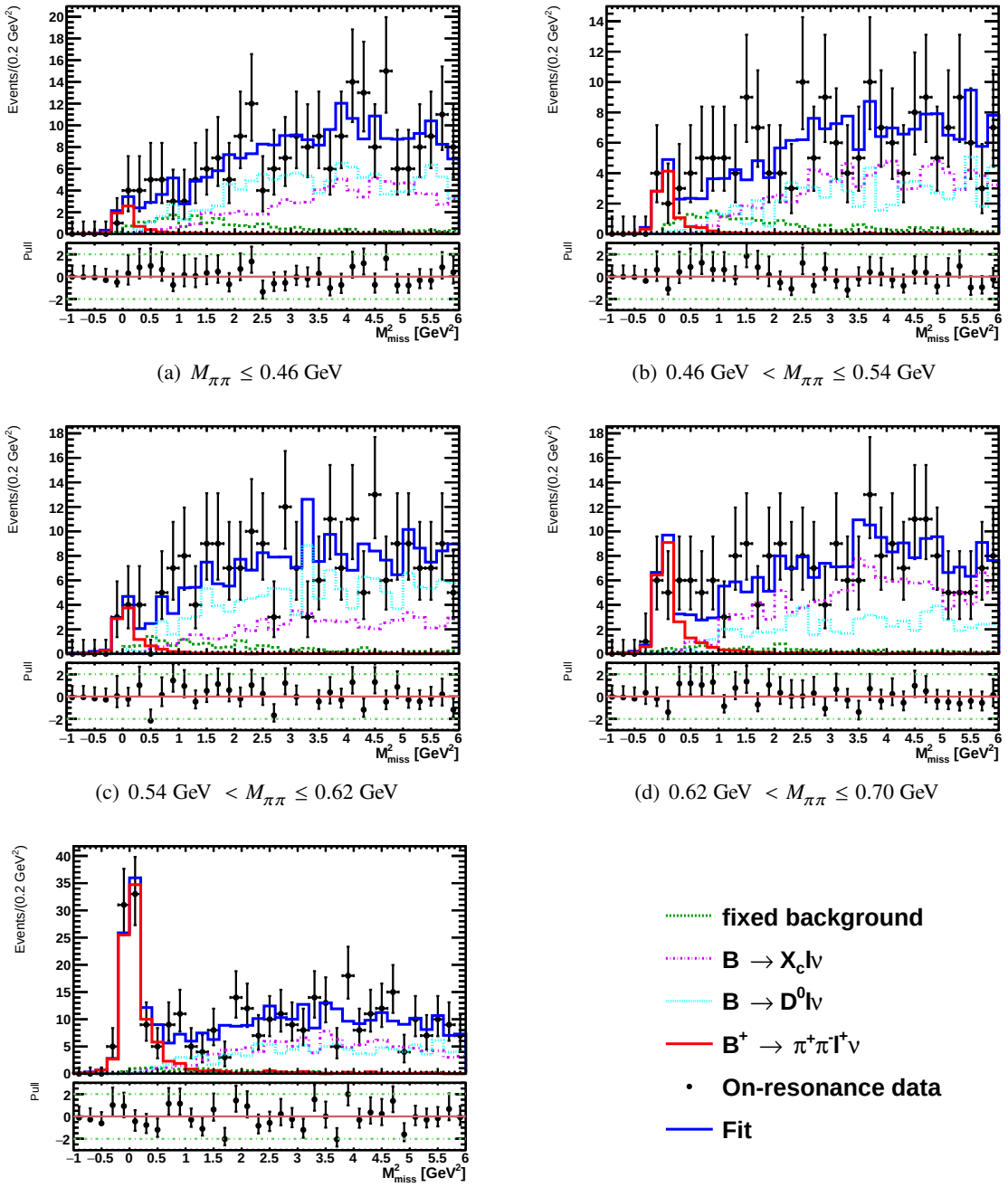
$$\begin{aligned} Y^{\text{cor}} &= Y^{\text{fit}} - \mu_{\text{pull}} \Delta Y^{\text{fit}}, \\ \Delta Y^{\text{cor}} &= \sigma_{\text{pull}} \Delta Y^{\text{fit}}, \end{aligned} \tag{6.9}$$

where Y^{fit} and ΔY^{fit} are the yields and errors from the nominal fit, μ_{pull} and σ_{pull} are the mean and width of the Gaussian fitted to the pull distribution, and Y^{cor} and ΔY^{cor} are the corrected values for the yield and errors. The bias is then given by:

$$\text{Bias} = \mu_{\text{pull}} \Delta Y^{\text{fit}}. \tag{6.10}$$

The mean and width of the pulls from the toy MC studies, as well as the bias and the corrected signal yields, are presented in Table 6.6 from the two binning configurations. After repeating the 20000 toy MC with the corrected values of the signal yields and their errors, it was confirmed that the fit procedure with the corrected values is free of bias. For completeness, the pull distributions for the remaining floating fit parameters are presented in figures A.6, A.7, A.8 and A.9.

The GSim and the linearity tests serve to validate the fit procedure, which means that the output of the fit behaves according to what is expected. However, a small bias is observed with the pull distribution from 20000 toy MC, and the signal event yields are corrected accordingly.



(a) $M_{\pi\pi} \leq 0.46$ GeV

(b) 0.46 GeV $< M_{\pi\pi} \leq 0.54$ GeV

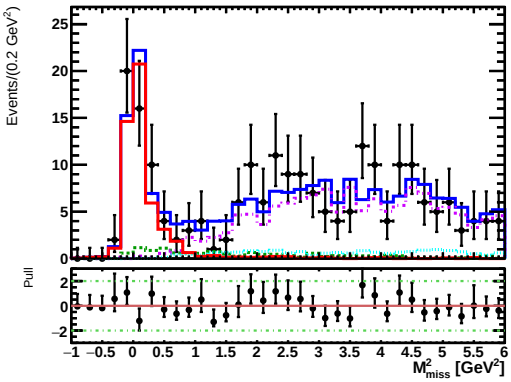
(c) 0.54 GeV $< M_{\pi\pi} \leq 0.62$ GeV

(d) 0.62 GeV $< M_{\pi\pi} \leq 0.70$ GeV

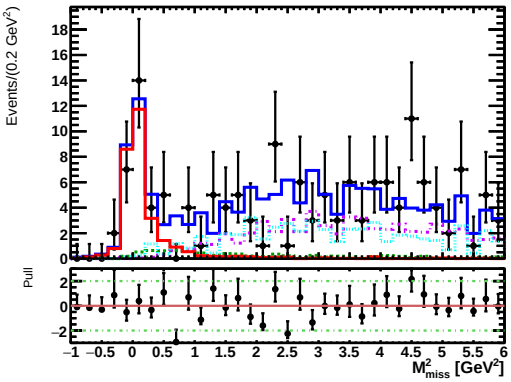
(e) 0.70 GeV $< M_{\pi\pi} \leq 0.78$ GeV

(f) Fit components

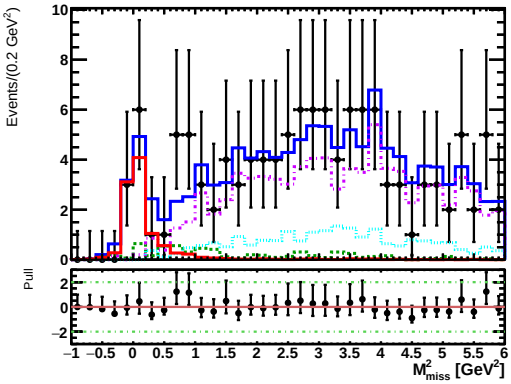
Figure 6.7: Distributions of the M^2_{miss} after a binned extended maximum likelihood fit in bins of the the di-pion mass in the range from $2M_\pi$ to 0.78 GeV. The fit components include the signal, $B^+ \to D^0 \ell^+ \nu_\ell$, other $B \to X_c \ell \nu$ and a fixed background. The latter contains contributions from continuum, $B \to X_u \ell \nu$ and rare B decays, that due to their poor statistics remains constant in the fit with a value equal to the prediction of the MC simulation. The pull values for each bin are presented underneath every plot, to show the accuracy of the fit with respect to the data. In most of the cases the pull values do not exceed 2σ, indicating a good consistency between the data and the fit result.



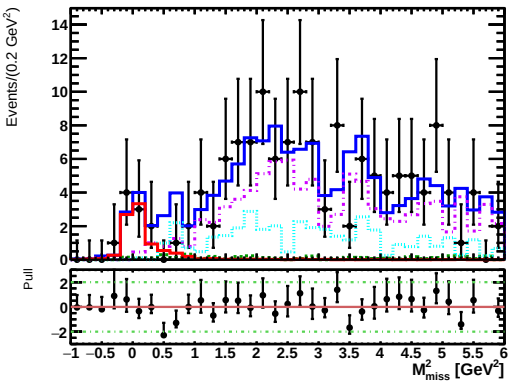
(a) 0.78 GeV $< M_{\pi\pi} \leq$ 0.86 GeV



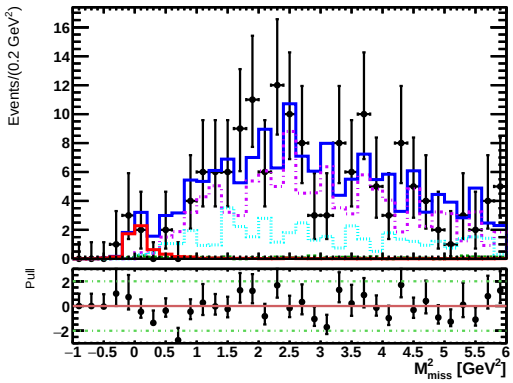
(b) 0.86 GeV $< M_{\pi\pi} \leq$ 0.94 GeV



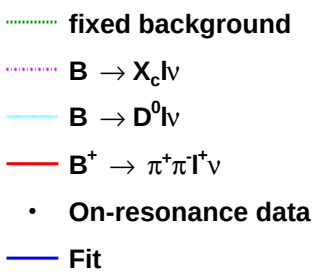
(c) 0.94 GeV $< M_{\pi\pi} \leq$ 1.02 GeV



(d) 1.02 GeV $< M_{\pi\pi} \leq$ 1.10 GeV

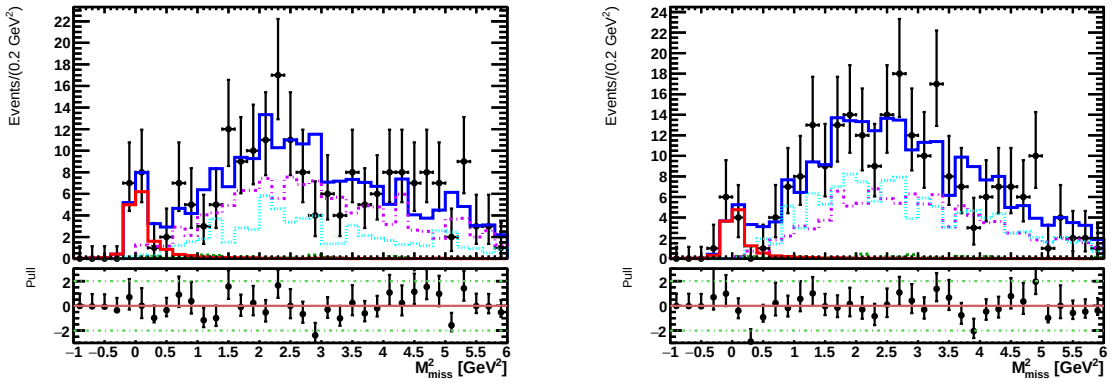


(e) 1.10 GeV $< M_{\pi\pi} \leq$ 1.18 GeV



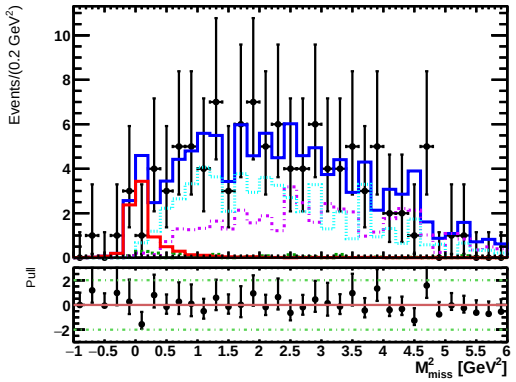
(f) Fit components

Figure 6.8: Distributions of the M^2_{miss} after a binned extended maximum likelihood fit in bins of the the di-pion mass in the range from 0.78 GeV to 1.18 GeV. The fit components include the signal, $B^+ \to D^0 \ell^+ \nu_\ell$, other $B \to X_c \ell \nu$ and a fixed background. The latter contains contributions from continuum, $B \to X_u \ell \nu$ and rare B decays, that due to their poor statistics remains constant in the fit with a value equal to the prediction of the MC simulation. The pull values for each bin are presented underneath every plot, to show the accuracy of the fit with respect to the data. In most of the cases the pull values do not exceed 2σ, indicating a good consistency between the data and the fit result.

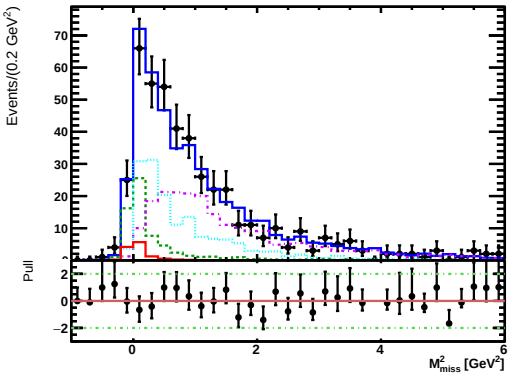


(a) 1.18 GeV $< M_{\pi\pi} \leq$ 1.26 GeV

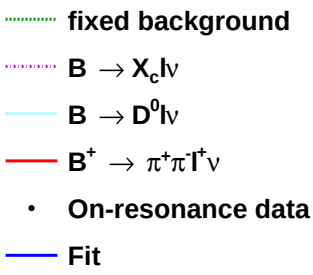
(b) 1.26 GeV $< M_{\pi\pi} \leq$ 1.34 GeV



(c) 1.34 GeV $< M_{\pi\pi} \leq$ 1.42 GeV

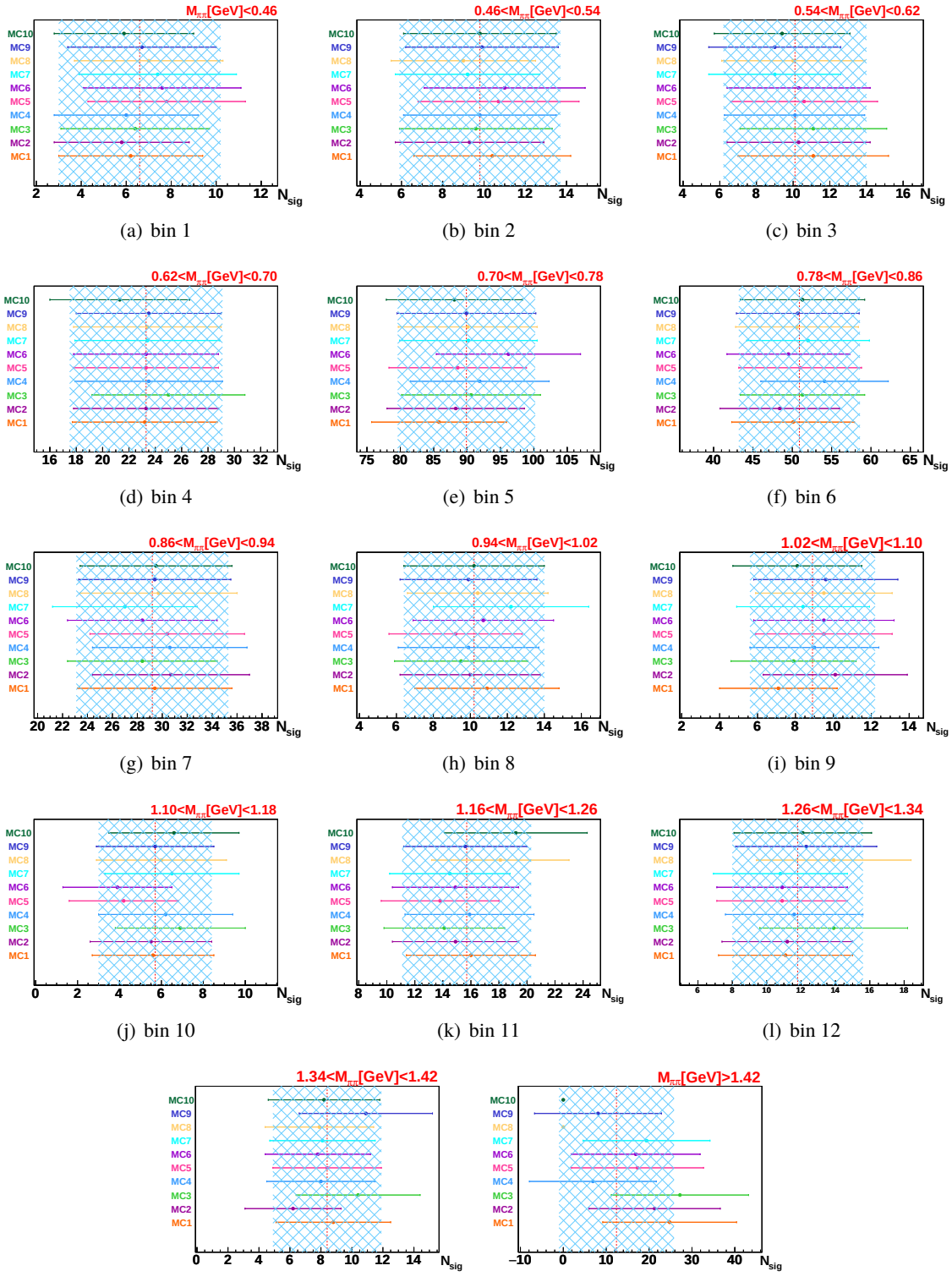


(d) $M_{\pi\pi} >$ 1.42 GeV



(e) Fit components

Figure 6.9: Distributions of the M^2_{miss} after a binned extended maximum likelihood fit in bins of the the di-pion mass above 1.18 GeV. The fit components include the signal, $B^+ \to D^0 \ell^+ \nu_\ell$, other $B \to X_c \ell \nu$ and a fixed background. The latter contains contributions from continuum, $B \to X_u \ell \nu$ and rare B decays, that due to their poor statistics remains constant in the fit with a value equal to the prediction of the MC simulation. The pull values for each bin are presented underneath every plot, to show the accuracy of the fit with respect to the data. In most of the cases the pull values do not exceed 2σ, indicating a good consistency between the data and the fit result.



(a) bin 1 (b) bin 2 (c) bin 3

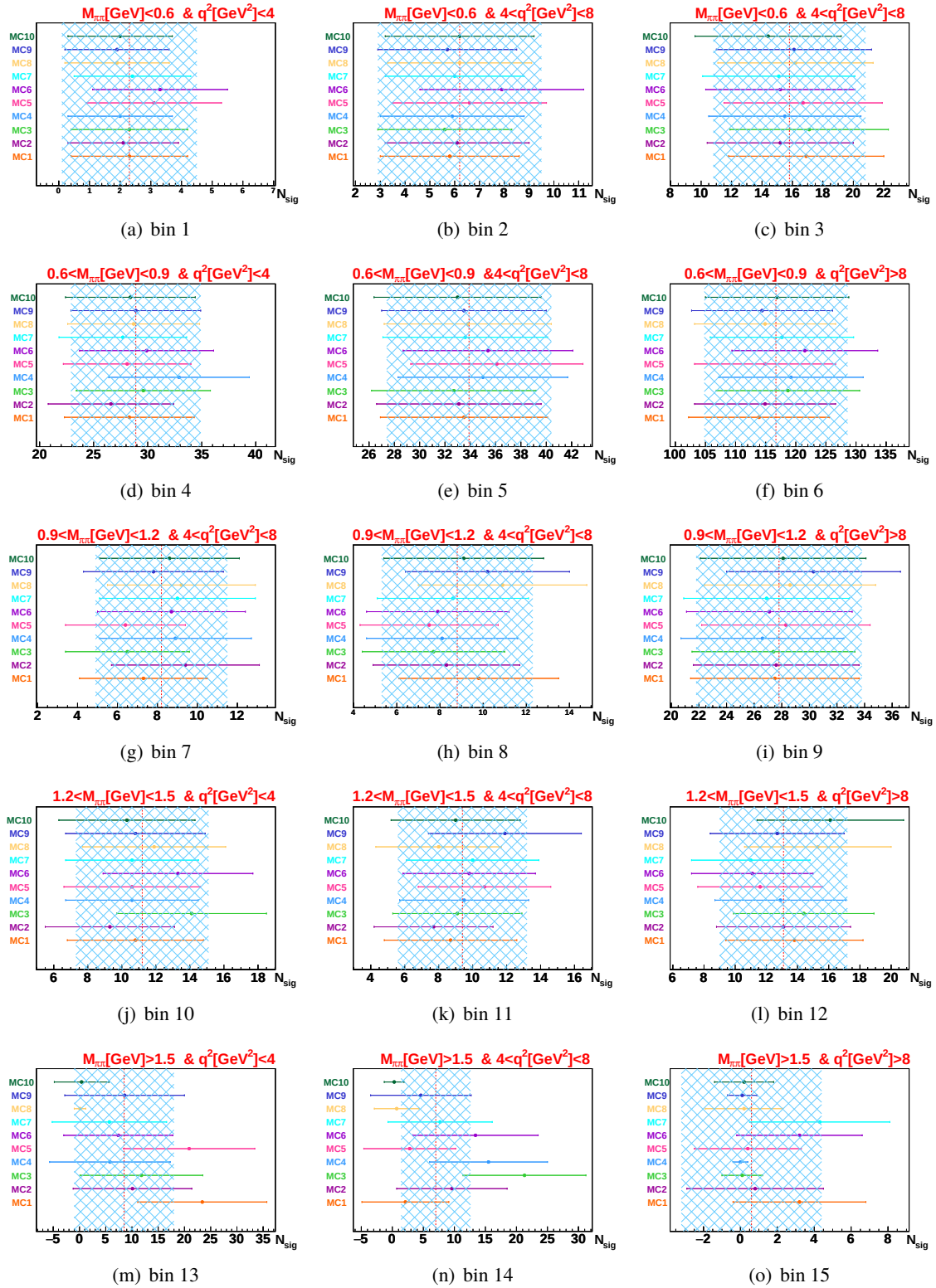
(d) bin 4 (e) bin 5 (f) bin 6

(g) bin 7 (h) bin 8 (i) bin 9

(j) bin 10 (k) bin 11 (l) bin 12

(m) bin 13 (n) bin 14

Figure 6.10: Comparison of the fitted signal yields from the ten fake data sets with the nominal result from real data (the central value is shown in the red dashed line and its statistical uncertainty is represented as the box). The fake data sets correspond to the sum of one stream of generic MC with the complete MC sample for the signal, $B \to X_u \ell \nu$ and rare samples. The histogram templates, used to perform the fit, are formed from the remaining streams of generic MC (nine for $B \to X_u \ell \nu$ and five for continuum MC) and the complete MC sample for the signal, $B \to X_u \ell \nu$ and rare samples. These results come from the one dimensional binning configuration.



(a) bin 1 (b) bin 2 (c) bin 3

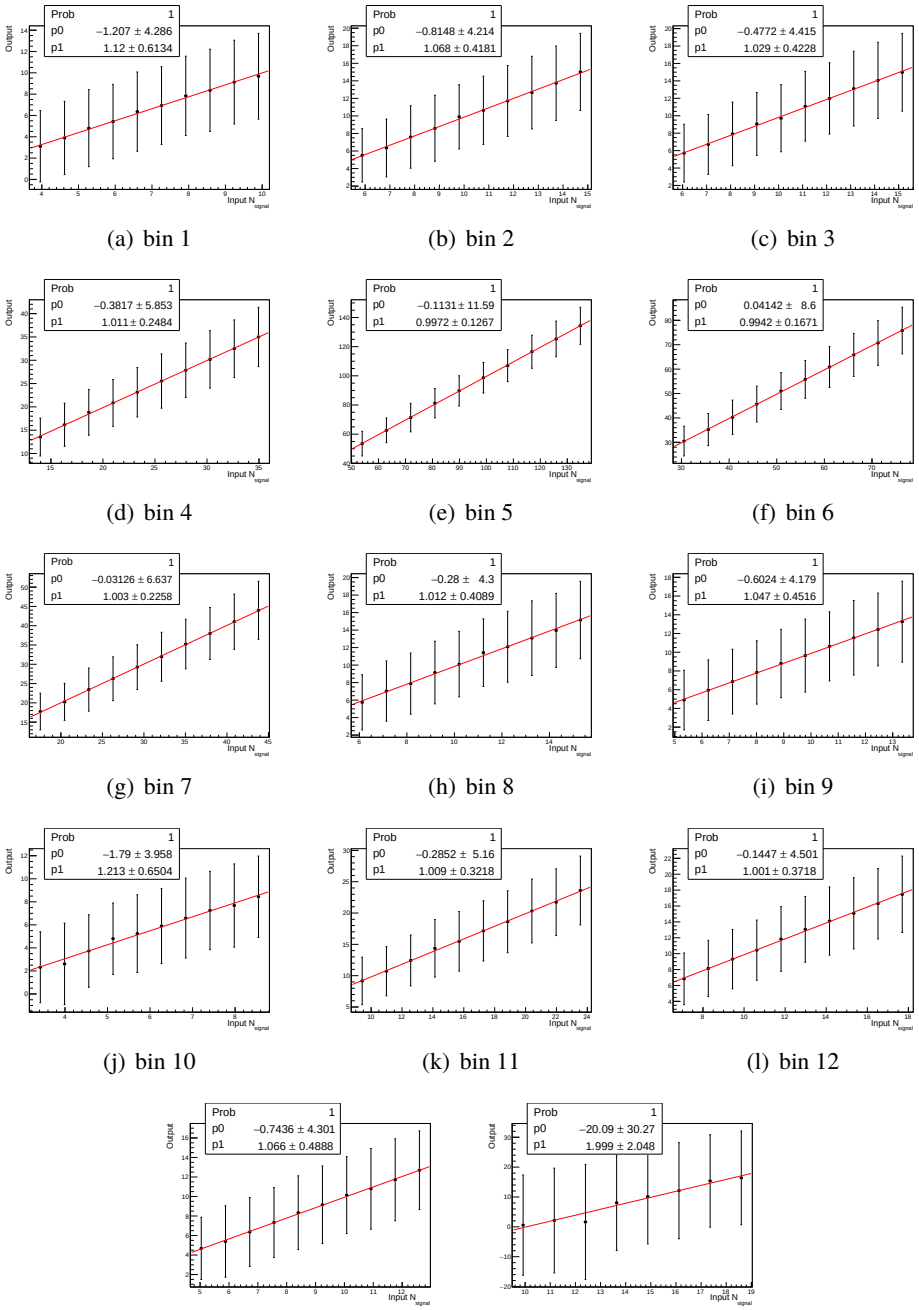
(d) bin 4 (e) bin 5 (f) bin 6

(g) bin 7 (h) bin 8 (i) bin 9

(j) bin 10 (k) bin 11 (l) bin 12

(m) bin 13 (n) bin 14 (o) bin 15

Figure 6.11: Comparison of the fitted signal yields from the ten fake data sets with the nominal result from real data (the central value is shown in the red dashed line and its statistical uncertainty is represented as the box). The fake data sets correspond to the sum of one stream of generic MC with the complete MC sample for the signal, $B \to X_u \ell \nu$ and rare samples. The histogram templates, used to perform the fit, are formed from the remaining streams of generic MC (nine for $B \to X_u \ell \nu$ and five for continuum MC) and the complete MC sample for the signal, $B \to X_u \ell \nu$ and rare samples. These results come from the two dimensional binning configuration.



(a) bin 1 (b) bin 2 (c) bin 3

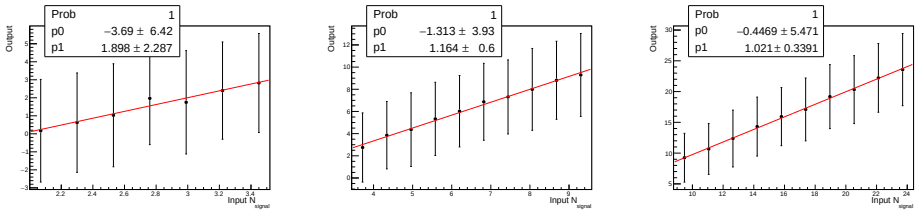
(d) bin 4 (e) bin 5 (f) bin 6

(g) bin 7 (h) bin 8 (i) bin 9

(j) bin 10 (k) bin 11 (l) bin 12

(m) bin 13 (n) bin 14

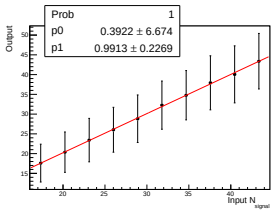
Figure 6.12: Linear regression for the output (fitted signal yield from Toy MC) as a function of the normalization of the signal yield from the one dimensional binning configuration.



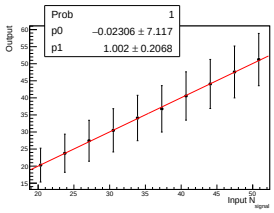
(a) bin 1

(b) bin 2

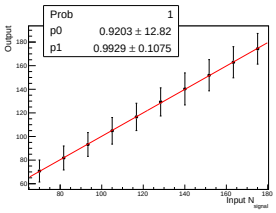
(c) bin 3



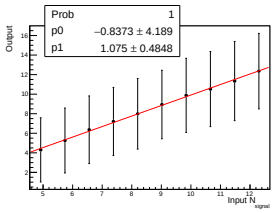
(d) bin 4



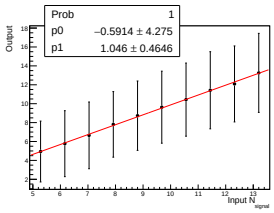
(e) bin 5



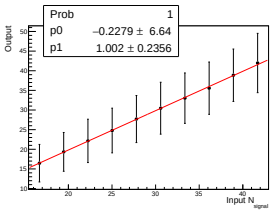
(f) bin 6



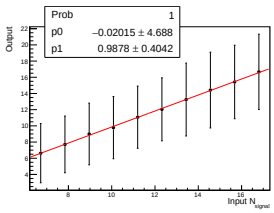
(g) bin 7



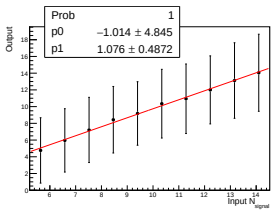
(h) bin 8



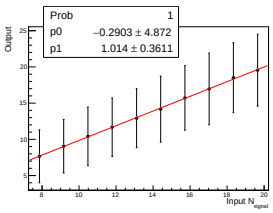
(i) bin 9



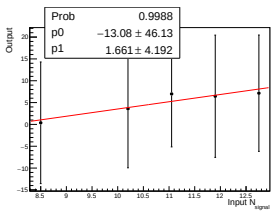
(j) bin 10



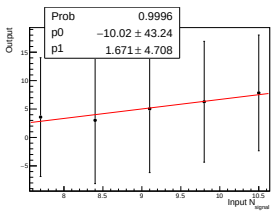
(k) bin 11



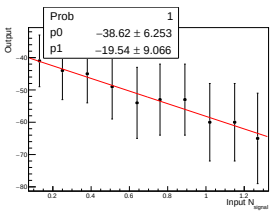
(l) bin 12



(m) bin 13



(n) bin 14



(o) bin 15

Figure 6.13: Linear regression for the output (fitted signal yield from Toy MC) as a function of the normalization of the signal yield from the two dimensional binning configuration. The signal yield is a positive defined quantity, however, when performing the toy MC analysis for different strengths of the signal yield for the last bin, the mean of the fitted Gaussian gave negative values and consequently a negative slope from the linearity test.

	1D binning					2D binning				
Bin	μ_{pull}	σ_{pull}	Bias [%]	Y^{cor}	ΔY^{cor}	μ_{pull}	σ_{pull}	Bias [%]	Y^{cor}	ΔY^{cor}
1	−0.096±0.009	0.916±0.005	5.24	6.9	3.3	−−	−−	–	2.3	2.2
2	−0.098±0.008	0.930±0.005	3.90	10.2	3.6	−0.110±0.009	0.921±0.005	5.85	6.6	3.0
3	−0.097±0.008	0.943±0.005	3.75	10.5	3.7	−0.108±0.008	0.956±0.005	3.42	16.3	4.8
4	−0.098±0.008	0.960±0.005	2.44	23.9	5.6	−0.066±0.007	0.960±0.005	1.35	29.3	5.7
5	−0.055±0.007	0.976±0.005	0.63	90.5	10.1	−0.067±0.008	0.972±0.005	1.28	34.3	6.3
6	−0.069±0.007	0.965±0.005	1.04	51.4	7.4	−0.055±0.007	0.985±0.005	0.56	117.4	11.7
7	−0.078±0.008	0.974±0.005	1.63	29.7	5.9	−0.106±0.008	0.933±0.005	4.27	8.5	3.1
8	−0.106±0.008	0.931±0.005	3.95	10.6	3.5	−0.108±0.008	0.984±0.005	4.30	9.2	3.4
9	−0.105±0.008	0.956±0.006	3.89	9.2	3.2	−0.110±0.008	0.972±0.005	2.37	28.5	5.8
10	−0.138±0.009	0.944±0.005	6.54	6.1	2.5	−0.114±0.008	0.954±0.005	3.97	11.6	3.7
11	−0.092±0.008	0.967±0.005	2.70	16.1	4.4	−0.108±0.009	0.912±0.005	4.37	9.8	3.5
12	−0.104±0.008	0.944±0.005	3.35	12.2	3.6	−0.1±0.008	0.962±0.005	3.13	13.5	3.9
13	−0.108±0.008	0.944±0.005	4.50	8.8	3.3	−−	−−	–	8.5	9.4
14	−−	−−	–	12.4	13.4	−−	−−	–	7.0	5.5
15	−−	−−	–	–	–	−−	−−	–	1.4	4.0

Table 6.6: Mean and width of the pulls from a 20000 toy MC for each bin in the two binning configurations. The bias of the fit procedure and the corrected signal yields are also presented. The highlighted values in red are the original values and have not been corrected since their pull distributions do not exhibit a Gaussian shape.

CHAPTER 7

SYSTEMATIC UNCERTAINTIES

In the measurement of physical parameters, the uncertainties associated with them are classified into two categories: statistical and systematic uncertainties. Statistical uncertainties appear due to random fluctuations originating from using data of finite sample size to carry out the measurement. On the other hand, systematic uncertainties are related to the nature of the experimental apparatus, the model used and the assumptions made to extract the physical parameters from data. In this chapter, an overview of the sources of systematic uncertainties in the measurement of the branching ratio of the $B^+ \to \pi^+\pi^-\ell^-\nu_\ell$ decay is given. Section 7.1 discusses the impact of the modeling of the signal and background processes in the measurement, section 7.2 presents the systematic uncertainties associated with the detector performance, section 7.3 describes the effect of the fitting procedure in the measurement and finally section 7.4 comments on other sources of systematic uncertainties.

In the majority of cases, the systematic uncertainties are computed by varying each parameter in the simulation by one standard deviation[1], up and down ($\pm 1\sigma$), and repeating the fit to the M^2_{miss}. The difference between the signal yields from the nominal fit and the ones with the varied parameter are taken as $\pm 1\sigma$ systematic error. Unless a different procedure is mentioned, this will be the standard procedure used to compute the systematic uncertainties throughout this chapter. The individual sources of systematic uncertainties are, to a good approximation, uncorrelated, so they can be added in quadrature to state the total relative contribution of these errors.

7.1 Modeling of Physical Processes

The use of a given model to simulate a physical process carries with it an implicit systematic uncertainty. They arise from the lack of precise knowledge of the parameters used in the simulation, for example the form factors that describe the model have errors associated with their theoretical calculations; when these form factors are varied according to their uncertainties, they cause a change in the shape of the distributions on which the measurement is performed. Another source of uncertainty is the relative contribution of each physical process, which is expressed as a branching ratio. A description of the procedure used to assess the contribution of the systematic uncertainties due to signal and background modeling is given in the following.

[1] These roughly corresponds to a 68% confidence level interval in Gaussian statistics.

7.1.1 Signal Modeling

To estimate the systematic uncertainties arising from the signal model, the signal reconstruction efficiency for each bin in $M_{\pi\pi}$, or $(M_{\pi\pi}, q^2)$, is calculated using the inclusive PHSP $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay. This value is then compared with the reconstruction efficiencies of semileptonic B decays with an intermediate resonance decaying to a $\pi^+\pi^-$ pair. As these channels simulate the same final state, the resulting efficiency should be similar. The resonant channels implemented for this comparison, as well as their valid $M_{\pi\pi}$ (q^2) range, were discussed in section 6.1. A weighted average between the signal and the resonant models and its error are calculated according to the following formulas:

$$x_{\text{avg}} = \frac{\sum_{i=1}^{N} \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}}, \quad \sigma_{\text{avg}} = \frac{1}{\sqrt{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}}}. \tag{7.1}$$

Here, the number of resonant channels vary from one bin to another depending on whether the range in $M_{\pi\pi}$ or q^2 is valid for that resonance. The percent difference between the signal reconstruction efficiency and the average efficiency is quoted as the systematic uncertainty due to the signal modeling.

7.1.2 Rare Decays

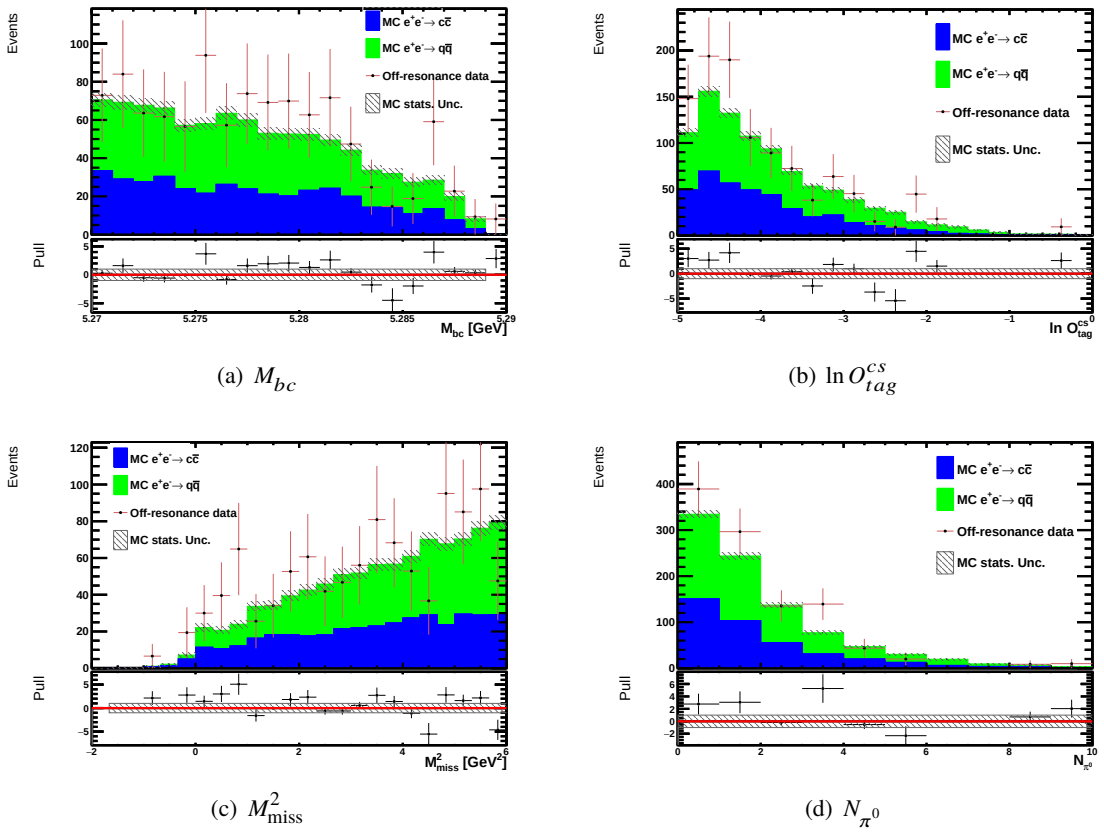
The amount of rare decay events in the final selection for each bin of phase space variables ($M_{\pi\pi}$ or q^2) is insignificant. At first, to evaluate their effect on the total systematic uncertainty, the difference between the nominal fit and a fit excluding the rare sample is taken as the error. However, this renders large fluctuations from bin to bin, especially for high di-pion mass regions. To reduce the statistical fluctuations effects on the estimation of this systematic uncertainty, 1000 toy MC are executed with two scenarios, one where the histograms templates for the fixed component contains the rare sample and another where it is absent. The average percent difference between these two scenarios is regarded as the uncertainty.

7.1.3 Continuum Description

After the complete selection, the MC expectation for continuum events is almost negligible. In order to assign an uncertainty on the modeling of this background, it is necessary to check the agreement with off-resonance data. Nevertheless, due to the limited sample size of the off-resonance data, it is not possible to do such a comparison after the signal selection. An alternative is to relax the selection at an earlier stage, e.g. the pre-selection stage, which is the step previous to applying the BDTG cut. Since the off-resonance data was taken just below the mass threshold used to produce the $B\bar{B}$ pairs, it has to be corrected by a scale factor in order to be compared with the continuum MC. This scale factor is defined as $\alpha = \frac{\mathcal{L}_{\text{on}}}{\mathcal{L}_{\text{off}}} \frac{S_{\text{off}}}{S_{\text{on}}}$, with $\mathcal{L}_{\text{on}} \sim 711$ fb^{-1}, $\mathcal{L}_{\text{off}} \sim 88$ fb^{-1}, $S_{\text{on}} = (10.58\text{ GeV})^2$ and $S_{\text{off}} = (10.52\text{ GeV})^2$.

Distributions of some variables after pre-selection are shown in Fig. 7.1. Based on these histogram comparisons, one can define a χ^2-like quantity to determine the possible normalization variation for the continuum MC to be used in the uncertainty calculation. This quantity is used to obtain an asymmetric error, one for the upper and another for the lower variation. They are given by

$$\begin{aligned} \chi^2_{\text{up}} &= \sum_i \frac{(D_i + \Delta D_i - MC_i^{\text{up}})^2}{MC_i^{\text{up}}} \\ \chi^2_{\text{down}} &= \sum_i \frac{(D_i - \Delta D_i - MC_i^{\text{down}})^2}{MC_i^{\text{down}}}, \end{aligned} \tag{7.2}$$



(a) M_{bc} (b) $\ln O^{cs}_{tag}$

(c) M^2_{miss} (d) N_{π^0}

Figure 7.1: Comparison between off-resonance data and continuum MC after pre-selection for some variables, showing a good consistency between both distributions as evidenced by the pull distributions underneath each plot. To get the appropriate variation on the MC to estimate the uncertainty due to the continuum modeling, the normalization of the MC is changed such that it matched the off-resonance data within error bars.

where i runs on the bins of the histogram, D and ΔD are the bin content of the data histogram and its error, MC_i^{up} is the bin content of the MC histogram with normalizations in the range (1,2) and MC_i^{down} is the bin content of the histogram with normalizations in the range (0,1). These quantities are minimized as a function of the normalization variation and are applied to different distributions with different bin widths. The minimum for the upper normalization variation falls in the range (40%,60%), whereas for the lower in (10%,25%). To be conservative, a down variation of 20% and an up variation of 50% are assigned.

7.1.4 Background Shapes

The general idea to determine the effects of the background shape on the calculation of the branching ratio is to reweigh a specific decay in the MC with another model. In this manner, the changes relative to the nominal fit are taken as the systematic uncertainty. The background shapes considered in this analysis are the following:

1. **Form factors for $B \to D^{(*)}\ell\nu$**: The Belle MC uses a generator for these decays based on HQET2 with outdated values for the form factors. This component is reweighted to the most recent values

for these form factors following the procedure indicated in reference [94], which was described in section 5.1.1. The contributions from each form factor are added in quadrature.

2. **Form factors for $B \to D^{**}\ell\nu$**: In an analogous fashion to the previous case, the four narrow excited charm states D_1, D_1', D_0^* and D_2 are reweighted from the ISGW2 to the LLSW model according to reference [94]. The systematic uncertainties due to each form factor are summed in quadrature to claim the total error associated with the $B \to D^{**}\ell\nu$ shape.

3. **Modeling of $B^+ \to \omega\ell^+\nu$ and $B^0 \to \rho^-\ell^+\nu$**: The Belle MC generator for these decays is based on LCSR calculations using the Ball98 model [98], which has the same form factor definition as the SLPOLE package from the EVTGEN package. The most recent calculations from LCSR are based on Ball05 [108], for which the form factor parametrization differs from the SLPOLE package. To simulate a decay with another form factor parametrization not available in the EVTGEN package, it is first produced with the FLATQ2 generator and then reweighted to the new model following the procedure indicated in reference [109]. Hence, 10 million events of these decays are generated according to the Ball98 and FLATQ2 models, the last one is reweighted to the Ball05 model and finally weights from the Ball98 to the Ball05 are calculated in terms of the true q^2 and the lepton momentum. The fit procedure is performed using the nominal MC sample (Ball98) on one hand and the reweighted sampled (from Ball98 to Ball05) on the other. The difference in the signal yield from these two models is taken as the systematic uncertainty.

4. **Modeling of $B \to \pi\ell^+\nu$**: This decay is generated with the Ball98 [98] form factor parametrization in the Belle MC. To study effects due to the shape of this background, the Bourrelly-Caprini-Lellouch (BCL) model [110] is implemented. This model combines the LCSR (valid for low q^2) and LQCD (restricted to the high q^2) calculations to form a model independent parametrization for $B \to \pi\ell^+\nu$. As in the previous cases, 10 million events of $B \to \pi\ell^+\nu$ decay are produced using the Ball98 and FLATQ2 generators, reweight the last one to the BCL parametrization, and produced weights in terms of $(q^2_{\text{true}}, p_\ell)$. The relative difference in the signal event yield between a fit using the nominal MC sample for $B \to \pi\ell^+\nu$ (Ball98) and the reweighted sample (from Ball98 to BCL) is taken as the uncertainty.

5. **Modeling of $B \to \eta^{(\prime)}\ell^+\nu$**: The Belle MC generator used to simulate these decays is based on the ISGW2 model. To assess the error due to the shape of this background, weights from the ISGW2 to the Ball07 [111] model are computed following an analogous procedure to the previous cases. These decays are simulated with 10 million events using the ISGW2 and FLATQ2 generators, reweigth the last one to the Ball07 and generate weights as a function of $(q^2_{\text{true}}, p_\ell)$. The relative difference in the signal event yield between a fit using the nominal MC sample for $B \to \eta'\ell^+\nu$ (ISGW2) and the reweighted sample (from ISGW2 to Ball07) is taken as the uncertainty.

6. **Modeling of other exclusive charmless semileptonic decays**: Other exclusive charmless semileptonic B decays are simulated in the Belle MC with the ISGW2 generator. These decays do not have a well established form factor parametrization, so the strategy is to generate 10 million events using the generators PHSP, FLATQ2 and ISGW2, and produce weights with respect to the last one in terms of $(q^2_{\text{true}}, p_\ell)$. To quote an error for the modeling of these decays, the average between the effects of reweighting the MC to the PHSP and to the FLATQ2 model is taken as the uncertainty.

7.1.5 Branching Ratios of Main Background Contributions

The decays $B \to (D, D^*, D^{**})\ell\nu_\ell$ as well as the exclusive $B \to (\pi, \eta, \rho^+, \omega, \eta')\ell\nu_\ell$ have already been corrected to the world average, as described in section 5.1.1. Since the world average involves an error, the branching ratio for each decay component in the MC is varied separately within their measured uncertainties, according to the HFLAV [41] or PDG [1]. As of 2018, there are no reported measurements of other exclusive charmless semileptonic decays, which constitute the rest of the $B \to X_u\ell\nu$ sample, to compare with. For these decays, the branching ratio is varied with a conservative uncertainty of 100%. The systematic uncertainties due to branching ratios are reported separately for $B \to X_c\ell\nu_\ell$ and $B \to X_u\ell\nu_\ell$, where the individual contributions for each decay were added in quadrature.

7.1.6 Secondary Leptons

Background leptons not stemming from semileptonic B decays are primarily produced in continuum events or as decay products of charm intermediate resonances such as the D or D^* mesons. To assess the magnitude of the effect of the secondary leptons to the total error, one has to determine the total uncertainty in the production of leptons via charm meson decays. This can be estimated using the decay chain $B^+ \to X_c\ell^+\nu_\ell$ with $X_c \to \ell^+ + \text{anything}$. The reference values [1] for these decays are: $\mathcal{B}(B^+ \to X_c\ell^+\nu_\ell) = (10.4 \pm 0.4)\%$, $\mathcal{B}(X_c \to e^+ + \text{anything}) = (6.49 \pm 0.11)\%$ and $\mathcal{B}(X_c \to \mu^+ + \text{anything}) = (6.7 \pm 0.6)\%$. The total branching ratio of the decay chain is $(1.37 \pm 0.10)\%$, from which a total uncertainty of 7.3% is obtained. Therefore, the number of secondary leptons in the MC is varied by 7.3%, and the deviation of the signal yield with respect to the nominal fit result is taken as the systematic uncertainty.

7.2 Detector Simulation

Given that the analysis is optimized using a MC simulation, the detection of final state particles affects the reconstruction of the decay and the subsequent extraction of the signal yields used in the calculation of the branching ratios. The efficiency for detecting these particles in data usually differs from that in the MC, for which a correction to the latter is applied. This correction is derived from independent control samples and the total uncertainty associated with it is taken as as systematic uncertainty. In the following, a discussion of the uncertainties dependent on the detector performance is presented.

7.2.1 Track Finding Efficiency

The recommended uncertainty at Belle of 0.35% per number of charged particles for track finding efficiency [112] is implemented, resulting in a total uncertainty of 1.05% for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay. This study was performed by the Belle tracking group comparing the track finding efficiency in data and MC using the decay chain $D^{*+} \to D^0\pi_{\text{slow}^+}$, $D^0 \to \pi^+\pi^- K^0_S$ and $K^0_S \to \pi^+\pi^-$. This channel offers a clean sample with enough statistics to carry out tracking studies. The reconstruction of the decay chain does not require the detection of one of the charged pions from the K^0_S decay, instead its four-momentum can be inferred from kinematic constraints. The track finding efficiency is then calculated as the ratio between the yields of the fully reconstructed D^{*+} to those partially reconstructed using kinematic constraints on the pion from the K^0_S decay. Finally, the ratio of the tracking efficiencies of data and MC is found to be consistent with unity and its uncertainty is assigned as the systematic uncertainty per track.

7.2.2 Lepton Identification

The lepton identification efficiency correction was discussed in section 5.1.1, and the error associated with this correction is taken as a systematic uncertainty.

7.2.3 Pion Identification

Similar to the previous case, the error associated with the charged pion efficiency correction is regarded as a systematic uncertainty (see section 5.1.1 for more details). The uncertainties are calculated for each pion individually and added linearly.

7.2.4 Fake Leptons

In a similar manner to secondary leptons, charged tracks misidentified as leptons, i.e. fake leptons, can also originate from continuum and charm semileptonic B decays. To inspect the misidentification probability, i.e. fake rate, for a hadron to be classified as a lepton, an enriched hadronic sample is utilized. This sample corresponds to the $D^{*+} \to D^0(K^-\pi^+)\pi^+_{\text{slow}}$ decay[2], that is studied using the on-resonance data and the generic MC (a composition of continuum events and B decays to charm mesons as discussed in section 5.1). The fake rates are calculated via:

$$\text{fake rate} = \frac{\text{No. of non-}\ell\text{ tracks found by tracking with } \mathcal{L}_{\ell id} > X}{\text{No. of non-}\ell\text{ tracks found by tracking}}, \tag{7.3}$$

where $\mathcal{L}_{\ell id}$ is the probability of identifying a track as a lepton and was defined in equations 3.1 and 3.2 for electrons and muons, respectively, and X is the cut-off value used in the analysis, that for both leptons is 0.9. Subsequently, the fake rates in data and MC for each lepton type to pions and kaons are calculated and a weight is assigned as the ratio between the fake rates in data and MC. Afterwards, the average between the weights for kaons and pions is prepared for each lepton type. Thus, the final weight for electron (muon) to hadrons is 1.42 ± 0.32 (1.06 ± 0.04). For each lepton type in the MC, the contribution of fake leptons is corrected by the central value of the weight and varied by its uncertainty. The systematic uncertainty associated with fake leptons is obtained by taking the difference between the fit to the samples corrected with the central values, and the ones with the variation by their errors.

7.3 Fit Procedure

As the central values and statistical errors from the signal yields are corrected due to a small bias in the fit procedure, a systematic uncertainty is assigned accordingly. The uncertainty is defined as half the bias shown in Table 6.6. However, the correction to the signal yields is only applicable to those bins where the pull distribution exhibits a Gaussian shape. For the bins without correction, the average of the bias is calculated as $< N_{\text{fit}} - N_{\text{input}} >$ with 20000 Toy MC and a systematic uncertainty is assigned as half the percent error with respect to nominal fit result (N_{input}).

7.4 Other Systematics

Other sources of systematic uncertainties are discussed in the following.

[2] The slow pion, π^+_{slow}, is characterized by its low momentum, and given the charge correlation with the other tracks allows to determine the flavor and charge of the other particles in the decay.

7.4.1 Number of $B\bar{B}$ Pairs

The total number of $B\bar{B}$ events is a quantity required to calculate the branching ratio for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay. The Belle collaboration has measured this value to be $N_{B\bar{B}} = (771.581 \pm 10.566) \times 10^6$, from which an uncertainty of 1.4% is ascribed.

7.4.2 Final State Radiation (FSR)

The signal MC includes QED radiation corrections due to processes such as Bremsstrahlung, which are simulated using the PHOTOS package. This simulation includes electroweak corrections for the charged leptons, but not for the quarks. To estimate the effects of the FSR, a sample with the same size as the signal is generated without PHOTOS. Histogram templates are produced with PHOTOS switched on (for the nominal fit) and switched off. However, to avoid an overestimation of the uncertainties associated with these processes, it has to be taken into consideration that corrections achieved with PHOTOS agreed with a precision around 1% with decays for which the radiative corrections have been calculated analytically [113]. Measurement of branching ratios of charmless semileptonic B decays performed by BaBar [43, 114], calculate the FSR uncertainty as 20% of the difference between the signal yields with PHOTOS switched on and switched off. A similar procedure is implemented in this thesis, but instead of using the fit results to on-resonance data, fake data are produced with the histogram templates scaled according to the fit results using the nominal signal MC and 4000 toy MC are carried out. Another 4000 toy MC are performed replacing the signal histogram template with one with PHOTOS switched off, without changing the fake data. A 20% of the difference between the mean values of the signal yields with PHOTOS switched on and switched off is taken as the systematic uncertainty. The idea of using a toy MC is to diminish the effect of the statistical fluctuations in the determination of the systematic uncertainty.

7.4.3 Tagging Efficiency Correction

The description of the hadronic tag efficiency correction was presented in section 5.1.1. For more details about this correction the reader is referred to reference [47]. Following the suggestion of this reference, the error associated with the tagging efficiency correction for B^+ mesons is assigned to be 4.2%.

7.5 Summary of Systematic Uncertainties

The breakdown of systematic uncertainties is presented in tables 7.1 and 7.2 for the one and two dimensional binning configurations, respectively. As the errors are uncorrelated, they are added in quadrature to quote a semi-total uncertainty. The latter does not include uncertainties due to the number of $B\bar{B}$ pairs, the track finding efficiency and the tagging efficiency correction, which altogether represents a 4.55% error. These three uncertainties are not computed for each bin, since they are common normalization errors to all the subsamples decided by the $M_{\pi\pi}$ and/or q^2 regions, otherwise they will be counted multiple times as a function of the number of bins.

Source	bin 1	bin 2	bin 3	bin 4	bin 5	bin 6	bin 7	bin 8	bin 9	bin 10	bin 11	bin 12	bin 13	bin 14
FF $B \to D^{(*)}\ell\nu$	$^{+0.05}_{-0.22}$	$^{+0.07}_{-0.08}$	$^{+0.11}_{-0.60}$	$^{+0.12}_{-0.10}$	$^{+0.05}_{-0.06}$	$^{+0.11}_{-0.13}$	±0.11	±0.16	$^{+0.16}_{-0.15}$	$^{+0.12}_{-0.02}$	$^{+0.17}_{-0.07}$	$^{+0.08}_{-0.10}$	$^{+0.30}_{-0.24}$	$^{+0.79}_{-0.60}$
FF $B \to D^{**}\ell\nu$	±0.34	±0.23	±0.22	±0.49	±0.24	±0.11	±0.46	±0.45	±0.67	±0.33	±0.18	±0.08	±0.57	±1.89
Shapes $B \to X_u\ell\nu$	±1.62	±1.35	±0.74	±1.09	±0.19	±0.45	±0.41	±0.62	±0.09	±1.52	±0.13	±0.09	±0.17	±0.43
$\mathcal{B}(B \to D^{(*)}\ell\nu)$	$^{+0.10}_{-0.37}$	±0.01	$^{+0.34}_{-0.28}$	$^{+0.17}_{-0.19}$	$^{+1.70}_{-0.04}$	$^{+0.12}_{-0.16}$	$^{+0.05}_{-0.03}$	±0.14	$^{+0.08}_{-0.23}$	±0.14	$^{+0.31}_{-0.22}$	$^{+0.12}_{-0.10}$	$^{+0.31}_{-0.41}$	$^{+14.24}_{-15.27}$
$\mathcal{B}(B \to D^{**}\ell\nu)$	$^{+0.12}_{-0.09}$	±0.08	$^{+0.38}_{-0.24}$	$^{+0.12}_{-0.21}$	±0.05	±0.08	$^{+0.14}_{-0.09}$	$^{+0.08}_{-0.12}$	$^{+0.10}_{-0.14}$	±0.09	$^{+0.14}_{-0.09}$	$^{+0.20}_{-0.11}$	$^{+0.31}_{-0.39}$	$^{+1.09}_{-1.06}$
$\mathcal{B}(B \to X_u\ell\nu)$	$^{+5.93}_{-6.54}$	$^{+2.99}_{-3.02}$	$^{+3.48}_{-4.14}$	$^{+3.21}_{-3.44}$	$^{+0.92}_{-0.87}$	$^{+1.74}_{-1.66}$	$^{+1.83}_{-5.13}$	$^{+3.10}_{-3.48}$	±0.55	$^{+0.19}_{-0.41}$	$^{+0.32}_{-0.43}$	$^{+0.19}_{-0.31}$	$^{+0.94}_{-0.33}$	±0.35
Continuum	$^{+0.91}_{-0.34}$	$^{+0.98}_{-0.29}$	$^{+0.09}_{-0.14}$	$^{+0.53}_{-0.19}$	$^{+0.17}_{-0.09}$	$^{+0.04}_{-0.00}$	$^{+0.41}_{-0.15}$	$^{+2.02}_{-0.76}$	$^{+0.00}_{-0.26}$	$^{+0.57}_{-0.20}$	$^{+0.03}_{-0.08}$	$^{+0.29}_{-0.23}$	$^{+1.19}_{-0.10}$	$^{+12.76}_{-6.85}$
Rare	±1.19	±1.60	±1.08	±1.23	±0.28	±0.47	±0.80	±0.79	±0.95	±3.20	±0.98	±0.78	±2.74	±4.11
Sec. Leptons	$^{+0.04}_{-0.28}$	$^{+0.16}_{-0.07}$	$^{+0.10}_{-0.18}$	$^{+0.21}_{-0.17}$	$^{+0.02}_{-0.06}$	$^{+0.09}_{-0.01}$	$^{+0.10}_{-0.08}$	$^{+0.14}_{-0.08}$	$^{+0.16}_{-0.21}$	$^{+0.03}_{-0.04}$	$^{+0.07}_{-0.10}$	$^{+0.21}_{-0.26}$	$^{+0.25}_{-0.06}$	$^{+2.78}_{-2.85}$
Fake leptons	±0.44	±0.10	$^{+0.07}_{-0.03}$	$^{+0.30}_{-0.17}$	$^{+0.01}_{-0.05}$	$^{+0.08}_{-0.10}$	$^{+0.02}_{-0.07}$	$^{+1.05}_{-0.81}$	$^{+0.04}_{-0.20}$	$^{+0.22}_{-0.21}$	$^{+0.56}_{-0.25}$	$^{+0.14}_{-0.23}$	$^{+0.59}_{-0.26}$	$^{+3.01}_{-3.83}$
ℓID	±1.85	±1.90	±1.90	±1.87	±1.93	±1.90	±1.93	±1.83	±1.89	±1.85	±1.89	±1.89	±1.91	±2.02
πID	±0.98	±0.98	±0.95	±1.00	±0.98	±1.00	±1.01	±0.99	±1.00	±1.01	±1.04	±1.02	±1.02	±1.19
FSR	±0.08	±0.72	±0.81	±1.13	±0.60	±0.02	±0.30	±0.15	±0.50	±0.72	±0.34	±0.91	±1.66	±1.15
Signal model	±1.65	±2.59	±9.94	±10.65	±9.11	±5.73	±6.64	±8.49	±1.29	±0.93	±4.47	±4.44	±0.41	±9.38
Fit procedure	±2.62	±1.95	±1.88	±1.22	±0.32	±0.52	±0.82	±1.98	±1.95	±3.27	±1.35	±1.68	±2.25	±5.26
Semitotal	$^{+7.37}_{-7.84}$	$^{+5.48}_{-5.41}$	$^{+11.03}_{-11.27}$	$^{+11.59}_{-11.65}$	$^{+9.60}_{-9.44}$	$^{+6.42}_{-6.40}$	$^{+7.36}_{-8.77}$	$^{+9.82}_{-9.74}$	$^{+3.46}_{-3.49}$	$^{+5.44}_{-5.43}$	$^{+5.31}_{-5.28}$	$^{+5.37}_{-5.37}$	$^{+4.85}_{-4.59}$	$^{+22.96}_{-21.16}$

Table 7.1: Relative systematic uncertainties in percent for the fits in $M_{\pi\pi}$ bins.

Source	bin 1	bin 2	bin 3	bin 4	bin 5	bin 6	bin 7	bin 8	bin 9	bin 10	bin 11	bin 12	bin 13	bin 14	bin 15
FF $B \to D^{(*)}\ell\nu$	$^{+0.37}_{-0.41}$	$^{+0.30}_{-0.04}$	$^{+0.07}_{-0.06}$	$^{+0.06}_{-0.15}$	$^{+0.06}_{-0.05}$	$^{+0.10}_{-0.10}$	$^{+0.13}_{-0.14}$	$^{+0.00}_{-0.18}$	$^{+0.12}_{-0.11}$	$^{+0.42}_{-0.33}$	$^{+0.31}_{-0.10}$	$^{+0.11}_{-0.24}$	$^{+1.86}_{-2.84}$	$^{+0.79}_{-0.83}$	$^{+5.00}_{-8.37}$
FF $B \to D^{**}\ell\nu$	±1.08	±0.35	±0.16	±0.14	±0.20	±0.13	±0.73	±0.47	±0.21	±0.63	±0.33	±0.24	±1.79	±2.05	±2.23
Shapes $B \to X_u\ell\nu$	±3.39	±1.54	±2.38	±0.85	±0.64	±0.59	±0.62	±0.67	±0.39	±0.59	±0.58	±0.34	±0.70	±0.24	±3.12
$\mathcal{B}(B \to D^{(*)}\ell\nu)$	$^{+0.25}_{-0.51}$	$^{+0.19}_{-0.30}$	±0.07	$^{+0.22}_{-0.23}$	$^{+0.06}_{-0.01}$	$^{+0.13}_{-0.08}$	$^{+0.09}_{-0.07}$	$^{+0.23}_{-0.39}$	$^{+0.17}_{-0.35}$	$^{+0.63}_{-0.86}$	$^{+0.50}_{-0.68}$	$^{+0.25}_{-0.33}$	$^{+8.30}_{-7.46}$	$^{+3.84}_{-4.10}$	$^{+6.72}_{-5.86}$
$\mathcal{B}(B \to D^{**}\ell\nu)$	$^{+0.38}_{-0.64}$	$^{+0.38}_{-0.12}$	±0.04	$^{+0.16}_{-0.17}$	$^{+0.02}_{-0.05}$	$^{+0.12}_{-0.05}$	$^{+0.25}_{-0.34}$	$^{+0.47}_{-0.19}$	$^{+0.11}_{-0.13}$	$^{+0.77}_{-0.81}$	$^{+0.53}_{-0.56}$	$^{+0.22}_{-0.13}$	$^{+0.65}_{-0.52}$	±0.79	$^{+2.43}_{-1.48}$
$\mathcal{B}(B \to X_u\ell\nu)$	$^{+19.80}_{-19.78}$	$^{+4.86}_{-6.34}$	$^{+2.96}_{-3.22}$	$^{+0.65}_{-0.92}$	$^{+1.56}_{-1.82}$	$^{+1.79}_{-1.94}$	$^{+1.01}_{-1.06}$	$^{+0.94}_{-1.56}$	$^{+1.51}_{-1.34}$	$^{+0.62}_{-0.46}$	$^{+0.65}_{-0.50}$	$^{+0.25}_{-0.18}$	$^{+1.25}_{-0.91}$	$^{+0.73}_{-0.28}$	$^{+5.99}_{-3.83}$
Continuum	$^{+4.39}_{-1.65}$	$^{+0.21}_{-0.04}$	$^{+0.43}_{-0.14}$	$^{+0.90}_{-0.49}$	$^{+0.02}_{-0.03}$	$^{+0.02}_{-0.10}$	$^{+0.97}_{-0.46}$	$^{+1.08}_{-0.71}$	$^{+0.82}_{-0.16}$	$^{+1.29}_{-0.01}$	$^{+0.26}_{-0.25}$	$^{+0.34}_{-0.02}$	$^{+1.83}_{-0.02}$	$^{+5.18}_{-2.76}$	±0.00
Rare	±4.17	±1.41	±0.58	±1.89	±0.44	±0.19	±4.25	±0.54	±0.22	±3.97	±0.02	±0.78	±5.95	±1.53	±1.01
Sec. Leptons	$^{+0.35}_{-0.04}$	$^{+0.16}_{-0.22}$	$^{+0.01}_{-0.07}$	$^{+0.09}_{-0.15}$	$^{+0.02}_{-0.05}$	$^{+0.02}_{-0.05}$	±0.03	$^{+0.30}_{-0.57}$	$^{+0.21}_{-0.01}$	$^{+0.56}_{-0.17}$	$^{+0.24}_{-0.27}$	$^{+0.11}_{-0.04}$	$^{+0.70}_{-1.43}$	$^{+0.17}_{-0.34}$	$^{+0.32}_{-0.77}$
Fake leptons	$^{+1.48}_{-0.27}$	$^{+0.01}_{-0.04}$	$^{+0.55}_{-0.57}$	$^{+0.23}_{-0.01}$	$^{+0.06}_{-0.16}$	$^{+0.14}_{-0.09}$	$^{+0.79}_{-0.83}$	$^{+0.39}_{-0.16}$	$^{+0.14}_{-0.46}$	$^{+0.23}_{-0.59}$	$^{+0.13}_{-0.12}$	$^{+0.10}_{-0.08}$	$^{+1.16}_{-1.03}$	$^{+0.81}_{-1.13}$	$^{+1.53}_{-0.38}$
ℓID	±1.78	±2.04	±1.85	±1.86	±2.00	±1.90	±1.83	±1.97	±1.83	±1.89	±1.97	±1.83	±2.08	±1.98	±1.81
πID	±1.14	±1.10	±0.81	±1.14	±1.04	±0.85	±1.10	±1.05	±0.90	±1.08	±1.08	±0.97	±1.23	±1.16	±1.11
FSR	±1.40	±2.40	±0.98	±0.48	±0.90	±0.24	±0.76	±0.54	±0.36	±1.24	±1.12	±0.76	±0.84	±0.64	±1.68
Signal model	±9.84	±1.44	±2.04	±4.76	±1.83	±5.42	±12.54	±3.04	±1.79	±5.07	±2.00	±1.44	±7.50	±5.67	±6.25
Fit procedure	±0.90	±2.93	±1.71	±0.68	±0.64	±0.28	±2.14	±2.15	±1.19	±1.99	±2.19	±1.57	±3.66	±4.42	±22.67
Semitotal	$^{+23.41}_{-23.00}$	$^{+7.09}_{-8.16}$	$^{+5.24}_{-5.37}$	$^{+5.81}_{-5.80}$	$^{+3.57}_{-3.69}$	$^{+6.12}_{-6.17}$	$^{+13.73}_{-13.71}$	$^{+4.77}_{-4.86}$	$^{+3.49}_{-3.37}$	$^{+7.48}_{-7.38}$	±4.09	$^{+3.24}_{-3.22}$	$^{+13.96}_{-13.54}$	$^{+10.39}_{-9.55}$	$^{+26.28}_{-26.42}$

Table 7.2: Relative systematic uncertainties in percent for the fits in $M_{\pi\pi}$ and q^2 bins.

CHAPTER 8

RESULTS AND DISCUSSION

The results obtained in chapters 6 and 7 are now employed to measure the total branching ratio of the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay using the formula defined in section 8.1. In the subsection 8.1.1, the $\mathcal{B}(B^+ \to D^0\ell^+\nu_\ell)$ with the D^0 meson decaying to a charged pion pair is calculated to validate the signal selection. The nominal result of this thesis is provided in subsection 8.1.2, which utilizes the results of the one dimensional fitting configuration (the measurement in bins of the dipion mass); the other branching ratios calculations in this subsection are intended to support the nominal result. An study with control samples to validate the branching ratio measurement is presented in subsection 8.1.3. In spite of the poor statistics after signal selection, an attempt to study possible resonances in the dipion mass spectrum is carried out using different approaches in section 8.2. The first approach uses shapes from MC simulation for the $B^+ \to \rho^0\ell^+\nu$ and $B^+ \to f_2(1270)\ell^+\nu$ decays, and scalar from factor derived from dispersion theory to model the remainder of the events. The second approach models the dipion mass as a incoherent sum of resonances. The last approach takes into account the interferences of the resonances in the second approach to analyze the dipion mass. The three approaches give consistent results for the $\mathcal{B}(B^+ \to \rho^0\ell^+\nu_\ell)$, while the branching ratios for the other assumed resonances are not statistically significant. In section 8.3, a comparison of the values of $|Vub|$ using the total branching ratio for $B^+ \to \rho^0\ell^+\nu$ decay obtained in this thesis and previous measurements is discussed. The final comments about the results for this thesis are addressed in section 8.4.

8.1 Branching Ratio calculation

The main result of this thesis is the measurement of the total branching ratio for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay, which as of 2018 has not been measured in the whole phase space available to the decay. Since this measurement is carried out in bins of the kinematic variables $M_{\pi\pi}$ or q^2, the total branching ratio is calculated as the sum over all bins of the partial branching ratios. The latter are defined as

$$\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell) = \frac{1}{4}\frac{Y_{\text{signal}}}{\epsilon}\frac{1}{\mathcal{B}(\Upsilon(4S) \to B^+B^-)N_{B\bar{B}}}, \tag{8.1}$$

where Y_{signal} is the corrected signal yield in a given bin as presented in Tab. 6.6, ϵ is the signal reconstruction efficiency for a specific bin, $\mathcal{B}(\Upsilon(4S) \to B^+B^-) = (51.4 \pm 0.6)\%$ a factor that corrects

$B \to D^0(\pi^+\pi^-)\ell\nu$	47.4±8.5
Other $B \to D^0\ell\nu$	0.0±2.8
Rest of $B \to X_c\ell\nu$	35.5±7.7
Fixed background	3.0
Data	86
χ^2/ndf	14.15/10
$\epsilon\,[10^{-4}]$	7.42±0.48

Table 8.1: Events yields and reconstruction efficiency for the $B^+ \to D^0(\pi^+\pi^-)\ell^+\nu_\ell$ decay.

for the number of charged B meson-pairs produced in the decay of the $\Upsilon(4S)$ resonance[1], and $N_{B\bar{B}} = (771.6 \pm 10.6) \times 10^6$ is the number of $B\bar{B}$-pairs produced for the complete Belle dataset. The signal reconstruction efficiency is determined from MC simulations, with corrections for differences between data and simulated detector performance. The factor of 4 in the denominator appears as an average among four channels $B^+ \to \pi^+\pi^- e^+\nu_e$, $B^+ \to \pi^+\pi^-\mu^+\nu_\mu$, $B^- \to \pi^+\pi^- e^-\bar{\nu}_e$ and $B^- \to \pi^+\pi^-\mu^-\bar{\nu}_\mu$.

8.1.1 Measurement of $\mathcal{B}(B^+ \to D^0\ell^+\nu_\ell)$ with $D^0 \to \pi^+\pi^-$ using the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ selection

The $B^+ \to D^0\ell^+\nu_\ell$ decay undergoes a $b \to c$ transition at the quark level, the dominant decay process at B decays. As such, this decay channel can be used as a control sample to validate the measurement of the signal. Usually, a decay channel can be used as a control sample if its topology is similar to that of the signal. Because the D^0 meson can decay into a pair of charged pions, the current data sample can be used to measure the branching ratio of the $B^+ \to D^0\ell^+\nu_\ell$ decay and compare this result with the world average. For that purpose, a small region in the dipion mass, where the D^0 meson is dominant, is selected. This region is limited to the interval [1.85 GeV, 1.88 GeV] corresponding to a 3σ mass window around the nominal D^0 mass; in this region the contributions from charmless semileptonic B decays are negligible. Given that the D^0 meson is an intermediate resonance in the dipion mass, Eq. 8.1 has to be modified to

$$\mathcal{B}(B^+ \to D^0(\pi^+\pi^-)\ell^+\nu_\ell) = \frac{1}{4}\frac{Y_{\text{signal}}}{\epsilon}\frac{1}{\mathcal{B}(\Upsilon(4S) \to B^+B^-)\, N_{B\bar{B}}\, \mathcal{B}(D^0 \to \pi^+\pi^-)}, \tag{8.2}$$

where $\mathcal{B}(D^0 \to \pi^+\pi^-) = (1.420 \pm 0.025) \times 10^{-3}$ is the world average [1] branching ratio of the $D^0 \to \pi^+\pi^-$ decay.

Following the same procedure as for the signal extraction, the event yields for the $B^+ \to D^0(\pi^+\pi^-)\ell^+\nu_\ell$ decay are obtained from an extended binned maximum likelihood fit to the M^2_{miss} distribution. The fitted M^2_{miss} distribution for this decay channel is shown in Fig. 8.1. The event yields for each process involve in this decay as well as the $B^+ \to D^0(\pi^+\pi^-)\ell^+\nu_\ell$ reconstruction efficiency is presented in Tab. 8.1. Plugging in these values into Eq. 8.2, one obtains $\mathcal{B}(B^+ \to D^0\ell^+\nu_\ell) = (2.83 \pm 0.54)\%$, where the quoted uncertainty is statistical only. This value is consistent with the HFLAV [41] value $\mathcal{B}(B^+ \to D^0\ell^+\nu_\ell) = (2.33 \pm 0.10)\%$.

[1] In the production of the MC samples samples at Belle, the production rate of neutral and charged B mesons from an $\Upsilon(4S)$ is assumed to be the same (50%).

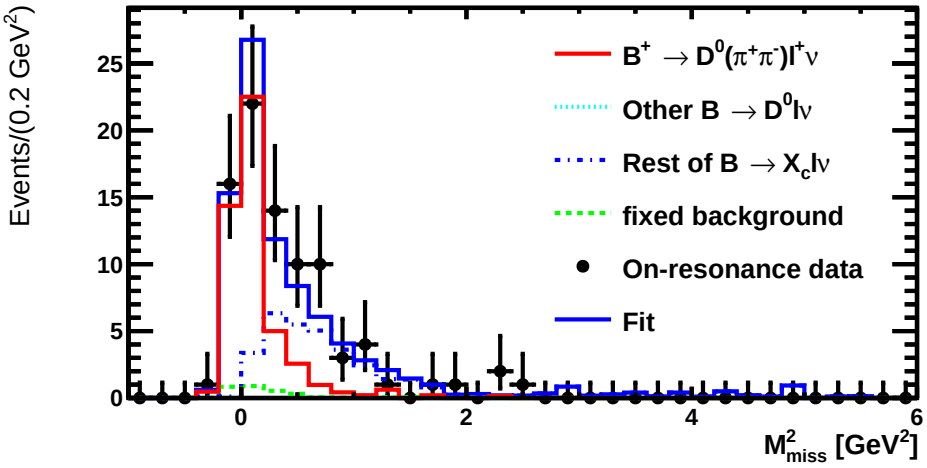


Figure 8.1: Distribution of M^2_{miss} after the fitting procedure for the $B^+ \to D^0(\pi^+\pi^-)\ell^+\nu_\ell$ decay.

8.1.2 Measurement of $\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$

The measurement of the total branching ratio for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay is made using two configurations. The one dimensional (1D) configuration is performed in 14 bins of the dipion mass, whereas the two dimensional (2D) configuration is carried out in 15 bins of the dipion mass and the momentum transfer. The values of the corrected signal yields, signal reconstruction efficiency and partial branching ratios (calculated according to Eq. 8.1) for each bin in the 1D and 2D configurations are presented in Tab. 8.2 and Tab. 8.3, respectively. The statistical significance, Z, for each measured partial branching ratio is also presented and is computed as $Z = \sqrt{-2\Delta\mathcal{L}}$. Here, $\Delta\mathcal{L} = \mathcal{L}_{S+B} - \mathcal{L}_B$ where $\mathcal{L}_{S+B}$ and $\mathcal{L}_B$ are the values of the constructed profile Log-Likelihood, with the inclusion of systematic uncertainties, at its minimum (signal plus background hypothesis) and for zero signal (background only hypothesis), respectively.

Adding the partial branching ratios in Tab. 8.2, the total branching ratio in the 1D configuration results in

$$\begin{aligned}\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)[\text{1D}] &= (22.4 \pm 1.9_{\text{stat}} \pm 0.8_{\text{syst1}} \pm 1.0_{\text{syst2}}) \times 10^{-5} \\ &= (22.4 \pm 1.9_{\text{stat}} \pm 1.3_{\text{syst}}) \times 10^{-5},\end{aligned} \tag{8.3}$$

where the index syst1 stands for the systematics uncertainties calculated bin by bin and added in quadrature, and syst2 represents the additional 4.55% systematic uncertainty that must be included in the total branching ratio calculation due to the combined effect of the tagging efficiency correction, track finding efficiency and the number of $B\bar{B}$-pairs. The index syst corresponds to the sum in quadrature of the of the two aforementioned systematic uncertainties. The total statistical significance of the measurement is calculated as $Z_{\text{total}} = \sqrt{-2\sum_i \Delta\mathcal{L}_i}$ and determined to be $Z_{\text{total}} = 28.4\sigma$ for the 1D configuration. Analogously, the total branching ratio in the 2D configuration is

$$\begin{aligned}\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)[\text{2D}] &= (23.2 \pm 1.8_{\text{stat}} \pm 0.6_{\text{syst1}} \pm 1.1_{\text{syst2}}) \times 10^{-5} \\ &= (23.2 \pm 1.8_{\text{stat}} \pm 1.2_{\text{syst}}) \times 10^{-5},\end{aligned} \tag{8.4}$$

with a statistical significance $Z_{\text{total}} = 29.8\sigma$. Both measurements provide consistent results. Hereafter,

the branching ratio from the one dimensional configuration is regarded as the nominal measurement from this thesis.

It is also useful to crosscheck these results using subsamples of the final dataset, for example separating the decay into contributions by lepton flavor or by the charge of the B meson. The tables containing the signal yields, signal reconstruction efficiency and partial branching ratios for each bin in the 1D configuration for these four subsamples are given in the Appendix A.4. The summary of the systematic uncertainties for these subsamples is also provided.

If the samples are separated into electron and muon channels, the total branching ratio reads as

$$\mathcal{B}(B^{\pm} \to \pi^+\pi^- e^{\pm}\nu_e) = (22.0 \pm 2.6_{\text{stat}} \pm 1.3_{\text{syst}}) \times 10^{-5} \tag{8.5}$$

and

$$\mathcal{B}(B^{\pm} \to \pi^+\pi^- \mu^{\pm}\nu_\mu) = (22.2 \pm 2.7_{\text{stat}} \pm 1.5_{\text{syst}}) \times 10^{-5}, \tag{8.6}$$

respectively. The two values are comparable and consistent with the nominal value in the 1D configuration, and thus supporting the SM assumption of lepton universality by which the interactions of the charged leptons differ only by their different masses [115]. As the masses of the electron and the muon are negligible compared to the mass of the B meson, their branching ratios have to be equal for lepton universality to hold.

If on the other hand, the samples are separated according to the charged of the B meson, the values of the total branching ratio for B^- and B^+ are:

$$\mathcal{B}(B^- \to \pi^+\pi^- \ell^- \bar{\nu}_\ell) = (20.2 \pm 2.7_{\text{stat}} \pm 1.2_{\text{syst}}) \times 10^{-5} \tag{8.7}$$

and

$$\mathcal{B}(B^+ \to \pi^+\pi^- \ell^+ \nu_\ell) = (23.2 \pm 3.1_{\text{stat}} \pm 1.2_{\text{syst}}) \times 10^{-5}, \tag{8.8}$$

respectively. Note the difference in the central values in both cases, but since they are consistent within error bars a conclusive statement cannot be formulated. In reference [116], the authors estimate that CP violation can be observed in the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ channel, but it would require a dataset containing around 10^{18} $B\bar{B}$-pairs to observe an asymmetry in the decay rates for B^+ and B^- at the 1σ level. The current Belle dataset contains around 10^9 $B\bar{B}$-pairs, for which a task to verify such an asymmetry is entrusted to Belle II.

A summary of the results obtained in this section is illustrated in Fig. 8.2. Notice that, within error bars all branching ratio calculations are consistent.

8.1.3 Study of Control Samples

To validate the signal selection, the fit procedure and the measurement of the signal branching ratio, one can use decays with a large sample size as control channels. Their main feature lies in the similarity in the event topology with that of the signal channel, in other words, the configuration of final state particles resembles that of the signal. Due to its large branching ratio, semileptonic decays to a D^0 meson ($B^+ \to D^0\ell^+\nu$) serve as control samples, with the final state particles consisting of three tracks as the signal, one lepton, and two charged hadrons, and the possibility to include a neutral pion. Hence, the D^0 meson is reconstructed in the following channels: $D^0 \to K^-\pi^+$, $D^0 \to K^-K^+$, $D^0 \to \pi^-\pi^+\pi^0$ and $D^0 \to K^-\pi^+\pi^0$. To ensure the reconstruction of a real D^0 meson a 3σ mass window is chosen around its nominal mass, corresponding to the interval (1.85 GeV, 1.88 GeV) for modes with two tracks, and (1.83 GeV, 1.9 GeV) for modes with two tracks and a neutral pion. The same variables as used in the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ analysis are used in the control samples to train a BDTG. These BDTGs have been

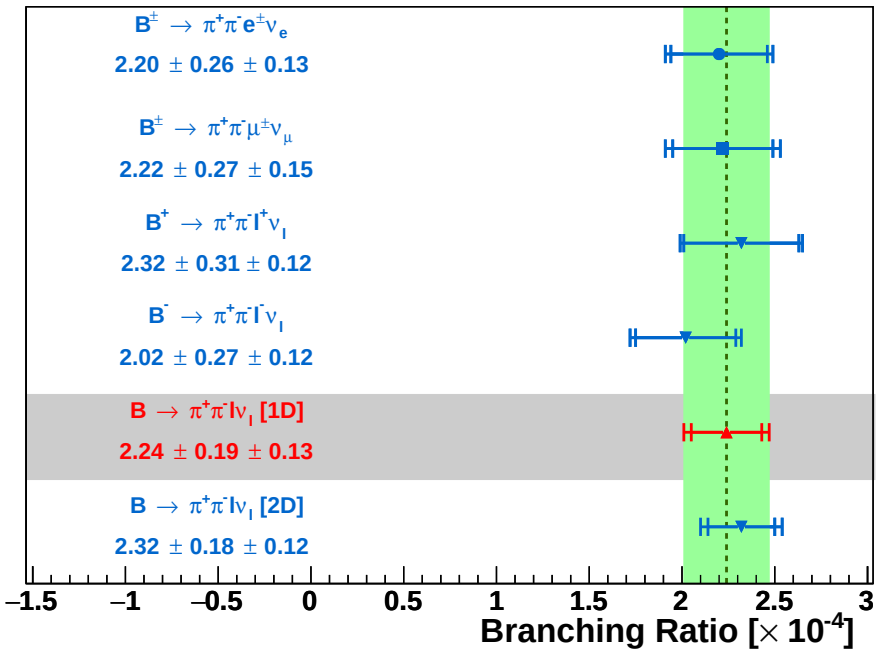


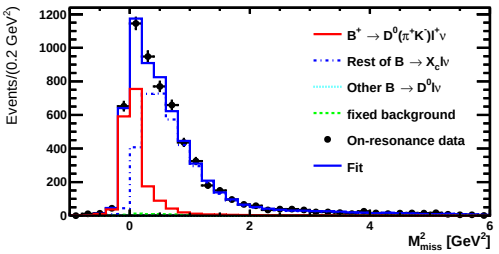
Figure 8.2: Comparison of the branching ratios for semileptonic B decays to two pions as function of the lepton flavor ($\ell = e$ or $\ell = \mu$), the charged of the B meson, and the combined results in the two fitting configurations studied in this thesis. The results from the one dimensional configuration are quoted as the main result of this analysis.

optimized for each mode using the figure of merit $FOM = S/\sqrt{S+B}$. The extraction of the event yields for each channel is carried out with an extended binned maximum likelihood fit to M^2_{miss} with a bin size of 0.2 GeV^2. As the D^0 exhibits a large variety of decay modes, the $B^+ \to D^0\ell^+\nu$ is separated into a sample containing the reconstructed channel and another with the rest of the $B^+ \to D^0\ell^+\nu$ decays. Thus, the fit parameters are : the reconstructed channel $B^+ \to (D^0 \to X)\ell^+\nu$, other $B^+ \to D^0\ell^+\nu$ decays, the rest of the $B \to X_c\ell^+\nu$ and the fixed background (continuum, $B \to X_u\ell\nu_\ell$ and rare decays). The fit results, as well as the calculated branching ratios for each control channel, are presented in Tab. 8.4. The fitted M^2_{miss} distributions for each control channel are shown in Fig. 8.3. Table 8.5 shows a summary of the systematic uncertainties for the branching ratio calculation.

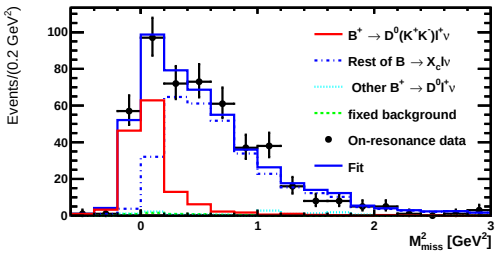
The individual measured branching ratios for each control channel are consistent with the world average [41] $\mathcal{B}(B^+ \to D^0\ell^+\nu_\ell) = (2.33 \pm 0.10)\%$, reassuring the branching ratio measurement of the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay.

Bin	Cor. Y_{signal}	ϵ [10^{-4}]	$\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$ [10^{-5}]	$\mathcal{L}_{S+B}$	$\mathcal{L}_B$	Z [σ]
1	6.9±3.3	7.86±0.66	$0.55 \pm 0.27 \pm 0.04$	-592.8	-588.7	2.9
2	10.2±3.6	8.13±0.77	$0.79 \pm 0.29 \pm 0.04$	-383.9	-366.5	5.9
3	10.5±3.7	7.75±0.68	$0.85 \pm 0.31\ ^{+0.09}_{-0.10}$	-514.9	-501.5	5.2
4	23.9±5.6	7.82±0.64	$1.93 \pm 0.48 \pm 0.22$	-522.5	-493.0	7.7
5	90.5±10.1	9.32±0.66	$6.12 \pm 0.81\ ^{+0.59}_{-0.58}$	-990.8	-793.3	19.9
6	51.4±7.4	7.76±0.58	$4.18 \pm 0.68 \pm 0.27$	-523.9	-442.3	12.8
7	29.7±5.9	8.18±0.57	$2.29 \pm 0.48\ ^{+0.17}_{-0.20}$	-278.1	-235.7	9.2
8	10.6±3.5	8.47±0.57	$0.79 \pm 0.27 \pm 0.08$	-119.5	-185.0	5.4
9	9.2±3.2	8.79±0.56	$0.66 \pm 0.23 \pm 0.02$	-265.8	-252.2	5.2
10	6.1±2.5	8.98±0.56	$0.43 \pm 0.18 \pm 0.02$	-323.2	-316.2	3.7
11	16.1±4.4	9.04±0.55	$1.12 \pm 0.31 \pm 0.06$	-484.0	-464.2	6.3
12	12.2±3.6	8.20±0.52	$0.94 \pm 0.28 \pm 0.05$	-617.5	-585.7	8.0
13	8.8±3.3	8.51±0.52	$0.65 \pm 0.25 \pm 0.03$	-168.4	-157.9	4.6
14	12.4±13.4	7.40±0.10	$1.06 \pm 1.14\ ^{+0.24}_{-0.22}$	-1694.7	-1694.2	1.0

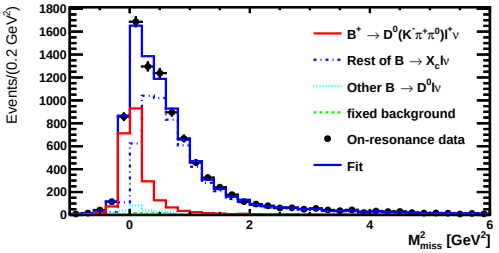
Table 8.2: Corrected signal yields (Y_{signal}), signal reconstruction efficiency (ϵ) and partial branching ratios ($\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$) for each bin in the one dimensional configuration. The first quoted uncertainty is statistical and the second is systematical. The values of the profile likelihood, including systematic uncertainties, for the signal plus background ($\mathcal{L}_{S+B}$) and for the background only ($\mathcal{L}_B$) hypotheses are also presented. They are used to calculate the statistical significance Z of the branching ratio measurement for each bin in terms of σ. The reader is referred to Tab. 6.3 to see the convention for the bin number.



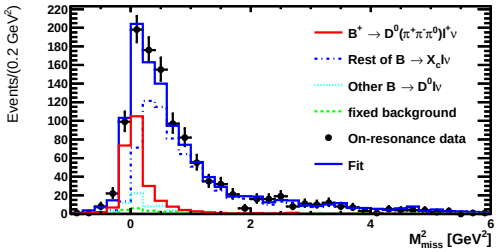
(a) $D^0 \to K^-\pi^+$



(b) $D^0 \to K^-K^+$



(c) $D^0 \to K^-\pi^+\pi^0$



(d) $D^0 \to \pi^-\pi^+\pi^0$

Figure 8.3: Fitted distributions of the missing mass squared for the four D^0 reconstructed modes used as control samples.

Bin	Cor. Y_{signal}	ϵ [10^{-4}]	$\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$ [10^{-5}]	$\mathcal{L}_{S+B}$	$\mathcal{L}_B$	Z [σ]
1	2.3±2.2	6.71±0.62	0.22 ± 0.21 ± 0.05	-148.9	-147.9	1.4
2	6.6±3.0	8.18±0.73	0.51 ± 0.24 ± 0.04	-278.7	-271.7	3.7
3	16.3±4.8	8.50±0.53	$1.21 \pm 0.36\ ^{+0.06}_{-0.07}$	-948.1	-933.3	5.5
4	29.3±5.7	8.61±0.50	2.14 ± 0.43 ± 0.12	-382.5	-337.9	9.5
5	34.3±6.3	8.35±0.53	$2.59 \pm 0.50\ ^{+0.09}_{-0.10}$	-692.8	-641.7	10.1
6	117.4±11.7	8.00±0.39	9.25 ± 1.03 ± 0.57	-1490.0	-1276.0	20.7
7	8.5±3.1	7.23±0.38	0.75 ± 0.27 ± 0.10	-210.5	-202.2	4.1
8	9.2±3.4	9.07±0.47	0.64 ± 0.24 ± 0.03	-483.6	-473.2	4.6
9	28.5±5.8	9.78±0.42	1.83 ± 0.38 ± 0.06	-497.0	-452.9	9.4
10	11.6±3.7	7.85±0.36	0.94 ± 0.30 ± 0.07	-312.8	-297.3	5.6
11	9.8±3.5	8.48±0.41	0.73 ± 0.26 ± 0.03	-622.3	-610.6	4.8
12	13.5±3.9	9.00±0.41	0.95 ± 0.28 ± 0.03	-337.1	-317.0	6.3
13	8.5±9.4	6.78±0.10	0.79 ± 0.87 ± 0.11	-506.8	-506.2	1.1
14	7.0±5.5	8.48±0.19	0.52 ± 0.41 ± 0.05	-449.3	-446.1	2.5
15	1.4±4.0	8.79±0.35	0.10 ± 0.29 ± 0.03	-137.039	-136.975	0.4

Table 8.3: Corrected signal yields (Y_{signal}), signal reconstruction efficiency (ϵ) and partial branching ratios ($\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$) for each bin in the two dimensional configuration. The first quoted uncertainty is statistical and the second is systematical. The values of the profile likelihood, including systematic uncertainties, for the signal plus background ($\mathcal{L}_{S+B}$) and for the background only ($\mathcal{L}_B$) hypotheses are also presented. They are used to calculate the statistical significance Z of the branching ratio measurement for each bin in terms of σ. The reader is referred to Tab. 6.4 to see the convention for the bin number.

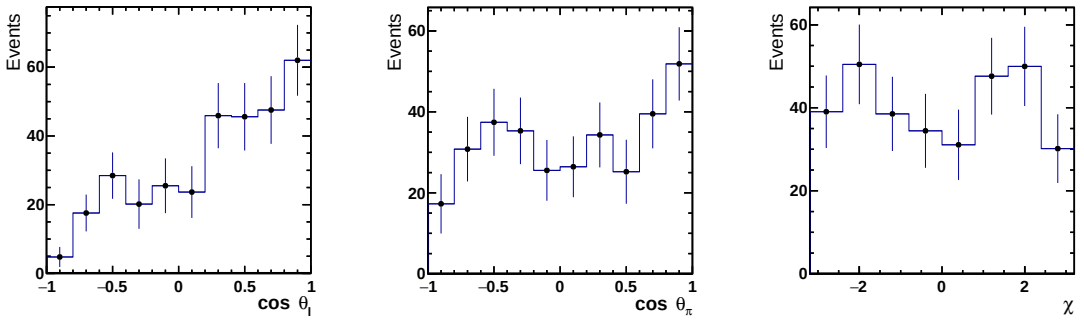
	$D^0 \to K^-\pi^+$	$D^0 \to K^-K^+$	$D^0 \to K^-\pi^+\pi^0$	$D^0 \to \pi^-\pi^+\pi^0$
Data	5912	484	8884	1123
Fixed Background	62.2	5.6	98.5	38.9
$B \to X_c\ell\nu$	4078±74	326±25	6086±111	753±41
Other $B \to D^0\ell\nu$	0±94	12.6±16.7	328±122	76.5±65.4
$B \to (D^0 \to X)\ell\nu$	1772±56	139±16	2371±96	256±48
$\mathcal{B}^{PDG}(D^0 \to X)$ [%]	3.89 ± 0.04	0.397 ± 0.007	14.2 ± 0.5	1.47 ± 0.06
ϵ [10^{-3}]	1.37 ± 0.01	1.07 ± 0.03	0.468 ± 0.004	0.50 ± 0.01
$\mathcal{B}^{\text{this}}(B \to D^0\ell\nu)$ [%]	2.10 ± 0.07 ± 0.11	2.06 ± 0.24 ± 012	2.25 ± 0.09 ± 0.15	2.20 ± 0.41 ± 0.15
$\mathcal{B}^{\text{HFLAV}}(B \to D^0\ell\nu)$ [%]	2.33±0.10			

Table 8.4: Fitted event yields, world average for the branching ratio of the D^0 decay modes and reconstruction efficiencies for the four D^0 decay channels used as control samples. The calculated branching ratios for the $B^+ \to D^0\ell^+\nu$) decay in each control sample are also shown, where the first and second uncertainties are of statistical and systematical nature, respectively.

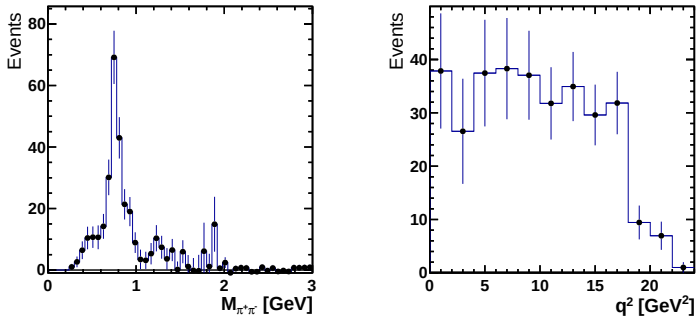
Source	$D^0 \to K^-\pi^+$	$D^0 \to K^-K^+$	$D^0 \to K^-\pi^+\pi^0$	$D^0 \to \pi^-\pi^+\pi^0$
FF $B \to D^{(*)}\ell\nu$	$^{+0.28}_{-0.33}$	$^{+0.33}_{-0.34}$	$^{+0.46}_{-0.52}$	$^{+0.57}_{-0.59}$
FF $B \to D^{**}\ell\nu$	± 0.05	± 0.08	$^{+0.15}_{-0.08}$	$^{+0.25}_{-0.57}$
Shapes $B \to X_u\ell\nu$	± 0.00	± 0.00	± 0.00	± 0.01
$\mathcal{B}(B \to D^{(*)}\ell\nu)$	$^{+0.17}_{-0.18}$	$^{+0.12}_{-0.11}$	$^{+0.43}_{-0.44}$	$^{+0.40}_{-0.39}$
$\mathcal{B}(B \to D^{**}\ell\nu)$	± 0.12	$^{+0.11}_{-0.08}$	$^{+0.37}_{-0.40}$	$^{+0.47}_{-0.50}$
$\mathcal{B}(B \to X_u\ell\nu)$	± 0.00	$^{+0.00}_{-0.03}$	± 0.01	$^{+0.29}_{-0.24}$
Continuum	± 0.01	$^{+0.01}_{-0.02}$	$^{+0.17}_{-0.07}$	$^{+0.32}_{-0.08}$
Rare	± 0.02	± 0.05	± 0.35	± 0.38
Sec. Leptons	± 0.20	$^{+0.14}_{-0.11}$	$^{+0.39}_{-0.42}$	$^{+0.20}_{-0.22}$
Fake leptons	$^{+0.00}_{-0.01}$	$^{+0.04}_{-0.01}$	$^{+0.02}_{-0.03}$	$^{+0.16}_{-0.21}$
ℓID	± 1.92	± 2.01	± 1.94	± 1.93
K/πID	± 0.93	± 1.86	± 0.83	± 0.89
π^0 reconstruction	± 0.00	± 0.00	± 2.50	± 2.50
$\mathcal{B}(D^0 \to X)$	± 1.03	± 1.76	± 3.52	± 4.08
Fit procedure	± 0.02	± 0.65	± 0.08	± 0.15
Tag efficiency	± 4.20	± 4.20	± 4.20	± 4.20
$N_{B\bar{B}}$	± 1.40	± 1.40	± 1.40	± 1.40
Track finding	± 1.05	± 1.05	± 1.05	± 1.05
Total	± 5.15	± 5.65	$^{+6.68}_{-6.69}$	$^{+7.02}_{-7.04}$

Table 8.5: Relative systematic uncertainties in percent for the branching ratio calculation for four D^0 decay channels used as control samples.

8.2 Search for resonance structures in the dipion mass



(a) $\cos\theta_\ell$ (b) $\cos\theta_\pi$ (c) χ



(d) $M_{\pi^+\pi^-}$ (e) q^2

Figure 8.4: Background-subtracted data distributions of the phase space variables describing the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay projected in the signal region ($M^2_{\text{miss}} < 0.8$ GeV^2): (a) the cosine of the of the polar angle θ_ℓ of the charged lepton ℓ^+ in the dilepton rest frame, (b) the cosine of the of the polar angle θ_π of the π^+ in the dipion rest frame, (c) the azimuthal angle χ between the dilepton and dipion decay planes, (d) the dipion mass $M_{\pi^+\pi^-}$ and (e) the momentum transfer q^2.

The ρ^0 resonance dominates the dipion invariant mass in the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ channel. Ideally, to disentangle this resonance from other contributions, an angular analysis has to be carried out. Such an analysis would require the five kinematic variables that describe the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay, namely, $M_{\pi^+\pi^-}$, q^2, and the angles θ_π, θ_ℓ and χ. The distributions of these variables, using the fit results of the one dimensional configuration, after background subtraction and projected in the signal region ($M^2_{\text{miss}} < 0.8$ GeV^2) are presented in Fig. 8.4. Note the limited statistics available, especially in the dipion mass region above 1 GeV, which makes the implementation of an angular analysis unfeasible. This section attempts to determine the ρ^0 content in the dipion mass using three alternative methods. These methods employ only the dipion mass assuming different line-shape distributions for the hypothesized particles that make up the $\pi^+\pi^-$ invariant mass spectrum. Furthermore, it also provides the results into two regions: first, the whole phase space region using the fit results from the one dimensional configuration, and the region defined by $q^2 < 8$ GeV^2 and $M_{\pi^+\pi^-} < 1.5$ GeV using the fit results from the two dimensional configuration. The general procedure is to perform an unbinned extended maximum likelihood fit to the dipion mass in the signal region, using different mass line-shapes for the possible $\pi^+\pi^-$ states and fixing the background

according to the fit results from the one (or two) dimensional configuration.

8.2.1 Study of particle content in the whole phase space

The background-subtracted data distribution of the dipion mass (see Fig. 8.4(d)) shows a large peak at $M_{\pi^+\pi^-} \approx 0.8$ GeV, a small bump at $M_{\pi^+\pi^-} \approx 1.3$ GeV and an excess near the $\pi^+\pi^-$ mass threshold. Based on this observation, any model of the dipion mass starts with a minimum of two resonances, the ρ^0 and the $f_2(1270)$ mesons, and an assumed shape for the remaining contributions. The latter is presumably due to nonresonant states dominated by an S-wave (or scalar) component [45], that can also receive contributions, to a lesser extent, from other states such as the $\pi\pi$ P-wave (vector) component. The nonresonant component may then originate from interferences among the different $\pi\pi$ states. In the following, three cases for modeling these components are studied. In the first case, the mass line-shapes of the ρ^0 and the $f_2(1270)$ are taken from MC simulations, and the third component is modeled as a scalar form factor derived from dispersion theory. The second case uses an incoherent sum of possible resonances decaying to $\pi^+\pi^-$ final states, modeled with an analytical mass line-shape, which is a modified version of a Breit-Wigner function. Finally, the resonances in the second case are allowed to interfere in a coherent sum.

Mass line-shapes for the ρ^0 and the $f_2(1270)$ resonances taken from MC and a scalar form factor

In the MC, the $B^+ \to \rho^0 \ell^+ \nu$ and $B^+ \to f_2(1270)\ell^+\nu$ decays are simulated using LCSR form factors from the references [98] and [99], respectively. The PDFs for these resonances are modeled in the unbinned fit using the RooKeysPdf module from the RooFit framework [117]. The ρ^0 resonance corresponds to a P-$\pi^+\pi^-$ state, also referred to as vector component, which is characterized by an angular momentum (J) and parity (P) set of $J^P = 1^-$. The $f_2(1270)$ resonance is a D-wave $\pi^+\pi^-$ state (tensor meson) with quantum numbers $J^P = 2^+$. The decay rate for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ process is defined in Eq. 2.44, with the hadronic matrix element expanded in terms of S- and P-wave contributions given in Eq. 2.45. The scalar (S-wave $J^P = 0^+$) form factor $f_0(M_{\pi^+\pi^-}, q^2)$ can be calculated in a model-independent way through dispersion theory, using the accurately known $\pi\pi$ phase shifts and a coupled channel treatment with the kaon-antikaon system [118]. Therefore, this approach removes the necessity of specifying any possible resonance or additional nonresonant states. In this formalism, the pion and kaon scalar form factors[2] are intertwined through the relation:

$$\begin{bmatrix} \Gamma_\pi(s) \\ \frac{2}{\sqrt{3}}\Gamma_K(s) \end{bmatrix} = \begin{bmatrix} \Omega_{11}(s) & \Omega_{12}(s) \\ \Omega_{21}(s) & \Omega_{22}(s) \end{bmatrix} \begin{bmatrix} \Gamma_\pi(0) \\ \frac{2}{\sqrt{3}}\Gamma_K(0) \end{bmatrix}, \tag{8.9}$$

where the matrix Ω is known as the Omnès matrix, and in the case of B^+ and B^0 mesons, the form factor normalizations are $\Gamma_\pi(0) = 0.984 \pm 0.006$ and $\Gamma_K(0) = 0.5 \pm 0.1$ [118]. This coupled channel formalism is applied as the scalar mesons can decay into a pair of pseudo-scalar mesons, such as $\pi\pi$, KK or $\eta\eta$, and some states can also decay to a four-pion channel. For instance, the scalar meson $f_0(980)$ can only decay into $\pi\pi$ and KK channel, and is produced near the mass threshold of the latter, and consequently affecting the phase of the scalar pion form factor. The pion scalar form factors in Eq. 8.9 can only be solved numerically[3] and its shape together with the impact of the uncertainties of $\Gamma_\pi(0)$ and $\Gamma_K(0)$ is shown in Fig. 8.5. Strictly speaking, the pion scalar form factor is valid up to $M_{\pi^+\pi^-} \approx 1.2$ GeV, since

[2] The pion (kaon) scalar form factor describes the B meson decay into a pair of pions (kaons) in a S-wave state.

[3] The numerical results of the scalar form factors for the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ used in this analysis were provided by Prof. Dr. Christoph Hanhart and Stefan Ropertz from the University of Bonn.

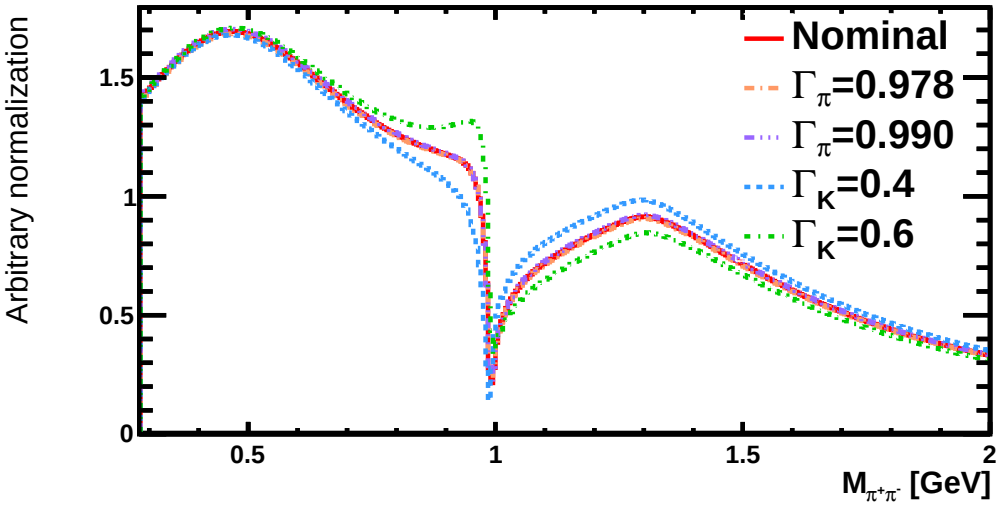


Figure 8.5: The pion scalar form factor describing the S-wave contribution of the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay as obtained from dispersion theory using the $\pi\pi$ phase shift and a kaon-antikaon coupled channel. The effect on the shape due to the uncertainties on the form factor normalizations $\Gamma_\pi(0)$ and $\Gamma_K(0)$ are also shown. This form factor is only valid up to $M_{\pi^+\pi^-} \approx 1.2$ GeV, but are used here as an approximation to model the S-wave $\pi^+\pi^-$ component.

X	Fit fraction [%]	Event yield	$\mathcal{B}(B^+ \to X\ell^+\nu_\ell) \times \mathcal{B}(X \to \pi^+\pi^-)$ [10^{-5}]	$Z[\sigma]$
ρ^0	68.4±9.1	209.8±19.2	$17.1 \pm 1.6_{\text{stat}} \pm 1.0_{\text{syst}}\,^{+0.1}_{-0.0\ \text{sff}}$	13.7
$f_2(1270)$	7.1±3.6	21.7±10.8	$1.8 \pm 0.9_{\text{stat}} \pm 0.1_{\text{syst}}\,^{+0.2}_{-0.1\ \text{sff}}$	2.2
Scalars	24.5±6.9	75.2±19.9	$6.1 \pm 1.6_{\text{stat}} \pm 0.4_{\text{syst}}\,^{+0.1}_{-0.0\ \text{sff}}$	4.4

Table 8.6: Fit fractions, event yields, the product of branching ratios $\mathcal{B}(B^+ \to X\ell^+\nu_\ell) \times \mathcal{B}(X \to \pi^+\pi^-)$ and statistical significance (without systematic uncertainties) for the $\pi^+\pi^-$ state X using shapes for the ρ^0 and $f_2(1270)$ mesons from MC and a scalar from factors derived from dispersion theory. The uncertainties in the product of branching ratios are statistical (stat), systematical (syst) and those due to the shape of the scalar form factors (sff), respectively. The signal reconstruction efficiency is assumed to be the same for all the $\pi^+\pi^-$ states, and calculated to be $\epsilon = (7.73 \pm 0.09) \times 10^{-4}$ using the PHSP $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ MC, as in the main measurement, in the signal region and fitted dipion mass.

other channels can also contribute to its shape. However, in this section, the scalar form factor is used up to $M_{\pi^+\pi^-} \approx 2.0$ GeV as an approximation to model the remaining contribution in the dipion mass. An additional systematic is calculated in this case to account for the variation in shape due to $\Gamma_\pi(0)$ and $\Gamma_K(0)$.

The contributions of the ρ^0 and $f_2(1270)$ mesons, as well as the scalar $\pi^+\pi^-$ are shown in Fig. 8.6. The corresponding event yields and branching ratios are presented in Tab. 8.6. According to these results, the ρ^0 resonance make up a $(68.4 \pm 9.1)\%$ of the $\pi^+\pi^-$ states in the semileptonic $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay. As the $\mathcal{B}(\rho^0 \to \pi^+\pi^-) = 100\%$, the $\mathcal{B}(B^+ \to \rho^0\ell^+\nu_\ell) = (17.1 \pm 1.6_{\text{stat}} \pm 1.0_{\text{syst}}) \times 10^{-5}$. The product of branching ratios can also be calculated by using

$$\mathcal{B}(B^+ \to X\ell^+\nu_\ell) \times \mathcal{B}(X \to \pi^+\pi^-) = \frac{f_{\text{fit}}}{\kappa}\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell), \tag{8.10}$$

Resonance	m_0 [GeV]	Γ_0 [GeV]	J^P	Line-shape	Source
$f_0(500)$	513±32	335 ± 67	0^+	Bugg	CLEO [119]
ρ^0	775.26±0.25	147.8 ± 0.9	1^-	GS	PDG [1]
$f_0(980)$	965±10	NA	0^+	Flatté	BES II [120]
$f_2(1270)$	1275.5±0.8	$186.7^{+2.2}_{2.5}$	2^+	RBW	PDG [1]
$\rho^0(1450)$	1493±15	427 ± 31	1^-	GS	BaBar [121]
$\rho^0(1700)$	1861±17	316 ± 26	1^-	GS	BaBar [121]

Table 8.7: Properties of the candidates for $\pi^+\pi^-$ final state, line-shapes that describe them and the source were they are taken from. The mass and decay width of the $f_0(500)$ do not enter as input parameters in the Bugg line-shape, they are just shown to give a sense of its position and large width in the dipion mass spectrum. The decay width of the $f_0(980)$ is not required by the Flatté parametrization; however, different experiments report decay widths ranging from 10 MeV to 100 MeV.

where f_{fit} is the fraction of the fit for a given resonance or nonresonant component in the dipion mass and κ is a factor that accounts for a change in efficiency due to the use of the signal region and dipion masses below 2.0 GeV in the unbinned fit. This equation also implies that the systematic uncertainties in the measurement of $\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$ can be propagated into the measurements of the product $\mathcal{B}(B^+ \to X\ell^+\nu_\ell) \times \mathcal{B}(X \to \pi^+\pi^-)$. Instead of determining κ, the latter is calculated using the fitted event yields and the signal reconstruction efficiency in the signal region and for dipion masses below 2.0 GeV, $\epsilon = (7.73 \pm 0.09) \times 10^{-4}$, in the same way as in the main measurement. Equation 8.10 is only shown to clarify the calculation of the systematic uncertainties from the main measurement. The shape of the scalar component does not have a significant impact in the systematic uncertainties for the ρ component but affect those involving the $f_2(1270)$ meson.

Incoherent sum of resonances decaying to a charged pion pair

In this case, instead of using the scalar form factor and shapes from MC simulation for the ρ and the $f_2(1270)$ resonances, one can assume contributions from explicit particles decaying into the $\pi^+\pi^-$ final state. Each particle is modeled with an analytical function to describe its mass line-shape, either with a Breit-Wigner function or a modified version of it. A list of these particles with their respective masses and decay widths is provided in Tab. 8.7. This list is not exhaustive as the majority of these particles are not well established states and the current statistics does not allow to use them all in the fit. In some cases, the values of the width and mass are taken from an experiment rather than the PDG, as there is not a consensus for their average. Dependent on the angular momentum of the particle, how broad its decay width is, or because it is produced near the $K\bar{K}$ mass threshold, different mass line-shapes are used. A list of these functions is given in the following:

1. **Relativistic Breit-Wigner (RWB)**. It is the standard function to describe the mass line-shape of a resonance, however, it fails to describe particles with a broad decay width such as the $f_0(500)$ and the ρ-family mesons. In this thesis, it is used to model the $f_2(1270)$ meson and given by [122, 123]

$$R(m) = \frac{m_0^2}{m_0^2 - m^2 - \imath m_0 \Gamma(m)}, \tag{8.11}$$

where

$$\Gamma(m) = \Gamma_0 \frac{m_0}{m} \left[\frac{k(m)}{k(m_0)} \right]^{2J+1} F(z, z_0),$$

$$F^{J=0}(z, z_0) = 1, \qquad F^{J=1}(z, z_0) = \sqrt{\frac{1+z_0}{1+z}}, \qquad F^{J=2}(z, z_0) = \sqrt{\frac{(z_0-3)^2 + 9z_0}{(z-3)^2 + 9z}},$$

$$k(m) = \frac{m}{2}\sqrt{1 - 4\frac{m_\pi^2}{m^2}}, \qquad z = r_{BW} k(m).$$

Here, m_0 and Γ_0 are the nominal mass and decay width of the resonance, J is the angular momentum of the particle, m_π is the mass of the π^+ meson and r_{BW} is the mean radius of the meson, taken as $r_{BW} = 1.5\ \text{GeV}^{-1}$ (≈ 0.3 fm) for the $f_2(1270)$ [124].

2. **Gounaris-Sakurai (GS)** [125]. This line-shape is used to describe vector mesons such as ρ^0, $\rho^0(1450)$ and $\rho^0(1700)$. It is defined as:

$$R(m) = \frac{m_0^2(1 + dm_0\Gamma_0/m_0)}{m_0^2 - m^2 + f(m) - \imath m_0 \Gamma(m)}, \tag{8.12}$$

where

$$f(m) = \Gamma_0 \frac{m_0^2}{k^3(m_0)} \left[k^2(m)[h(m) - h(m_0)] + (m_0^2 - m^2)k^2(m_0) \frac{dh}{dm^2}\Big|_{m^2=m_0^2} \right],$$

$$h(m) = \frac{2k(m)}{\pi m} \ln\left(\frac{m + 2k(m)}{2m_\pi} \right),$$

$$\frac{dh}{dm^2}\Big|_{m^2=m_0^2} = h(m_0^2)\left[\frac{1}{8k^2(m_0) - \frac{1}{2m_0^2}} \right] + \frac{1}{2\pi m_0^2},$$

$$d = \frac{3m_\pi^2}{\pi k^2(m_0)} \ln\left(\frac{m_0 + 2k(m_0)}{2m_\pi} \right) + \frac{m_0}{2\pi k(m_0)} - \frac{m_\pi^2 m_0}{\pi k^3(m_0)},$$

$$\Gamma(m) = \Gamma_0 \frac{m_0}{m} \left[\frac{k(m)}{k(m_0)} \right]^{2J+1}.$$

3. **Flatté** [126]. This parametrization is used to model the $f_0(980)$ line-shape. This resonance is produced near the mass threshold of K^+K^- and can also decay to this final state, as such the Flatté function incorporates a coupled channel treatment for this meson with the following form:

$$R(m) = \frac{m_0^2}{m_0^2 - m^2 - \imath m_0(\rho_{\pi\pi} g_1 + \rho_{KK} g_2)}, \tag{8.13}$$

where

$$\rho_{\pi\pi} = \frac{2}{3}\sqrt{1 - \frac{4m_{\pi^\pm}^2}{m^2}} + \frac{1}{3}\sqrt{1 - \frac{4m_{\pi^0}^2}{m^2}},$$
$$\rho_{KK} = \frac{1}{2}\sqrt{1 - \frac{4m_{K^\pm}^2}{m^2}} + \frac{1}{2}\sqrt{1 - \frac{4m_{K^0}^2}{m^2}}.$$

Here, the parameters g_1, g_2 and m_r are taken from reference [120] with the following values: $m_r = 965 \pm 10$ MeV, $g_1 = 165 \pm 18$ MeV and $g_2 = 695 \pm 53$ MeV.

4. **Bugg** [127]. This function is used to describe the broad resonance $f_0(500)$ that is dominant at the low dipion mass, it is defined as:

$$R(m) = \frac{M^2}{M^2 - m^2 - g_1^2 \frac{m^2 - m_A^2}{M^2 - m_A^2}\left[j_1(m^2) - j_1(M^2)\right] - \imath M \sum_{i=1}^{4} \Gamma_i(m^2)}, \tag{8.14}$$

where

$$j_1(m^2) = \frac{1}{\pi}\left[2 + \rho_1 \ln\left(\frac{1-\rho_1}{1+\rho_1}\right)\right],$$
$$g_1^2(m^2) = M\left(b_1 + b_2(m^2)\right)\exp\left(-(m^2 - M^2)/A\right),$$
$$M\Gamma_1(m^2) = g_1^2(m^2)\frac{m^2 - m_A^2}{M^2 - m_A^2}\rho_1(m^2),$$
$$M\Gamma_2(m^2) = 0.6 g_1^2(m^2)\frac{m^2}{M^2}\exp\left(-\alpha\left|m^2 - 4m_K^2\right|\right)\rho_2(m^2),$$
$$M\Gamma_3(m^2) = 0.2 g_1^2(m^2)\frac{m^2}{M^2}\exp\left(-\alpha\left|m^2 - 4m_\eta^2\right|\right)\rho_3(m^2),$$
$$M\Gamma_4(m^2) = M g_{4\pi}\frac{\rho_{4\pi}(m^2)}{\rho_{4\pi}(M^2)}$$
$$\rho_{4\pi} = \frac{1}{1 + \exp\left(7.082 - 2.845m^2\right)},$$
$$\rho_i(m^2) = \sqrt{1 - 4\frac{m_i^2}{m^2}} \quad \text{with 1,2 and 3} = \pi,\ K \text{ and } \eta.$$

The parameters are taken from reference [127] and fixed to $m_r = 0.953$, $S_A = 0.41 m_\pi^2$, $b_1 = 1.302$ GeV^2, $b_2 = 0.340$, $A = 2.426$ GeV^2 and $g_{4\pi} = 0.011$ GeV.

The total PDF for the resonant states is taken as the sum of the magnitude squared of each line-shape. An additional systematic is calculated taking into account the uncertainties on the input parameters for each mass line-shape function. Different options were tested to model the nonresonant contribution, as there is not an analytical function available for the $\pi^+\pi^-$ state. These options include a constant, a linear, a polynomial or an exponential function. The shape of the PHSP $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ MC was also tested as

X	Fit fraction [%]	Event yield	$\mathcal{B}(B^+ \to X\ell^+\nu_\ell) \times \mathcal{B}(X \to \pi^+\pi^-)$ $[10^{-5}]$	$Z[\sigma]$
$f_0(500)$	5.3±3.7	16.1±11.2	$1.3 \pm 0.9_{\text{stat}} \pm 0.1_{\text{syst}} \pm 0.1_{\text{ls}}$	1.5
ρ^0	70.8±10.6	216.0±20.0	$17.6 \pm 1.6_{\text{stat}} \pm 1.1_{\text{syst}} \pm 0.7_{\text{ls}}$	14.9
$f_0(980)$	6.9±4.0	21.0±11.8	$1.7 \pm 1.0_{\text{stat}} \pm 0.1_{\text{syst}} \pm 0.2_{\text{ls}}$	2.0
$f_2(1270)$	11.1±4.8	34.0±14.0	$2.8 \pm 1.1_{\text{stat}} \pm 0.2_{\text{syst}}\ ^{+0.4}_{-0.7\ \text{ls}}$	2.6
$\rho^0(1450)$	2.4±6.1	7.4±18.5	$0.6 \pm 1.5_{\text{stat}} \pm 0.0_{\text{syst}}\ ^{+0.1}_{-0.0\ \text{ls}}$	0.4
$\rho^0(1700)$	3.4±3.3	10.5±10.0	$0.9 \pm 0.8_{\text{stat}} \pm 0.1_{\text{syst}} \pm 0.1_{\text{ls}}$	1.2

Table 8.8: Fit fractions, event yields, the product of branching ratios $\mathcal{B}(B^+ \to X\ell^+\nu_\ell) \times \mathcal{B}(X \to \pi^+\pi^-)$ and statistical significance (without systematic uncertainties) for the $\pi^+\pi^-$ state X using an incoherent sum of resonances. The uncertainties in the product of branching ratios are statistical (stat), systematical (stat) and those due to the variation of the parameters of the line-shape functions within their uncertainties, respectively. The signal reconstruction efficiency is assumed to be the same for all the $\pi^+\pi^-$ states, and calculated to be $\epsilon = (7.73 \pm 0.09) \times 10^{-4}$ using the PHSP $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ MC, as in the main measurement, in the signal region and fitted dipion mass.

a nonresonant model. These models did not affect the results from the incoherent sum of resonances[4], and consequently, are ignored for the rest of the analysis.

The resonance content in the dipion mass, as well as their corresponding product of branching ratios is given in Tab. 8.8. In this case, the ρ^0 meson make up a (70.8±10.6)% of the total $\pi^+\pi^-$ states and the branching ratio is measured to be $\mathcal{B}(B^+ \to \rho^0\ell^+\nu_\ell) = (17.6 \pm 1.6_{\text{stat}} \pm 1.3_{\text{syst}}) \times 10^{-5}$. The measured product of branching ratios for the other resonances is not statistically significant. The functional shape for each resonance is shown in Fig. 8.7.

Coherent sum of resonances decaying to a charged pion pair

In this case the resonances are allowed to interfere by the introduction of a phase in the mass line-shape. The total PDF of the $\pi^+\pi^-$ states is given by

$$PDF(m) = \left| \sum_i c_i R_i(m) e^{\imath\theta_i} \right|^2, \tag{8.15}$$

where c_i is a positive real number that accounts for the contribution of each resonance and is referred to as the amplitude in this context, and $\theta_i \in [-\pi, \pi]$ is the phase. The Fit fraction for the resonance i is calculated as

$$f^i_{\text{fit}} = \frac{\int \left| c_i R(m) e^{\imath\theta_i} \right|^2 dm}{\int \left| \sum_i c_i R(m) e^{\imath\theta_i} \right|^2 dm}, \tag{8.16}$$

and the interference between resonances i and j is given by

$$f^{ij}_{\text{fit}} = \frac{\int 2\Re \left[c_i c_j^* R_i(m) R_j^*(m) e^{\imath(\theta_i - \theta_j)} \right] dm}{\int \left| \sum_i c_i R(m) e^{\imath\theta_i} \right|^2 dm}. \tag{8.17}$$

[4] The event yield with nonresonant models were consistent with zero in the fit.

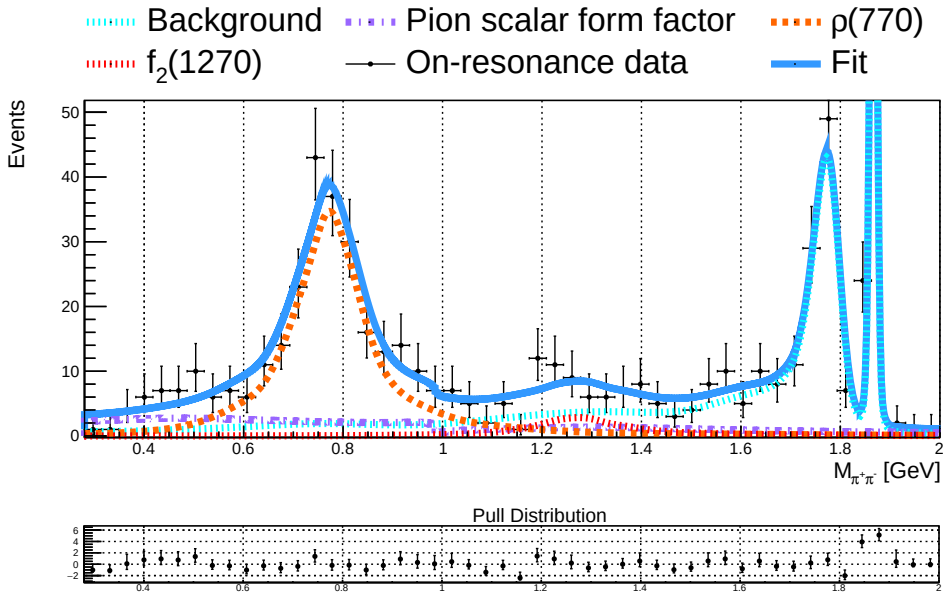
X	Sc1	Sc2	Sc4	Sc4	Sc5
$f_0(500)$	✓	✓	✓	✓	✓
ρ^0	✓	✓	✓	✓	✓
$f_0(980)$	✗	✓	✗	✗	✓
$f_2(1270)$	✓	✓	✓	✓	✓
$\rho^0(1450)$	✗	✗	✓	✓	✓
$\rho^0(1700)$	✗	✗	✗	✓	✓

Table 8.9: Scenarios to combine the mass line-shapes of different resonance candidates in a coherent way.

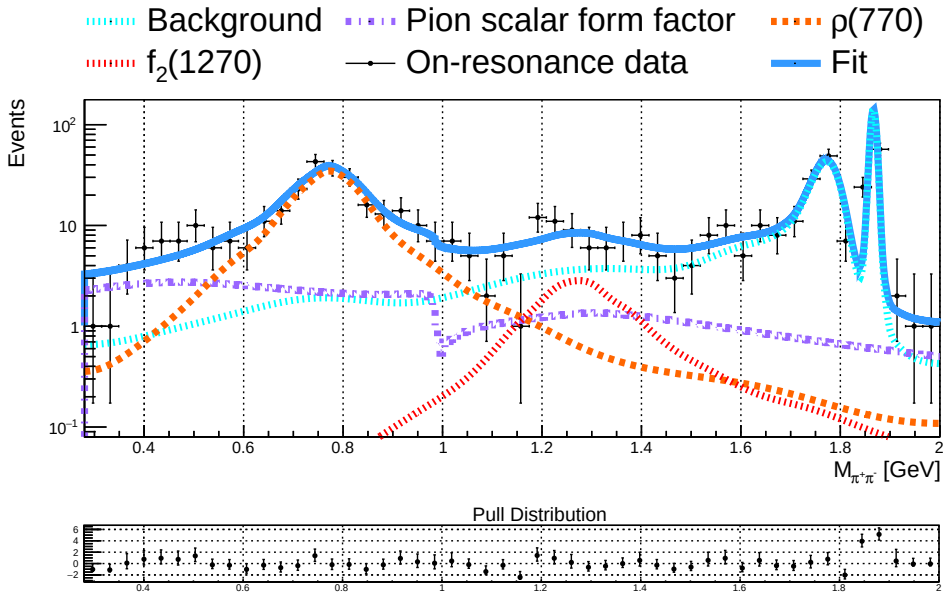
The inclusion of the phase for each resonance increases the number of free parameters and consequently the statistical error in the fit fraction, which is larger compared to the previous cases. Since the dominant resonance in the dipion mass is the ρ^0 meson, the values for its coefficient and phase are fixed to one and zero, respectively.

Five scenarios are considered with different combinations of possible $\pi^+\pi^-$ states with three common resonances: $f_0(500)$, ρ^0 and $f_2(1270)$. These scenarios are summarized in Tab. 8.9. The fit results containing the amplitude, phase and fit fraction for each resonance in the different scenarios is given in 8.10. The corresponding shape of the total contribution to the line-shape from a coherent sum of resonances using the fifth scenario is visualized in Fig. 8.8. The statistical uncertainties in the amplitude, phase and total (coherent) yield are propagated into the fit fraction, leading to asymmetrical uncertainties. As these uncertainties are large, no additional systematic uncertainties are considered.

The calculated product of branching ratios $\mathcal{B}(B^+ \to X\ell^+\nu_\ell) \times \mathcal{B}(X \to \pi^+\pi^-)$ for each resonance in the considered scenarios are presented in Tab. 8.11. Despite the fluctuation in these values from one scenario to the other, they agree within error bars. For instance, the branching ratio of semileptonic $B^\pm$ to the ρ meson varies from $\mathcal{B}(B^+ \to \rho^0\ell^+\nu_\ell) = (14.8^{+4.7}_{-4.1\,\mathrm{stat}} \pm 0.9_{\mathrm{syst}}) \times 10^{-5}$ in the fifth scenario to $\mathcal{B}(B^+ \to \rho^0\ell^+\nu_\ell) = (17.5^{+4.3}_{-5.8\,\mathrm{stat}} \pm 1.0_{\mathrm{syst}}) \times 10^{-5}$ in the second scenario. One can argue that the difference in the central values is due to the low statistics in the high dipion mass region (above 1.4 GeV), and consequently resonances in this region, such as $\rho^0(1450)$ and $\rho^0(1700)$, can be neglected in the modeling of contributing $\pi^+\pi^-$ states. On the other hand, as the study of resonances is carried out using only the dipion mass without angular variables, the interference effects can be ignored. Thus, the previous two cases are enough to study the resonance component of the dipion mass. Overall, the three cases studied, provide consistent results for the ρ^0 channel. Apart from the ρ^0 meson, the other states are measured with a small statistical significance, indicating that either a larger data sample is needed to confirm these decays with a hadronic tag B meson reconstruction or another analysis must be carried out using another B meson reconstruction method such as the untagged technique.



(a) Linear scale



(b) Logarithmic scale

Figure 8.6: Contributions of the ρ^0 and $f_2(1270)$ mesons, as well as the scalar $\pi^+\pi^-$ states after performing an unbinned extended maximum likelihood fit to the dipion mass in the signal region ($M^2_{\text{miss}} < 0.8\ \text{GeV}^2$). The event yields of the background are fixed according to the fit results in the one dimensional configuration. The distributions in the dipion mass are presented in a linear (a) and logarithmic (b) scale to appreciate the mass line-shape of each contribution.

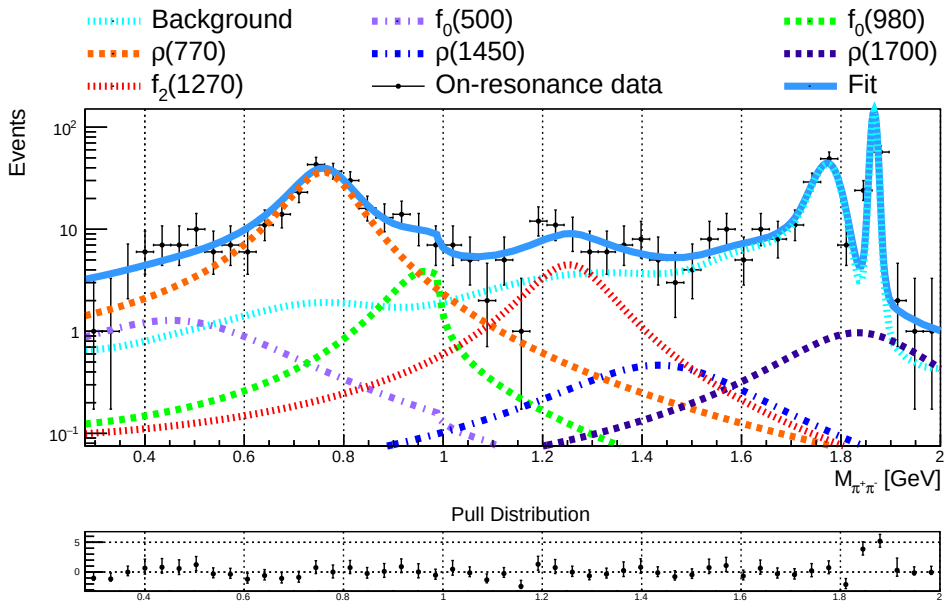


Figure 8.7: Mass line-shapes for possible particles decaying to a $\pi^+\pi^-$ final state after performing an unbinned extended maximum likelihood fit to the dipion mass in the signal region ($M^2_{\text{miss}} < 0.8$ GeV^2) using a incoherent sum of resonances. The event yields of the background are fixed according to the fit results in the one dimensional configuration.

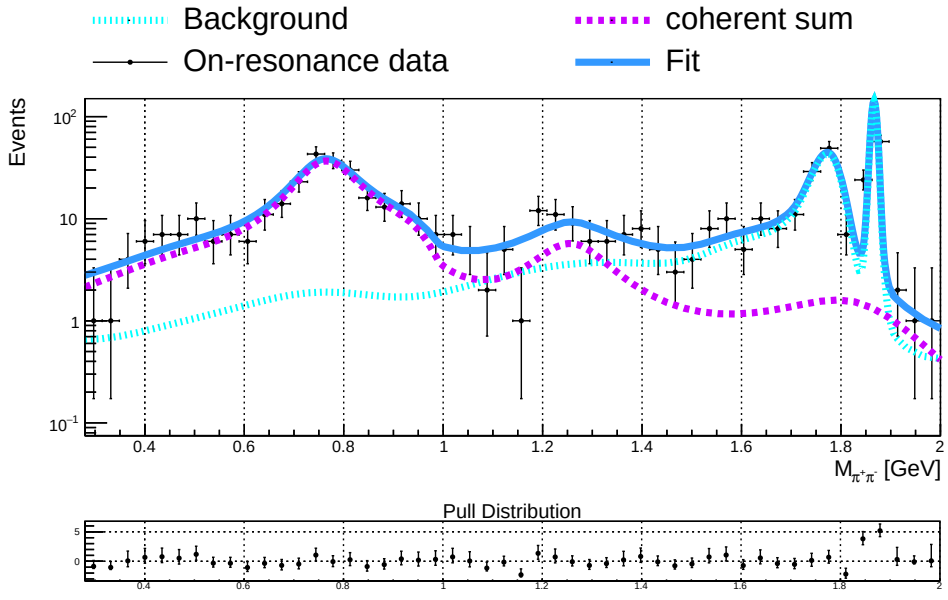


Figure 8.8: Results from using scenario 5 after performing an unbinned extended maximum likelihood fit to the dipion mass in the signal region ($M^2_{\text{miss}} < 0.8$ GeV^2) using a coherent sum of resonances. The event yields of the background are fixed according to the fit results in the one dimensional configuration.

	Sc1			Sc2			Sc3		
	c	θ [rad]	f_{fit} [%]	c	θ [rad]	f_{fit} [%]	c	θ [rad]	f_{fit} [%]
$f_0(500)$	0.73±0.28	−1.08±0.33	$7.5^{+5.6}_{-4.6}$	0.48±0.31	−0.55±1.00	$3.4^{+4.8}_{-3.1}$	0.62±0.30	−0.86±0.65	$4.8^{+4.9}_{-3.7}$
ρ^0	1.00	0.00	$69.4^{+12.3}_{-9.1}$	1.00	0.00	$72.5^{+17.8}_{-24.0}$	1.00	0.00	$61.9^{+16.9}_{-20.7}$
$f_0(980)$	–	–	-	0.22±0.15	−1.80±0.77	$3.6^{+7.7}_{-3.5}$	–	–	-
$f_2(1270)$	0.23±0.05	−1.20±0.45	$6.8^{+2.7}_{-2.6}$	0.30±0.08	−0.97±0.50	$12.0^{+5.9}_{-6.1}$	0.16±0.10	−1.32±0.70	$2.9^{+3.9}_{-2.6}$
$\rho^0(1450)$	–	–	-	–	–	-	0.15±0.16	−1.75±1.44	$1.8^{+4.6}_{-1.8}$
$\rho^0(1700)$	–	–	-	–	–	-	–	–	-
Sum of interferences	–	–	$18.3^{+7.0}_{-10.6}$	–	–	$8.5^{+23.0}_{-28.0}$	–	–	$7.3^{+57.5}_{-35.6}$
Coherent yield		295.5±19.4			295.6±19.5			300.2±20.4	

	Sc4			Sc5		
	c	θ [rad]	f_{fit} [%]	c	θ [rad]	f_{fit} [%]
$f_0(500)$	0.63±0.26	−0.49±0.59	$4.8^{+4.1}_{-3.1}$	0.72±0.37	0.42±0.88	$6.2^{+6.2}_{-4.7}$
ρ^0	1.00	0.00	$59.6^{+17.7}_{-11.6}$	1.00	0.00	$59.3^{+18.7}_{-16.3}$
$f_0(980)$	–	–	-	0.27±0.18	−2.30±0.59	$4.5^{+9.1}_{-4.2}$
$f_2(1270)$	0.18±0.06	−1.45±0.58	$3.6^{+2.5}_{-2.0}$	0.27±0.09	−1.41±0.58	$8.0^{+5.3}_{-4.5}$
$\rho^0(1450)$	0.00±0.12	−0.55±4.32	0.0	0.00±0.35	−0.72±4.26	0.0
$\rho^0(1700)$	0.20±0.08	−3.14±1.16	$5.1^{+3.9}_{-3.5}$	0.20±0.07	−3.14±1.20	$5.1^{+4.1}_{-3.2}$
Sum of interferences	–	–	$7.0^{+3.2}_{-6.6}$	–	–	$17.0^{+21.3}_{-24.7}$
Coherent yield		308.7±19.2			306.9±19.5	

Table 8.10: Amplitude, phase and fraction of the fit for each resonance for the five different scenarios considered in the coherent sum of mass line-shapes. The total event yield for the coherent sum of resonances is also provided as well as the total fraction of the interferences. The amplitude (c) and phase for the ρ meson are fixed to one and zero, respectively.

X	Sc1	Sc2	Sc4	Sc4	Sc5
$f_0(500)$	$1.8\,^{+1.3}_{-1.1} \pm 0.05$	$0.8\,^{+1.2}_{-0.7} \pm 0.1$	$1.2\,^{+1.2}_{-0.9} \pm 0.1$	$1.2\,^{+1.0}_{-0.8} \pm 0.1$	$1.6\,^{+1.6}_{-1.2} \pm 0.1$
ρ^0	$16.7\,^{+3.0}_{-2.2} \pm 1.0$	$17.5\,^{+4.3}_{-5.8} \pm 1.0$	$15.2\,^{+4.1}_{-5.1} \pm 0.9$	$15.0\,^{+4.5}_{-2.9} \pm 0.9$	$14.8\,^{+4.7}_{-4.1} \pm 0.9$
$f_0(980)$	-	$0.9\,^{+1.9}_{-0.8} \pm 0.1$	-	-	$1.1\,^{+2.3}_{-1.1} \pm 0.1$
$f_2(1270)$	$1.6\,^{+0.7}_{-0.6} \pm 0.1$	$2.9\,^{+1.4}_{-1.5} \pm 0.2$	$0.7\,^{+1.0}_{-0.6} \pm 0.04$	$0.9\,^{+0.6}_{-0.5} \pm 0.1$	$2.0\,^{+1.3}_{-1.1} \pm 0.1$
$\rho^0(1450)$	-	-	$0.4\,^{+1.1}_{-0.4} \pm 0.03$	0.0	0.0
$\rho^0(1700)$		-	-	$1.3\,^{+1.0}_{-0.9} \pm 0.1$	$1.3\,^{+1.0}_{-1.3} \pm 0.1$

Table 8.11: Product of branching ratios $\mathcal{B}(B^+ \to X\ell^+\nu_\ell) \times \mathcal{B}(X \to \pi^+\pi^-)$ $[10^{-5}]$ for each resonance in the different scenarios of the coherent sum case. The first uncertainty is statistical and the second is systematical.

X	Fit fraction [%]	Event yield	$\mathcal{B}(B^+ \to X\ell^+\nu_\ell) \times \mathcal{B}(X \to \pi^+\pi^-)$ $[10^{-5}]$
ρ^0	56.9±12.9	71.8±11.8	$6.0 \pm 1.0_{\text{stat}} \pm 0.3_{\text{syst}}\,^{+0.05}_{-0.007\ \text{sff}}$
$f_2(1270)$	15.7±7.3	19.8±8.6	$1.6 \pm 0.7_{\text{stat}} \pm 0.1_{\text{syst}}\,^{+0.05}_{-0.07\ \text{sff}}$
Scalars	27.4±11.5	34.5±13.4	$2.9 \pm 1.1_{\text{stat}} \pm 0.1_{\text{syst}}\,^{+0.17}_{-0.24\ \text{sff}}$

Table 8.12: Fit fractions, event yields and the product of branching ratios $\mathcal{B}(B^+ \to X\ell^+\nu_\ell) \times \mathcal{B}(X \to \pi^+\pi^-)$ for the $\pi^+\pi^-$ state X in the low q^2 region ($q^2 < 8$ GeV^2), using shapes for the ρ^0 and $f_2(1270)$ mesons from MC and a scalar from factors derived from dispersion theory. The uncertainties in the product of branching ratios are statistical (stat), systematical (stat) and those due to the shape of the scalar form factors (sff), respectively. The signal reconstruction efficiency is assumed to be the same for all the $\pi^+\pi^-$ states, and calculated to be $\epsilon = (7.60 \pm 0.12) \times 10^{-4}$ using the PHSP $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ MC, as in the main measurement, in the signal region and fitted dipion mass.

A comparison of previous measurements of the total branching ratio of semileptonic B decays to a ρ^0 meson, with the one obtained in the first case (the use of the scalar form factor) is shown in Fig. 8.9.

8.2.2 Study of dipion mass in the low momentum transfer region ($q^2 < 8$ GeV^2)

As pointed out in the previous section, the modeling of the dipion mass should be restricted to resonances below 1.4 GeV because of the limited statistics. Since the three cases studied for the whole phase space gave consistent results for the branching ratio of the $B^+ \to \rho^0\ell^+\nu_\ell$ channel, one can restrict the analysis of the $\pi^+\pi^-$ states in the low q^2 region to the use of the scalar form factor and shapes derived from MC simulation. This method has the advantage of reducing the statistical uncertainties of the event yields as the number of fit parameters is smaller compared to the other cases. The fit fraction, event yields and product of branching ratios for each resonance are presented in Tab. 8.12.

In the case of the ρ^0 meson, the partial branching ratio in the low q^2 is

$$\Delta\mathcal{B}(B^+ \to \rho^0\ell^+\nu_\ell)\,|_{q^2 < \text{GeV}^2} = (6.0 \pm 1.0_{\text{stat}} \pm 0.3_{\text{syst}}) \times 10^{-5}. \tag{8.18}$$

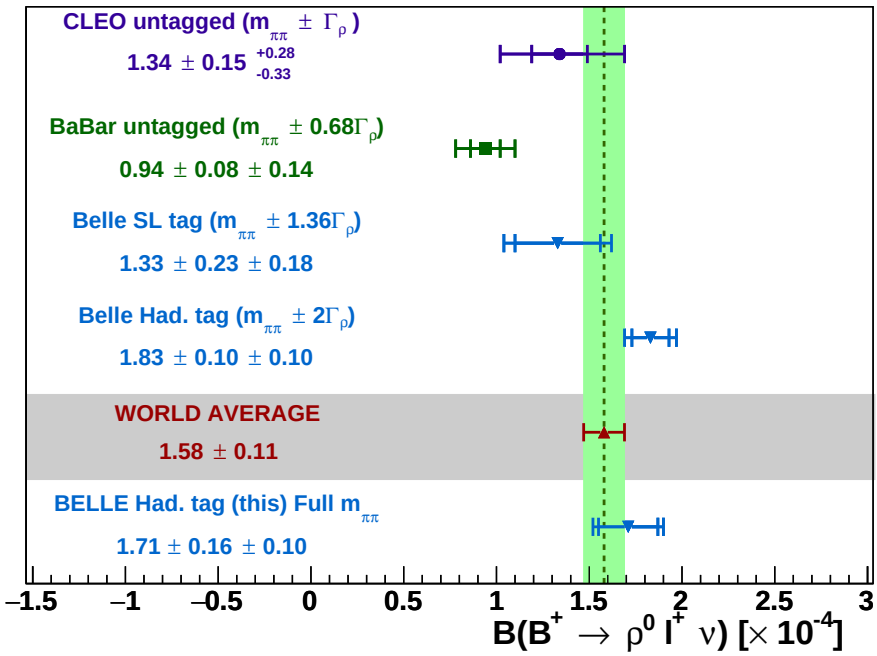


Figure 8.9: Previous measurements of $\mathcal{B}(B^+ \to \rho^0 \ell^+ \nu_\ell)$ used in the calculation of the world average [1]. They include an untagged measurement by CLEO [128], an untagged measurement by BaBar [43], a semileptonic tag measurement by Belle [129] and a hadronic tag measurement by Belle [47]. They are compared with the result from the first case studied using scalar form factor and shapes from MC. The result from this thesis treats the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ channel as a four body decay using the full phase space, whereas the reported measurements for the ρ^0 are optimized based on a three body decay using a mass window around the nominal mass of this meson. This mass window is specified in the plot for each measurement in terms of the ρ^0 decay width (Γ_ρ).

8.3 Implications for $|V_{ub}|$ from the $B^+ \to \rho^0 \ell^+ \nu_\ell$ decay

The magnitude of the CKM matrix element V_{ub} is related to the partial branching ratio $\Delta\mathcal{B}$ within a certain region of q^2 through

$$|V_{ub}| = \sqrt{\frac{c_\nu \Delta\mathcal{B}}{\tau_{B^+}\Delta\zeta}}, \tag{8.19}$$

where $C_\nu = 2$ for B^+ decay modes, τ_{B^+} is the lifetime of the B^+ meson, and $\Delta\zeta$ is the normalized decay rate calculated in the same q^2 region as $\Delta\mathcal{B}$ within a theoretical framework. Values of $\Delta\zeta$ are available for the $B^+ \to \rho^0\ell^+\nu_\ell$ decay from form factor predictions such as LCSR (low q^2), LQCD (high q^2) or quark models (some calculations cover the full q^2 range). To reduce the statistical uncertainty in the branching ratio calculation, the values obtained from a fit using the scalar form factors and the MC shapes for the decays are employed in this section. The predictions from the ISGW2 [39] and UKQCD [130][5] are used for the measurement in the full q^2 range, whereas the LCSR [108] is used for the measurement in the $q^2 < 8$ GeV^2. The theoretical uncertainties from the ISGW2 calculation are not available. The lifetime of the B^+ meson is taken as $\tau_{B^+} = 1.638 \pm 0.004$ ps [1]. The calculated values of $|V_{ub}|$ using the

[5] The UKQCD calculation uses a quenched lattice QCD and other assumptions to extrapolate $\Delta\zeta$ to the full q^2 range.

Theory	q^2 [GeV^2]	$\Delta\mathcal{B}$ [10^{-5}]	$\Delta\zeta$ ps^{-1}	$\lvert V_{ub}\rvert$ [10^{-3}]
UKQCD [130]	all	$17.1 \pm 1.6 \pm 1.0$	$16.5^{+3.5}_{-2.3}$	$3.56 \pm 0.17 \pm 0.10\ ^{+0.38}_{-0.25}$
ISGW2 [39]	all	$17.1 \pm 1.6 \pm 1.0$	14.1	$3.85 \pm 0.18 \pm 0.11$
LCSR [108]	$q^2 < 8$	$6.0 \pm 1.0 \pm 0.3$	5.4 ± 0.4	$3.68 \pm 0.31 \pm 0.09 \pm 0.14$

Table 8.13: Calculated values of $|V_{ub}|$ for different predictions of form factors for the $B^+ \to \rho^0 \ell^+ \nu_\ell$ decay, using the results from a fit to MC shapes and scalar form factors. The uncertainties are statistical, systematical and theoretical, respectively.

aforementioned predictions are provided in Tab. 8.13. These results are consistent at the 0.6σ level. Note that the theoretical uncertainty for the UKQCD calculation is larger than the statistical uncertainty in the full q^2 range. On the other hand, the result at low q^2 is dominated by the statistical uncertainty due to the small sample size at this region.

Using the UKQCD calculation for $\Delta\zeta$ and previous measurements of $\mathcal{B}(B^+ \to \rho^0 \ell^+ \nu_\ell)$, one can compare the central values for the $|V_{ub}|$. The values for this matrix element for each experiment are lower compared with its exclusive value ($|V_{ub}| = (3.65 \pm 0.14) \times 10^{-3}$) [41]. The latter is currently the most precise measurement of $|V_{ub}|$ derived from global fits to the $B \to \pi \ell \nu$ channel. In the literature [44–46], this difference in central values has been associated to the use of a limited dipion mass region around the nominal ρ^0 mass, suggesting access to a small fraction of the phase space available to the decay. This argument would imply that the experimentalists did not correct their analyses for signal efficiencies. What truly happens is that previous analyses could not separate the S-wave/nonresonant contribution from the ρ resonance, due to the high levels of background contamination in untagged analyses. From the previous measurements, the most precise result for $|V_{ub}|$ comes from a hadronic tag measurement by Belle [47] using a dipion mass window around twice the decay width of the ρ^0 meson, for which the central value is consistent with the exclusive determination. The result from this thesis is also consistent with the exclusive value, and even though it uses a hadronic tag method, the statistical error is larger than that in the latter. This difference is mainly due to the treatment of the final state particles, in this thesis the $B^+ \to \pi^+ \pi^- \ell^+ \nu_\ell$ channel is considered as a four body decay, while the other cases used a three body decay approach. This combined with the restricted phase space give rise to differences in the signal selection and background level contamination. The exclusive results hold up the disagreement at the 3σ level with the inclusive value ($|V_{ub}| = (4.52 \pm 0.15_{\text{exp}} \pm 0.13_{\text{theo}}) \times 10^{-3}$) [41].

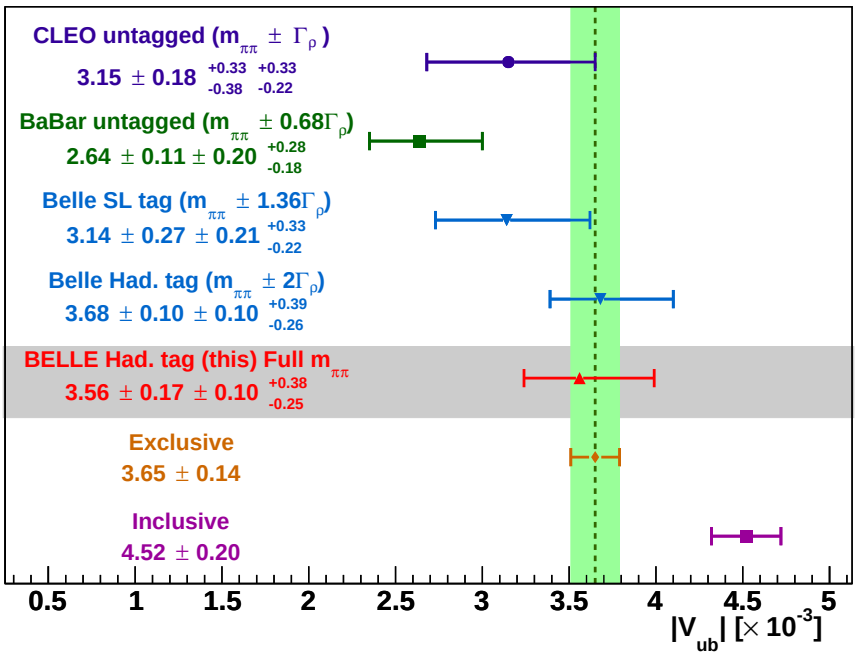


Figure 8.10: Calculated values of $|V_{ub}|$ using the UKQCD [130] normalized decay rate and previous measurements of $\mathcal{B}(B^+ \to \rho^0 \ell^+ \nu_\ell)$, together with the result from this thesis, in the full q^2 range. The branching ratio measurements include an untagged measurement by CLEO [128], an untagged measurement by BaBar [43], a semileptonic tag measurement by Belle [129] and a hadronic tag measurement by Belle [47]. They are compared with the exclusive, based on a global fit using the $B \to \pi \ell \nu_\ell$ channel, and the inclusive values of $|V_{ub}|$ as reported by HFLAV [41]. Note that the mass window used in the analysis to select ρ^0 candidates affects the $|V_{ub}|$ value, measurements that take into account the broad width of this meson (Γ_ρ) give $|V_{ub}|$ values in agreement with the exclusive value.

8.4 Final comments

The present analysis of the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ is by no means complete given the poor statistics obtained after signal selection, especially for dipion mass regions above 1 GeV. The hadronic tag method, implemented in this thesis, yields a signal reconstruction efficiency for this channel at the 7×10^{-4} level. This implies, that in order to study and measure relations between form factors and other resonance states as suggested in [48], another B meson reconstruction method has to be utilized, at the expense of large systematic uncertainties. For instance, the untagged (inclusive B reconstruction) method will provide an increased signal data sample to accomplish such studies.

In summary, a total branching ratio of the four-body decay $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ has been measured to be

$$\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell) = (22.4 \pm 1.9_{\text{stat}} \pm 1.3_{\text{syst}}) \times 10^{-5},$$

from which approximately a $(69 \pm 9)\%$ corresponds to the resonance state involving the ρ^0 vector meson. The branching ratio for the latter is measured to be

$$\mathcal{B}(B^+ \to \rho^0 \ell^+ \nu_\ell) = (17.1 \pm 1.6_{\text{stat}} \pm 1.0_{\text{syst}}) \times 10^{-5}, \tag{8.20}$$

this result can be used together with the theoretical calculation of the normalized decay rate from the UKQCD collaboration to compute

$$|V_{ub}| = (3.56 \pm 0.17 \pm 0.10\ ^{+0.38}_{-0.25}) \times 10^{-3}, \tag{8.21}$$

in accordance with the exclusive result using global fits to the $B \to \pi \ell \nu_\ell$ channel. Other resonances lack evidential value as can be seen, for example,in the case of the semileptonic B^+ decay to the tensor meson $f_2(1270)$ has a branching ratio of

$$\mathcal{B}(B^+ \to f_2(1270)\ell^+\nu) \times \mathcal{B}(f_2(1270) \to \pi^+\pi^-) = (1.8 \pm 0.9_{\text{stat}}{}^{+0.2}_{-0.1\,\text{syst}} \times 10^{-5}, \tag{8.22}$$

with a statistical significance (without including systematic uncertainties) of 2.2σ.

CHAPTER 9

SUMMARY AND OUTLOOK

9.1 Summary

In this thesis, exclusive charmless semileptonic B decays to two charged pions were studied using a 711 fb^{-1} data sample collected at the $\Upsilon(4S)$ resonance with the Belle detector. The events are selected by fully reconstructing one B meson into hadronic decay modes. The other B meson is further studied to extract the signal. As the dominant processes in B decays involve a $b \to c$ quark transition, they become an irreducible source of background to the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay. Since the kinematics of both processes are similar, this thesis implements a Boosted Decision Tree algorithm to reduce this background efficiently.

The main advantage of the hadronic tag in reconstructing B mesons is that it allows determining the properties of the signal via energy and momentum conservation laws. In this manner, one can infer the neutrino kinematics from the missing four-momentum of the event. As the neutrino is the only particle no reconstructed, the missing mass squared is zero. In contrast, the background usually contains intermediate decays, giving rise to massive particles not reconstructed in the event, and consequently, a missing mass squared different from zero. Therefore, the missing mass squared becomes the primary discriminator to obtain the signal. The measurement is carried out in bins of the dipion mass $M_{\pi\pi}$ to reduce model dependence. For each bin, an extended binned maximum likelihood fit to the missing mass squared is performed to determine the signal yields.

The measured total branching ratio, derived from the sum of partial branching ratios for each bin in the dipion mass, is

$$\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu) = (2.24 \pm 0.19_{\text{stat}} \pm 0.13_{\text{syst}}) \times 10^{-4}.$$

Here, the dominant uncertainty to the branching ratio is of a statistical nature due to the small size of the signal samples as a direct consequence of the hadronic tag reconstruction. The most significant source of systematic uncertainty comes from the tagging efficiency correction, which is assigned to be 4.2%. This branching ratio is the first reported measurement for this decay. As such, a comparison with other experiments is yet not possible. Instead, different tests were carried out to validate this measurement. One trial consisted of repeating the procedure in subsamples for this channel, by lepton type or by the charge of the B meson, yielding compatible results with the one previously quoted. Another inspection was carried out using control samples. In this case, semileptonic B decays to a charmed meson (the D^0), in four decay channels, were analyzed using a selection procedure similar to the one used in the signal.

The measured branching ratios are in agreement with the ones quoted by the Particle Data Group [1].

An attempt to resolve possible $\pi^+\pi^-$ states was also carried out using the dipion mass spectrum solely in the signal region. To that end, three approaches were implemented. The first approach used mass line-shapes from MC simulations for the ρ^0 and $f_2(1270)$ mesons, generated with form factor predictions from the Light Cone Sum Rule calculations [98, 99]. The remaining states are assumed to originate from $S-$wave contributions to the dipion mass, and modeled with a scalar form factor derived from dispersion theory. These scalar form factors were provided by Prof. Dr. Christoph Hanhart and Stefan Ropertz from the University of Bonn. The second approach employed analytical mass line-shapes for each possible resonance, such that the incoherent sum of these functions describes the dipion mass spectrum. These functions correspond to a modified version of a Breit-Wigner functions adapted to a particular resonance. The last approach modeled the dipion mass spectrum as a coherent sum of resonances, using the shapes from the second approach. The three methods gave consistent results for the ρ^0 meson, which accounts for $(69 \pm 9)\%$ of the total $\pi^+\pi^-$ states. However, the other particles lacked evidential value due to the poor statistics of the signal sample.

Using the first approach, the branching ratio of semileptonic B decays involving the ρ^0 meson reads as:

$$\mathcal{B}(B^+ \to \rho^0 \ell^+ \nu) = (1.71 \pm 0.16_{\text{stat}} \pm 0.10_{\text{syst}}) \times 10^{-4},$$

in agreement with previous measurements. The preceding result together with the normalized decay rate from the UKQCD [130] calculation are used to compute the CKM matrix element V_{ub}, which is measured to be:

$$|V_{ub}| = (3.56 \pm 0.17_{\text{stat}} \pm 0.10_{\text{syst}} \,{}^{+0.38}_{-0.25}{}_{\text{theo}}) \times 10^{-3}.$$

This result supports the exclusive value, $|V_{ub}| = (3.65 \pm 0.14) \times 10^{-3}$, from global fits to the $B \to \pi \ell \nu_\ell$ channel and holds up the tension at the 3σ level with the inclusive value, $|V_{ub}| = (4.52 \pm 0.15_{\text{exp}} \pm 0.13_{\text{theo}}) \times 10^{-3}$ [41]. Some measurements have reported branching ratios that resulted in $|V_{ub}|$ values much lower than the average exclusive value. Contrary to the discussion in the literature [44–46], this effect is not a consequence of selecting a narrow window in the dipion mass for the analyses, but a result of the treatment of the background underneath the ρ resonance, which has a large impact in untagged analyses.

The product of branching ratios for the $f_2(1270)$ meson is measured to be

$$\mathcal{B}(B^+ \to f_2(1270)\ell^+\nu) \times \mathcal{B}(f_2(1270) \to \pi^+\pi^-) = (1.8 \pm 0.9_{\text{stat}} \,{}^{+0.2}_{-0.1}{}_{\text{syst}} \times 10^{-5},$$

with a 2.2σ statistical significance without taking into account systematic uncertainties.

9.2 Outlook

The signal sample size attained with the hadronic tag method poses an impediment for further analysis of the data. Angular studies, resolution of resonances, measurement of ratios between form factors or even form factors themselves, are not possible with the current statistics. With the full dataset accumulated by the Belle detector, this task can be pursued using an untagged method. Typical signal reconstruction efficiencies achieved with this method are about 2% [43, 131], rendering around 7000 $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ events to perform such studies. However, as the background contamination is higher than in the hadronic method, the systematic uncertainties are expected to dominate the experimental error. To obtain that same amount of signal events with the hadronic tag, a data sample around 18 ab^{-1} would be necessary.

Such a large dataset will be possible to collect in the not so distant future with the Belle II detector.

Complementary studies with other four-body decay channels such as $B^+ \to \pi^0\pi^0\ell^+\nu_\ell$ and $B^0 \to \pi^-\pi^0\ell^+\nu_\ell$, can add more information to the resonance composition in $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$. Because of the arrangement of the angular momentum of the dipion system, the $B^+ \to \pi^0\pi^0\ell^+\nu_\ell$ channel admits only scalar mesons, while the $B^0 \to \pi^-\pi^0\ell^+\nu_\ell$ contains only vector mesons. In contrast, both meson types are found in the $B^+ \to \pi^+\pi^-\ell^+\nu_\ell$ decay [48]. Research on these channels can also be useful in the establishment of some resonances, that currently are still controversial such as the $f_0(500)$ or the excited states of the ρ meson ($\rho(1450)$, $\rho(1700)$). Thus, semileptonic decays of heavy mesons like B or D constitute a probe to study the internal structure of light mesons due to the QCD factorization [132], by which one can disentangle the leptonic and hadronic part of the decay. The information gained on these mesons, for example form factors, can be extended to improve predictions for hadronic decays involving these mesons [49].

Finally, one of the primary concerns in the flavor physics sector is the striking difference in the measured values of the CKM matrix element $|V_{ub}|$ between the inclusive and exclusive approaches. Global fits using SM parameters by the CKMfitter [20] and UTfit [133] groups, as well as measurement in decays of bottom baryons [88], reinforce the latter. This fact may suggest a reconsideration of the theoretical model for inclusive charmless semileptonic B decays. On the experimental side, there is still much work to do, as the measured exclusive modes correspond to approximately a quarter of the total inclusive rate [1]. Therefore, the remaining states demand a separate study to reduce this gap and help to improve the inclusive modeling, since the lack of knowledge on all exclusive $b \to u\ell\nu$ modes is one of the contributions to the systematic uncertainty in the inclusive approach. These mesons have masses above 1 GeV and decay to multi-pion states that, from a theoretical point of view, represents a challenge. Consequently, a prediction of form factors for such particles treated as multi-body states may not be available soon from lattice QCD or sum rules.

BIBLIOGRAPHY

[1] C. Patrignani et al., *Review of Particle Physics*, Chin. Phys. **C40** (2016) 100001.

[2] V. B. Golubev, *B-meson semileptonic decays at BaBar*, Nucl. Phys. Proc. Suppl. **156** (2006) 178.

[3] P. Ramond, "The Five Instructions (TASI 2011)", *The Dark Secrets of the Terascale*, ed. by K. Matchev and et al., 2013 1, eprint: 1201.0396 (hep-ph).

[4] G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1.

[5] S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30.

[6] C. Davies and S. Playfer, eds., *Proc. 55th Scottish Universities Summer School in Physics (St. Andrews, August 2001)*, Scottish Universities Summer School in Physics and Institute of Physics Publishing, 2002.

[7] M. Thomson, *Modern Particle Physics*, 1st ed., Cambridge University Press, 2013.

[8] W. Cottingham and D. Greenwodd, *An Introduction to the Standard Model of Particle Physics*, 2nd ed., Cambridge University Press, 2007.

[9] D. Griffiths, *Introduction to Elementary Particle Physics*, 2nd ed., Cambridge University Press, 2008.

[10] M. Robinson, *Symmetry and the Standar Model. Mathematics and Particle Physics*, 1st ed., Springer, 2011.

[11] A. Bettini, *Introduction to Elementary Particle Physics*, 1st ed., Cambridge University Press, 2008.

[12] C. S. Wu et al., *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (4 1957) 1413.

[13] U. Sarkar, *Particle and Astroparticle Physics*, 1st ed., Taylor and Francis, 2008.

[14] A. Pich, *The Standard Model of Electroweak Interactions*, 2012, arXiv: 1201.0537 [hep-ph].

[15] F. Halzen and A. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, 1st ed., John Wiley and Sons, 1984.

[16] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (12 1963) 531.

[17] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) 652.

[18] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (21 1983) 1945.

[19] R. Aleksan, B. Kayser and D. London, *Determining the Quark Mixing Matrix from* CP*-violating Asymmetries*, Phys. Rev. Lett. **73** (1 1994) 18.

[20] J. Charles et al., *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1.

[21] *The Upsilon System*, `https://www.classe.cornell.edu/public/lab-info/upsilon.html`, Accessed: 2018-04-12.

[22] S. W. Herb, D. C. Hom and e. a. Lederman, *Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (5 1977) 252.

[23] C. W. Darden et al., *Observation of a Narrow Resonance at 9.46-GeV in electron-Positron Annihilations*, Phys. Lett. **76B** (1978) 246.

[24] C. Berger et al., *Observation of a Narrow Resonance Formed in e^+e^- Annihilation at 9.46-GeV*, Phys. Lett. **76B** (1978) 243.

[25] D. Andrews et al., *Observation of Three Upsilon States*, Phys. Rev. Lett. **44** (1980) 1108.

[26] T. Bohringer et al., *Observation of Υ, Υ', and $Upsilon''$ at the Cornell electron Storage Ring*, Phys. Rev. Lett. **44** (1980) 1111.

[27] E. H. Thorndike, "Bottom quark physics: Past, present, future", *Probing luminous and dark matter. Proceedings, Symposium in Honor of Adrian Melissinos, Rochester, USA, September 24-25, 1999*, 1999 127, arXiv: `hep-ex/0003027 [hep-ex]`.

[28] D. Andrews et al., *Observation of a Fourth Upsilon State in e^+e^- Annihilations*, Phys. Rev. Lett. **45** (4 1980) 219.

[29] G. Finocchiaro et al., *Observation of the Υ''' at the Cornell Electron Storage Ring*, Phys. Rev. Lett. **45** (4 1980) 222.

[30] C. Bebek et al., *Evidence for New-Flavor Production at the $\Upsilon(4S)$*, Phys. Rev. Lett. **46** (2 1981) 84.

[31] L. J. Spencer et al., *Measurement of B-Meson Semileptonic Decay*, Phys. Rev. Lett. **47** (11 1981) 771.

[32] S. Behrends et al., *Observation of Exclusive Decay Modes of b-Flavored Mesons*, Phys. Rev. Lett. **50** (12 1983) 881.

[33] S. Stone, "B phenomenology", *Heavy flavor physics: Theory and experimental results in heavy quark physics and CP violation. Proceedings, 55th Scottish Universities Summer School in Physics, SUSSP 2001, St. Andrews, UK, August 7-23, 2001*, 2001 237, arXiv: `hep-ph/0112008 [hep-ph]`.

[34] J. Dingfelder and T. Mannel, *Leptonic and semileptonic decays of B mesons*, Rev. Mod. Phys. **88** (3 2016) 035008.

[35] A. J. Bevan et al., *The Physics of the B Factories*, Eur. Phys. J. **C74** (2014) 3026.

[36] D. Boutigny et al., "The BABAR physics Book: Physics at an asymmetric *B* factory", *Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting) Pasadena, California, September 22-24, 1997*, 1998.

[37] J. D. Richman and P. R. Burchat, *Leptonic and semileptonic decays of charm and bottom hadrons*, Rev. Mod. Phys. **67** (4 1995) 893.

[38] N. Isgur et al., *Semileptonic B and D decays in the quark model*, Phys. Rev. D **39** (3 1989) 799.

[39] D. Scora and N. Isgur, *Semileptonic meson decays in the quark model: An update*, Phys. Rev. D **52** (5 1995) 2783.

[40] T. Huang et al., *Semileptonic B(Bs, Bc) decays in the light-cone QCD sum rules*, International Journal of Modern Physics A **23** (2008) 3237.

[41] Y. Amhis et al., *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, Eur. Phys. J. **C77** (2017) 895.

[42] C. Albertus, E. Hernández and J. Nieves, *$B \to \rho$ semileptonic decays and $|V_{ub}|$*, Phys. Rev. D **90** (1 2014) 013017.

[43] P. del Amo Sanchez et al., *Study of $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ decays and determination of $|V_{ub}|$*, Phys. Rev. D **83** (3 2011) 032007.

[44] C. S. Kim, G. López Castro and S. L. Tostado, *Evaluation of CKM matrix elements from exclusive $P_{\ell 4}$ decays*, Phys. Rev. D **95** (7 2017) 073003.

[45] A. Bharucha, D. M. Straub and R. Zwicky, *$B \to V\ell^+\ell^-$ in the Standard Model from light-cone sum rules*, Journal of High Energy Physics **2016** (2016) 98.

[46] X.-W. Kang et al., *$B_{\ell 4}$ decays and the extraction of $|V_{ub}|$*, Phys. Rev. D **89** (5 2014) 053015.

[47] A. Sibidanov et al., *Study of Exclusive $B \to X_u\ell\nu$ Decays and Extraction of $\|V_{ub}\|$ using Full Reconstruction Tagging at the Belle Experiment*, Phys. Rev. **D88** (2013) 032005.

[48] S. Faller et al., *Disentangling the decay observables in $B^- \to \pi^+\pi^-\ell^-\overline{\nu}_\ell$*, Phys. Rev. D **89** (1 2014) 014015.

[49] S. Cheng, A. Khodjamirian and J. Virto, *$B \to \pi\pi$ form factors from light-cone sum rules with B-meson distribution amplitudes*, Journal of High Energy Physics **2017** (2017) 157.

[50] N. Cabibbo and A. Maksymowicz, *Angular Correlations in K_{e4} Decays and Determination of Low-Energy $\pi - \pi$ Phase Shifts*, Phys. Rev. **137** (2B 1965) B438.

[51] B. Wang, *The Belle II Experiment and SuperKEKB Upgrade*, 2016, arXiv: `1511.09434 [physics.ins-det]`.

[52] S. Kurokawa and E. Kikutani, *Overview of the KEKB accelerators*, Nucl. Instrum. Meth. **A499** (2003) 1.

[53] J. Haba, *KEKB and the BELLE experiment*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **368** (1995) 74, Proceedings of the Third International Workshop on B-Physics at Hadron Machines.

[54] `http://belle.kek.jp`.

[55] B. et al, *Physics achievements from the Belle experiment*, Progress of Theoretical and Experimental Physics **2012** (2012) 04D001.

[56] A. Abashian et al., *The Belle Detector*, Nucl. Instrum. Meth. **A479** (2002) 117.

[57] M. Hazumi, *Performance of the Belle silicon vertex detector*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **473** (2001) 1, Proceedings of the 9th International Workshop on Vertex Detectors.

[58] K. A. et al., *Letter of Intent for KEK Super B Factory. Part II: Detector.*, Report 04-4, KEK, 2004.

[59] H. H. et al., *A high-resolution cylindrical drift chamber for the KEK B-factory*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **455** (2000) 294.

[60] M. A. et al., *Cathode image readout in the BELLE central drift chamber*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **454** (2000) 322.

[61] Shoji and Uno, *The BELLE central drift chamber*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **379** (1996) 421, Proceedings of the Sixth International Conference on Instrumentation for Experiments at e+ e- Colliders.

[62] S. Tavernier, *Experimental Techniques in Nuclear and Particle Physics*, 1st ed., Springer-Verlag Berlin Heidelberg, 2010.

[63] C. Grupen and I. Buvat, *Handbook of Particle Detection and Imaging*, 1st ed., Springer-Verlag Berlin Heidelberg, 2012.

[64] F. Sauli, *Instrumentation in High Energy Physics*, World Scientific Publishing Co, 1992.

[65] K. Miyabayashi, *Belle electromagnetic calorimeter*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **494** (2002) 298, Proceedings of the 8th International Conference on Instrumentatio n for Colliding Beam Physics.

[66] K Hanagaki et al., *Electron identification in Belle*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **485** (2002) 490.

[67] A Abashian et al., *Muon identification in the Belle experiment at KEKB*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **491** (2002) 69.

[68] W. R. Leo, *Techniques for nuclear and particle physics experiments: a how-to approach*, 2nd ed., Springer-Verlag Berlin Heidelberg GmbH, 1994.

[69] O. Behnke et al., eds., *Data analysis in high energy physics*, Wiley-VCH, 2013, ISBN: 9783527410583, 9783527653447, 9783527653430.

[70] P. C. Bhat, *Multivariate Analysis Methods in Particle Physics*,
Annual Review of Nuclear and Particle Science **61** (2011) 281.

[71] I. Narsky and F. C. Porter, *Statistical analysis techniques in particle physics*, Wiley-VCH, 2014, ISBN: 9783527410866, 9783527677313.

[72] M. Feindt and U. Kerzel, *The NeuroBayes neural network package*,
Nucl. Instrum. Meth. **A559** (2006) 190.

[73] M. Feindt et al.,
A hierarchical NeuroBayes-based algorithm for full reconstruction of B mesons at B factories,
Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **654** (2011) 432.

[74] A. Hocker et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040.

[75] G. James et al., *An Introduction to Statistical Learning: With Applications in R*,
Springer Publishing Company, Incorporated, 2014, ISBN: 1461471370, 9781461471370.

[76] Y. Freund and R. E. Schapire,
A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting,
Journal of Computer and System Sciences **55** (1997) 119.

[77] J. H. Friedman, *Greedy Function Approximation: A Gradient Boosting Machine*,
Annals of Statistics **29** (2000) 1189.

[78] L. Breiman, *Bagging Predictors*, Mach. Learn. **24** (1996) 123.

[79] J. H. Friedman, *Stochastic Gradient Boosting*, Comput. Stat. Data Anal. **38** (2002) 367.

[80] G. C. Fox and S. Wolfram,
Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes,
Phys. Rev. Lett. **41** (23 1978) 1581.

[81] K. e. a. Abe, *Evidence for $B^0 \to \pi^0\pi^0$*, Phys. Rev. Lett. **91** (26 2003) 261801.

[82] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*,
Nucl. Instrum. Meth. **A389** (1997) 81.

[83] D. J. Lange, *The EvtGen particle decay simulation package*,
Nucl. Instrum. Meth. **A462** (2001) 152.

[84] R. Brun et al., *GEANT3*, (1987).

[85] E. Barberio, B. van Eijk and Z. Was,
PHOTOS: A Universal Monte Carlo for QED radiative corrections in decays,
Comput. Phys. Commun. **66** (1991) 115.

[86] E. Barberio and Z. Was,
PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0,
Comput. Phys. Commun. **79** (1994) 291.

[87] A. Dobrovolskaya et al., *Multiparticle production in the decays of the B-meson*,
Physics Letters B **229** (1989) 293.

[88] R. Aaij et al., *Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays*,
Nature Phys. **11** (2015) 743.

[89] M. Ladisa et al., *$B \to \pi\pi\ell\nu$ decays in a QCD relativistic potential model*,
Physics Letters B **471** (1999) 81.

[90] S. Nishida, *Study of Kaon and Pion Identification Using Inclusive D^* Sample*, tech. rep., BELLE Note 779: Belle Collaboration.

[91] J. W. et al., *Lepton ID Efficiency Correction and Systematic Error for Experiment 21 to 27*, tech. rep., BELLE Note 777: Belle Collaboration.

[92] J. W. et al., *Lepton ID Efficiency Correction and Systematic Error*, tech. rep., BELLE Note 954: Belle Collaboration.

[93] A. Sibidanov and K. Varvell, *Exclusive $B \to \pi\ell\nu$, $B \to \ell\nu$ and $B \to \omega\ell\nu$ Decays using New Full Reconstruction Tagging*, tech. rep., BELLE Note 1206: Belle Collaboration.

[94] C. O. et al, *Modeling of semileptonic $B_{(S)}$ decays*, tech. rep., BELLE Note 1335: Belle Collaboration.

[95] A. K. Leibovich et al., *Semileptonic B decays to excited charmed mesons*, Phys. Rev. D **57** (1 1998) 308.

[96] B. Casey, *HadronB*, tech. rep., BELLE Note 390: Belle Collaboration.

[97] K. Hara, *Measurement of $B \to \tau\nu$ decay with full reconstruction tagging method*, tech. rep., BELLE Note 1105: Belle Collaboration.

[98] V. B. P. Ball, *Exclusive semileptonic and rare B meson decays in QCD*, Phys. Rev. **D58 (1998) 094016** ().

[99] K.-C. Yang, *B to light tensor meson form factors derived from light-cone sum rules*, Physics Letters B **695** (2011) 444.

[100] J. E. Gaiser, *Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ'*, PhD thesis: SLAC, 1982.

[101] M. Oreglia, *A Study of the Reactions $\psi' \to \gamma\gamma\psi$*, PhD thesis: SLAC, 1980.

[102] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*, PhD thesis: Cracow, INP, 1986.

[103] L. Lyons, *Statistics for nuclear and particle physicists*, Cambridge, Uk: Univ. Pr., 1986, ISBN: 9780521379342.

[104] G. Cowan, *Statistical data analysis*, Oxford, UK: Clarendon, 1998, ISBN: 9780198501565.

[105] R. J. Barlow, *Extended maximum likelihood*, Nucl. Instrum. Meth. **A297** (1990) 496.

[106] S. Baker and R. D. Cousins, *Clarification of the use of CHI-square and likelihood functions in fits to histograms*, Nuclear Instruments and Methods in Physics Research **221** (1984) 437.

[107] L. Lyons and L. Demortier, *Everything you always wanted to know about pulls*, tech. rep., Note-5776: CDF, 2002.

[108] P. Ball and R. Zwicky, *$B_{d,s} \to \rho, \omega, K^*, \phi$ decay form factors from light-cone sum rules reexamined*, Phys. Rev. D **71** (1 2005) 014029.

[109] D. e. a. Côté, *Reweighting of the form factors in exclusive $B \to X\ell\nu$ decays*, The European Physical Journal C - Particles and Fields **38** (2004) 105.

[110] C. Bourrely, L. Lellouch and I. Caprini,
Model-independent description of $B \to \pi l \nu$ decays and a determination of $|V_{ub}|$,
Phys. Rev. D **79** (1 2009) 013008.

[111] P. Ball and G. W. Jones, *$B \to \eta^{(\prime)}$ form factors in QCD*, Journal of High Energy Physics (2007).

[112] B. Bhuyan, *High p_T tracking efficiency using partially reconstructed D^* decays*,
BELLE Note 1165 ().

[113] E. Richter-Was, *QED bremsstrahlung in semileptonic B and leptonic τ decays*,
Phys. Lett. **B303** (1993) 163.

[114] J. P. Lees et al., *Branching fraction and form-factor shape measurements of exclusive charmless semileptonic B decays, and determination of $|V_{ub}|$*, Phys. Rev. D **86** (9 2012) 092004.

[115] G. Ciezarek et al., *A Challenge to Lepton Universality in B Meson Decays*,
Nature **546** (2017) 227.

[116] C. S. Kim, J. Lee and W. Namgung,
CP *violation in the semileptonic B_{l4} decays ($B^\pm \to \pi^+\pi^- l^\pm \nu$)*, Phys. Rev. D **60** (9 1999) 094019.

[117] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*,
eConf **C0303241** (2003) MOLT007, [,186(2003)].

[118] J. T. Daub, C. Hanhart and B. Kubis,
A model-independent analysis of final-state interactions in $\overline{B}^0_{d/s} \to J/\psi\pi\pi$,
Journal of High Energy Physics **2016** (2016) 9.

[119] H. Muramatsu et al., *Dalitz analysis of $D^0 \to K^0_S\pi^+\pi^-$*,
Phys. Rev. Lett. **89** (2002) 251802, [Erratum: Phys. Rev. Lett.90,059901(2003)].

[120] M. Ablikim et al., *Resonances in $J/psi \to \phi\pi^+\pi^-$ and ϕK^+K^-*, Phys. Lett. **B607** (2005) 243.

[121] J. P. Lees et al., *Precise Measurement of the $e^+e^- \to \pi^+\pi^-(\gamma)$ Cross Section with the Initial-State Radiation Method at BABAR*, Phys. Rev. **D86** (2012) 032013.

[122] R. Aaij et al., *Dalitz plot analysis of $B^0 \to \overline{D}^0\pi^+\pi^-$ decays*, Phys. Rev. **D92** (2015) 032002.

[123] Y. Li et al., *The S-wave resonance contributions to the three-body decays $B^0_{(s)} \to \eta_c f_0(X) \to \eta_c\pi^+\pi^-$ in perturbative QCD approach*, Eur. Phys. J. **C76** (2016) 675.

[124] R Aaij et al., *Analysis of the resonant components in $B_s \to J/\psi\pi^+\pi^-$*,
Phys. Rev. **D86** (2012) 052006.

[125] G. J. Gounaris and J. J. Sakurai,
Finite-Width Corrections to the Vector-Meson-Dominance Prediction for $\rho \to e^+e^-$,
Phys. Rev. Lett. **21** (4 1968) 244.

[126] S. M. Flatte, *Coupled - Channel Analysis of the $\pi\eta$ and $K\bar{K}$ Systems Near $K\bar{K}$ Threshold*,
Phys. Lett. **63B** (1976) 224.

[127] D. V. Bugg, *The Mass of the sigma pole*, J. Phys. **G34** (2007) 151.

[128] B. H. Behrens et al., *Measurement of $B \to \rho\ell\nu$ decay and $|V_{ub}|$*, Phys. Rev. **D61** (2000) 052001.

[129] T. Hokuue et al., *Measurements of branching fractions and q^2 distributions for $B \to \pi\ell\nu$ and $B \to \rho\ell\nu$ decays with $B \to D^{(*)}\ell\nu$ decay tagging*, Phys. Lett. **B648** (2007) 139.

[130] L. Del Debbio et al., *Lattice constrained parametrizations of form-factors for semileptonic and rare radiative B decays*, Phys. Lett. **B416** (1998) 392.

[131] C. Schwanda et al., *Evidence for $B^+ \to \omega \ell^+ \nu$*, Phys. Rev. Lett. **93** (2004) 131803.

[132] Y.-J. Shi, W. Wang and S. Zhao,
Chiral dynamics, S-wave contributions and angular analysis in $D \to \pi\pi\ell\bar{\nu}$,
Eur. Phys. J. **C77** (2017) 452.

[133] M. Bona et al., *The Unitarity Triangle Fit in the Standard Model and Hadronic Parameters from Lattice QCD: A Reappraisal after the Measurements of $\Delta m(s)$ and $\mathcal{B}(B \to \tau \nu_{tau})$*,
JHEP **10** (2006) 081.

APPENDIX A

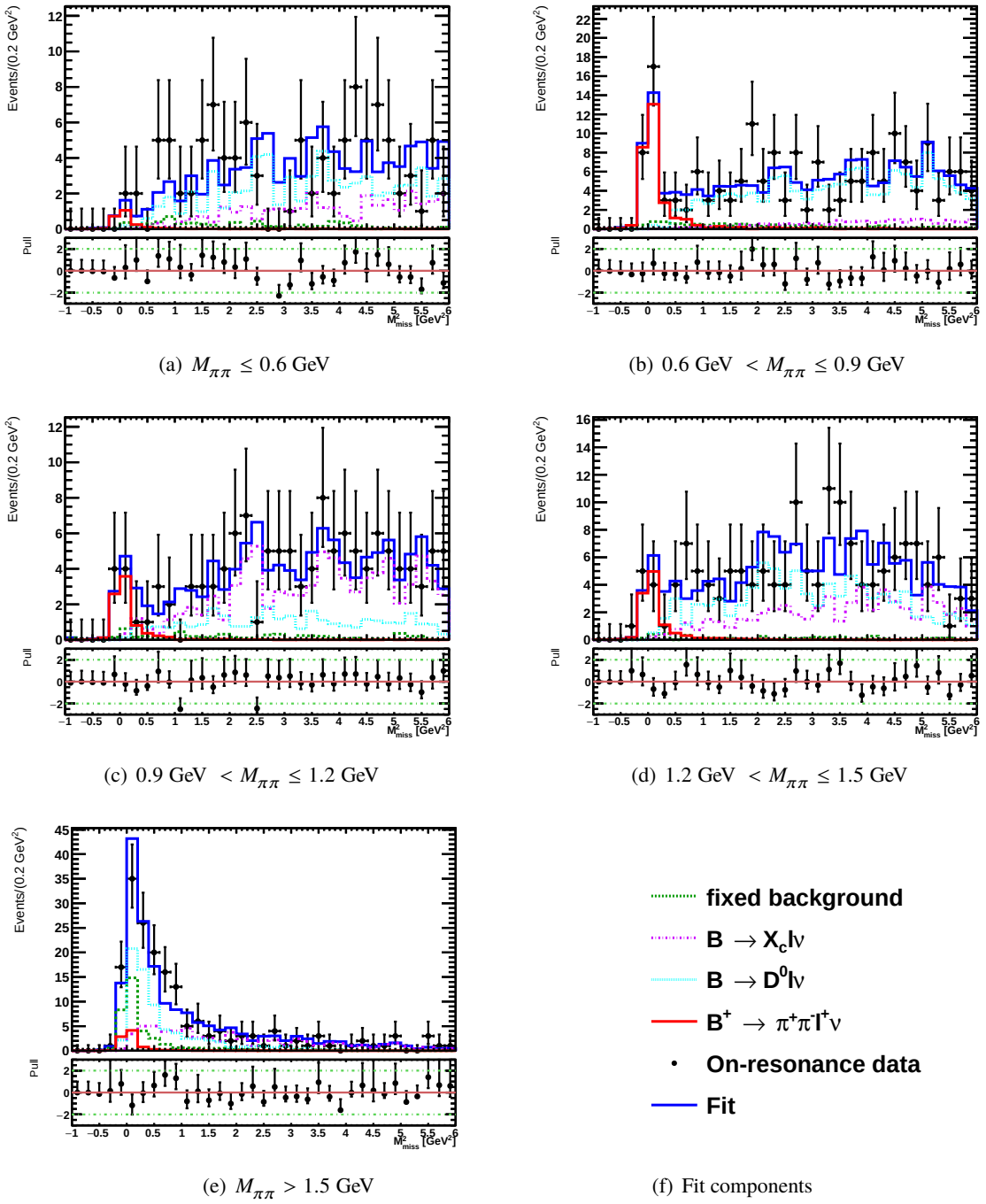
APPENDIX

A.1 MC modelling

B^+ Mode	$\mathcal{B}(\times 10^{-2})$ (MC)	$\mathcal{B}(\times 10^{-2})$ (PDG)	FF model and FF's
$D^{*0}\ell\nu$	5.79	5.68±0.19	$\rho^2 = 1.24 \pm 0.04$, $R_1 = 1.41 \pm 0.049$, $R_2 = 0.844 \pm 0.027$
$D^0\ell\nu$	2.31	2.23±0.11	$\rho^2 = 1.18 \pm 0.04 \pm 0.04$
$D_1^0\ell\nu$	0.81		ISGW2
$D_2^{*0}\ell\nu$	0.39		ISGW2
$D_0^{*0}\ell\nu$	0.91		ISGW2
$D_1^{*0}\ell\nu$	0.81		ISGW2
$D^0(2S)\ell\nu$	0.02		ISGW2
$D^{*0}(2S)\ell\nu$	0.05		ISGW2
$D^{(*)}_{\text{non-res}}\ell\nu$			Goity-Roberts
B^0 Mode	$\mathcal{B}(\times 10^{-3})$ (MC)	$\mathcal{B}(\times 10^{-3})$ (PDG)	FF model and FF's
$D^{*+}\ell\nu$	5.33	5.11±0.23	
$D^+\ell\nu$	2.13	2.17±0.12	$\rho = 1.18 \pm 0.04 \pm 0.04$
$D_1^+\ell\nu$	0.74		ISGW2
$D_2^{*+}\ell\nu$	0.36		ISGW2
$D_0^{*+}\ell\nu$	0.84		ISGW2
$D_1^{*+}\ell\nu$	0.74		ISGW2
$D^+(2S)\ell\nu$	0.02		ISGW2
$D^{*+}(2S)\ell\nu$	0.05		ISGW2
$D^{(*)}_{\text{non-res}}\ell\nu$			Goity-Roberts

Table A.1: Summary of the $b \to c\ell\nu$ modes used in the Monte Carlo simulation.

A.2 Fitted M^2_{miss} in the two dimensional configuration



(a) $M_{\pi\pi} \leq 0.6$ GeV

(b) 0.6 GeV $< M_{\pi\pi} \leq 0.9$ GeV

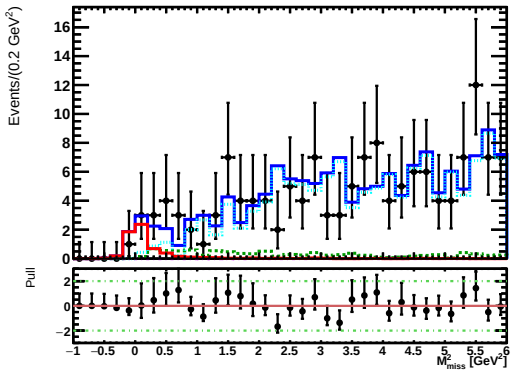
(c) 0.9 GeV $< M_{\pi\pi} \leq 1.2$ GeV

(d) 1.2 GeV $< M_{\pi\pi} \leq 1.5$ GeV

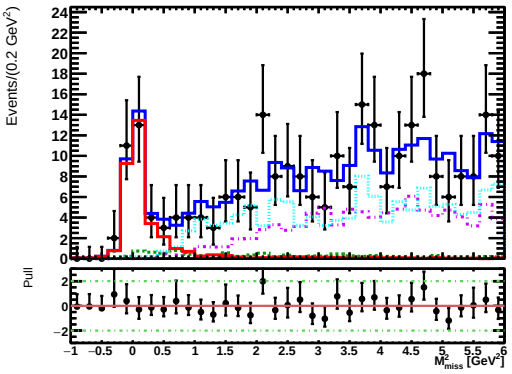
(e) $M_{\pi\pi} > 1.5$ GeV

(f) Fit components

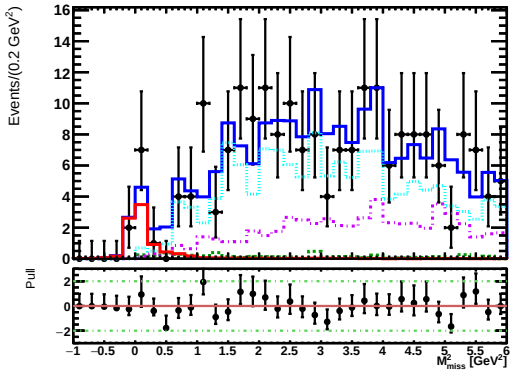
Figure A.1: Distributions of the M^2_{miss} after a binned extended maximum likelihood fit in bins of the the di-pion mass and q^2. The plots are presented in different regions of di-pion mass with a bin width of 300 MeV, and for $q^2 < 4$ GeV^2. The fit components include the signal, $B^+ \to D^0 \ell^+ \nu_\ell$, other $B \to X_c \ell \nu$ and a fixed background. The latter contains contributions from continuum, $B \to X_u \ell \nu$ and rare B decays, that due to their poor statistics remains constant in the fit with a value equal to the prediction of the MC simulation. The pull values for each bin are presented underneath every plot, to show the accuracy of the fit with respect to the data. In most of the cases the pull values do not exceed 2σ indicating a good consistency between the data and the fit result.



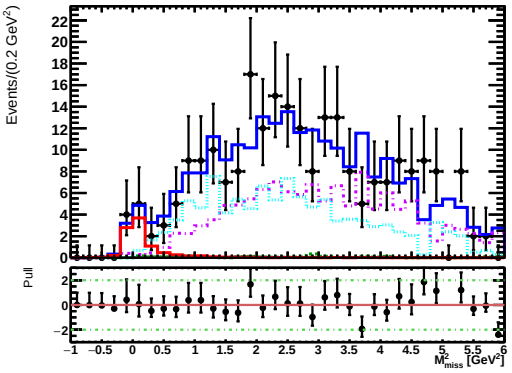
(a) $M_{\pi\pi} \leq 0.6$ GeV



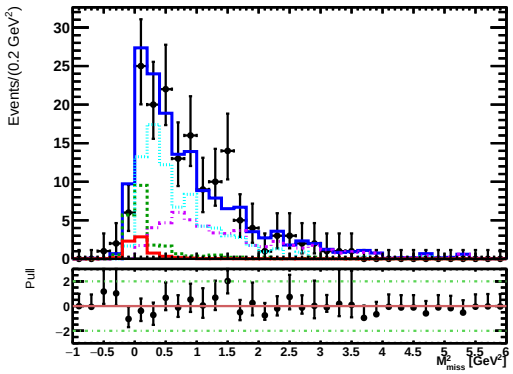
(b) 0.6 GeV $< M_{\pi\pi} \leq 0.9$ GeV



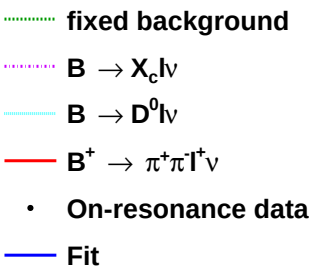
(c) 0.9 GeV $< M_{\pi\pi} \leq 1.2$ GeV



(d) 1.2 GeV $< M_{\pi\pi} \leq 1.5$ GeV

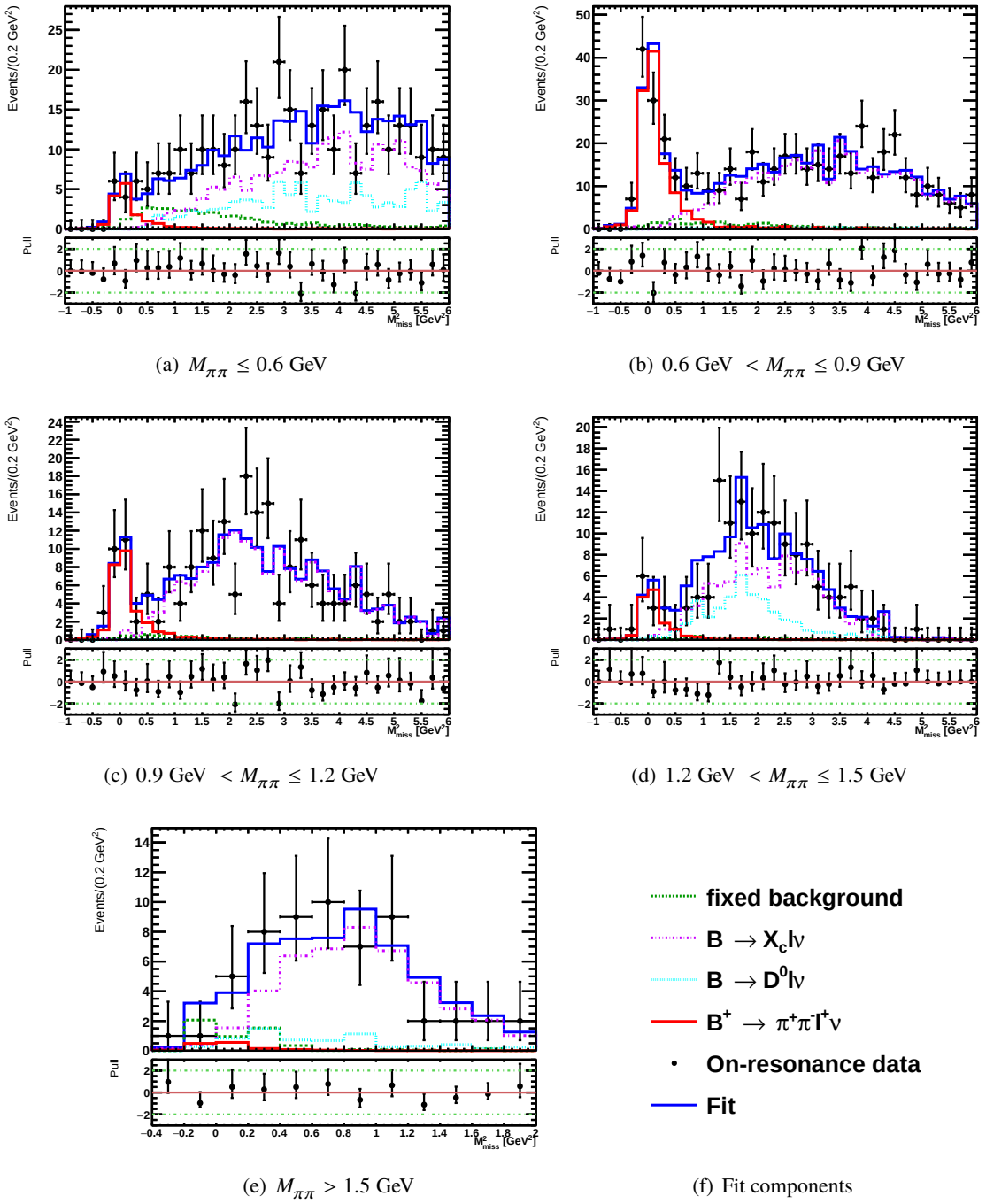


(e) $M_{\pi\pi} > 1.5$ GeV



(f) Fit components

Figure A.2: Distributions of the M^2_{miss} after a binned extended maximum likelihood fit in bins of the the di-pion mass and q^2. The plots are presented in different regions of di-pion mass with a bin width of 300 MeV, and for q^2 in the range from 4 GeV^2 to 8 GeV^2. The fit components include the signal, $B^+ \to D^0 \ell^+ \nu_\ell$, other $B \to X_c \ell \nu$ and a fixed background. The latter contains contributions from continuum, $B \to X_u \ell \nu$ and rare B decays, that due to their poor statistics remains constant in the fit with a value equal to the prediction of the MC simulation. The pull values for each bin are presented underneath every plot, to show the accuracy of the fit with respect to the data. In most of the cases the pull values do not exceed 2σ indicating a good consistency between the data and the fit result.



(a) $M_{\pi\pi} \leq 0.6$ GeV

(b) 0.6 GeV $< M_{\pi\pi} \leq 0.9$ GeV

(c) 0.9 GeV $< M_{\pi\pi} \leq 1.2$ GeV

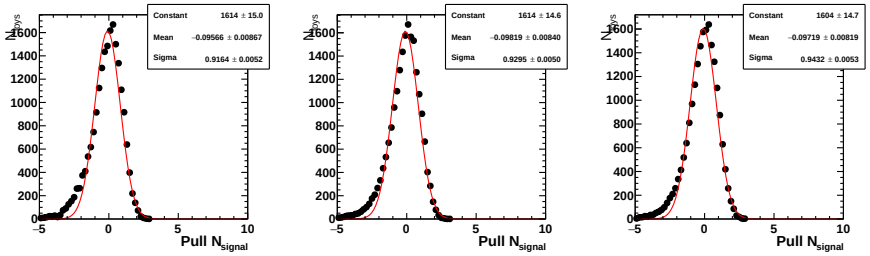
(d) 1.2 GeV $< M_{\pi\pi} \leq 1.5$ GeV

(e) $M_{\pi\pi} > 1.5$ GeV

(f) Fit components

Figure A.3: Distributions of the M^2_{miss} after a binned extended maximum likelihood fit in bins of the the di-pion mass and q^2. The plots are presented in different regions of di-pion mass with a bin width of 300 MeV, and for $q^2 > 8$ GeV^2. The fit components include the signal, $B^+ \to D^0 \ell^+ \nu_\ell$, other $B \to X_c \ell \nu$ and a fixed background. The latter contains contributions from continuum, $B \to X_u \ell \nu$ and rare B decays, that due to their poor statistics remains constant in the fit with a value equal to the prediction of the MC simulation. The pull values for each bin are presented underneath every plot, to show the accuracy of the fit with respect to the data. In most of the cases the pull values do not exceed 2σ indicating a good consistency between the data and the fit result.

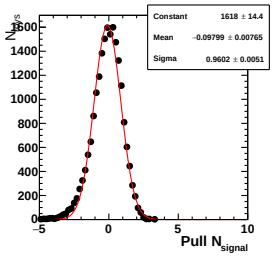
A.3 Pull distributions



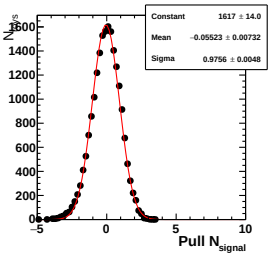
(a) bin 1

(b) bin 2

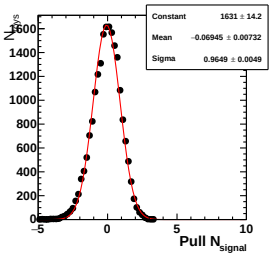
(c) bin 3



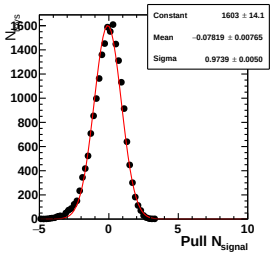
(d) bin 4



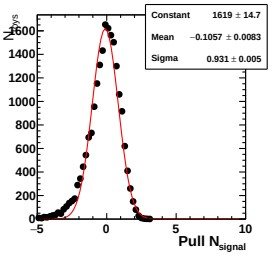
(e) bin 5



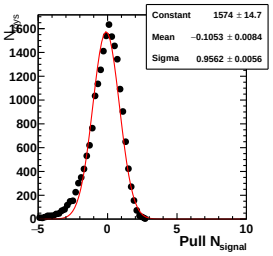
(f) bin 6



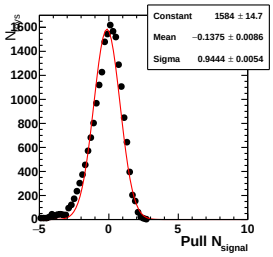
(g) bin 7



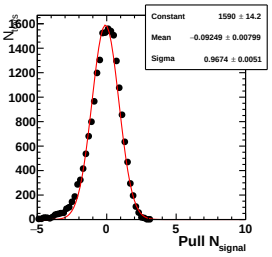
(h) bin 8



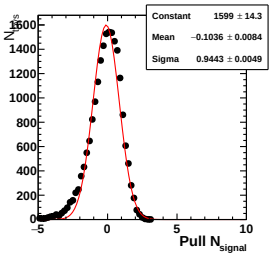
(i) bin 9



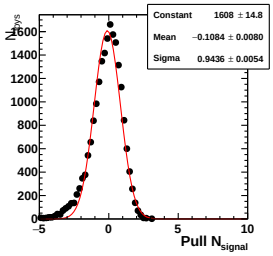
(j) bin 10



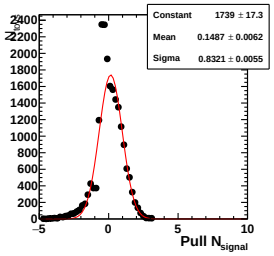
(k) bin 11



(l) bin 12

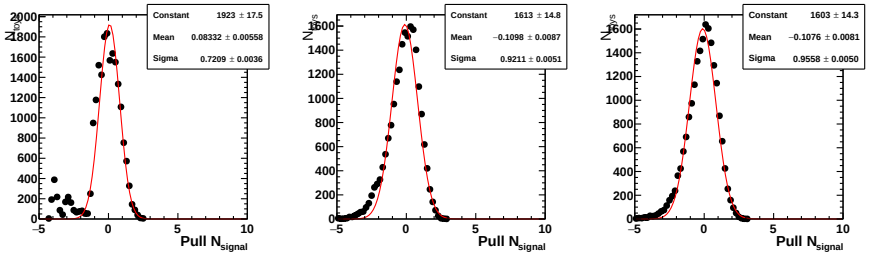


(m) bin 13



(n) bin 14

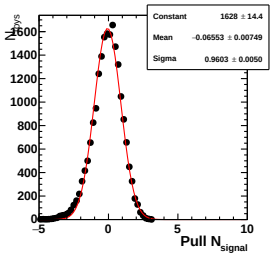
Figure A.4: Pull distribution for the signal yield drawn from 20000 toy MC for the one dimensional binning configuration. The mean and width of the pull are obtained from a fit to a Gaussian function.



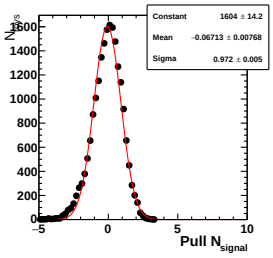
(a) bin 1

(b) bin 2

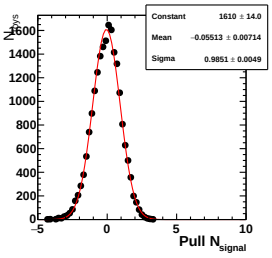
(c) bin 3



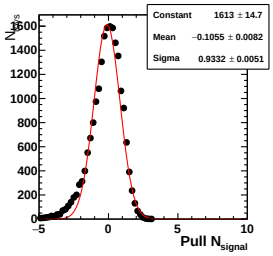
(d) bin 4



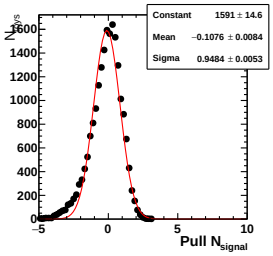
(e) bin 5



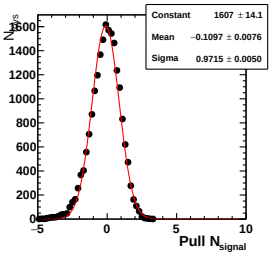
(f) bin 6



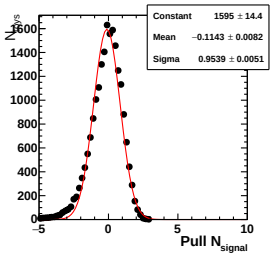
(g) bin 7



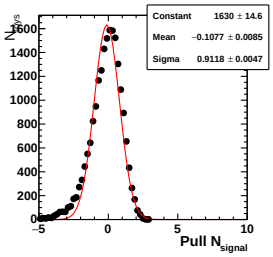
(h) bin 8



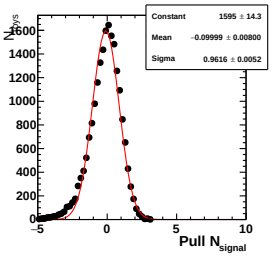
(i) bin 9



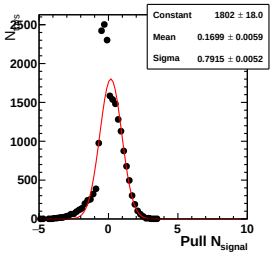
(j) bin 10



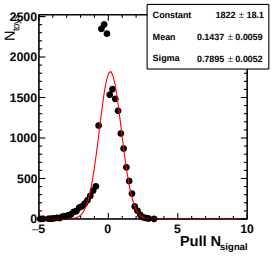
(k) bin 11



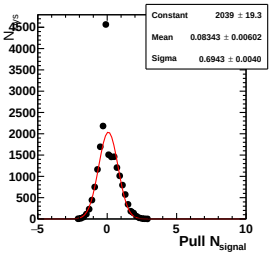
(l) bin 12



(m) bin 13

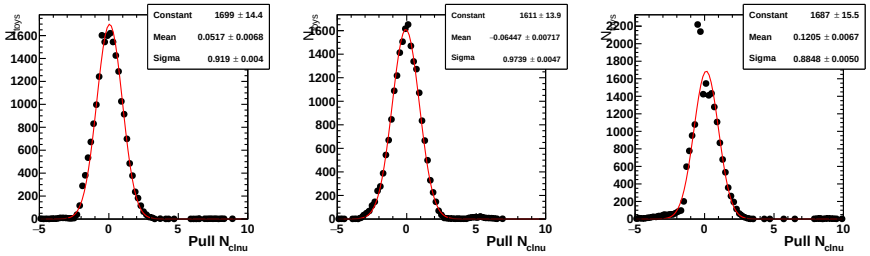


(n) bin 14



(o) bin 15

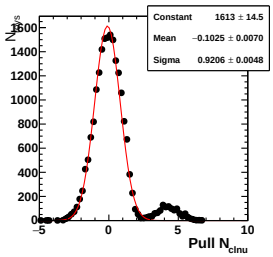
Figure A.5: Pull distribution for the signal yield drawn from 20000 toy MC for the two dimensional binning configuration. The mean and width of the pull are obtained from a fit to a Gaussian function.



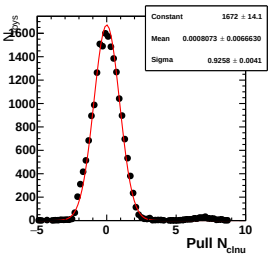
(a) bin 1

(b) bin 2

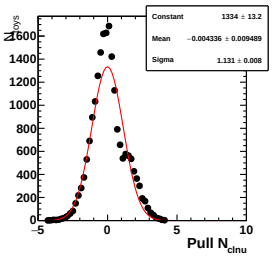
(c) bin 3



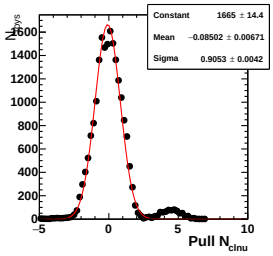
(d) bin 4



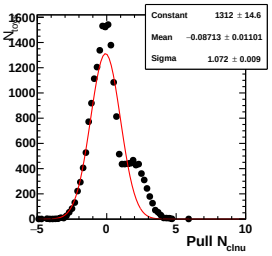
(e) bin 5



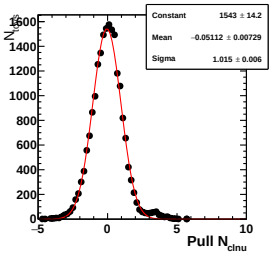
(f) bin 6



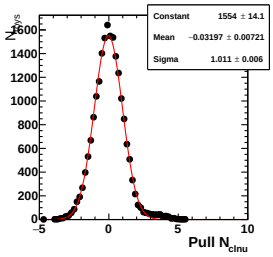
(g) bin 7



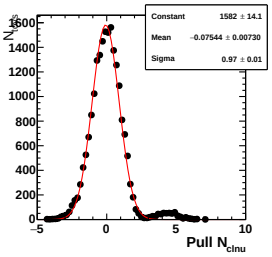
(h) bin 8



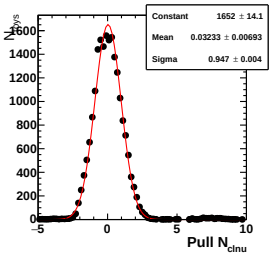
(i) bin 9



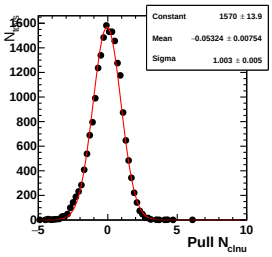
(j) bin 10



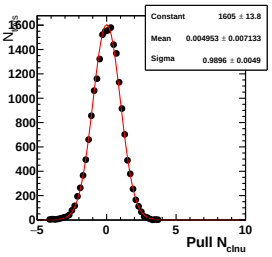
(k) bin 11



(l) bin 12

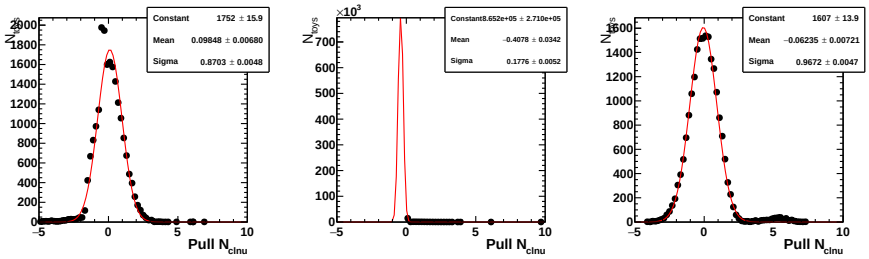


(m) bin 13



(n) bin 14

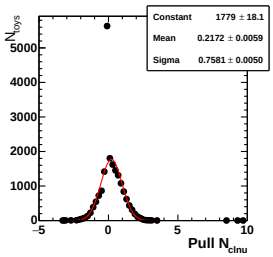
Figure A.6: Pull distribution for the rest of the $B \to X_c \ell \nu$ fit component drawn from 20000 toy MC for the one dimensional binning configuration. The mean and width of the pull are obtained from a fit to a Gaussian function.



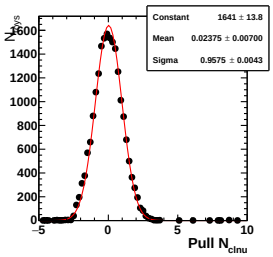
(a) bin 1

(b) bin 2

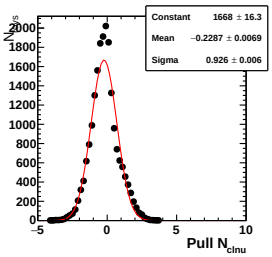
(c) bin 3



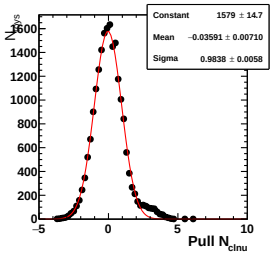
(d) bin 4



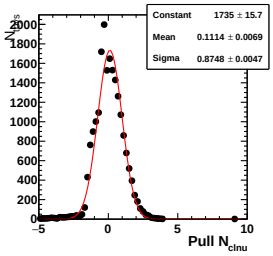
(e) bin 5



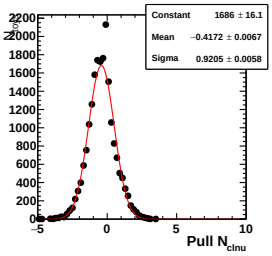
(f) bin 6



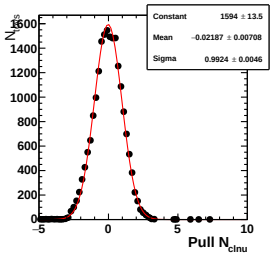
(g) bin 7



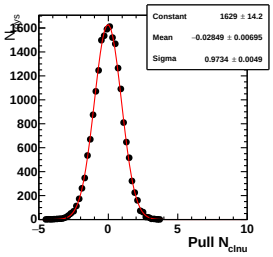
(h) bin 8



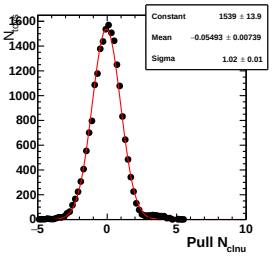
(i) bin 9



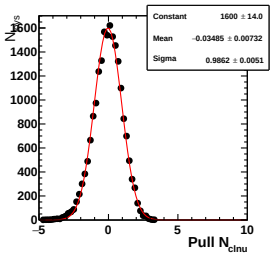
(j) bin 10



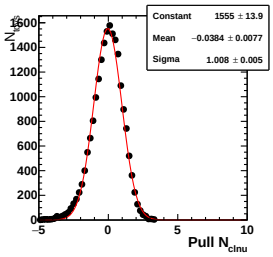
(k) bin 11



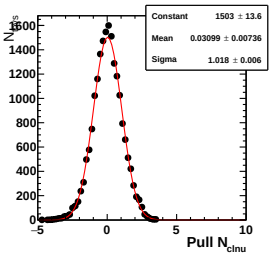
(l) bin 12



(m) bin 13

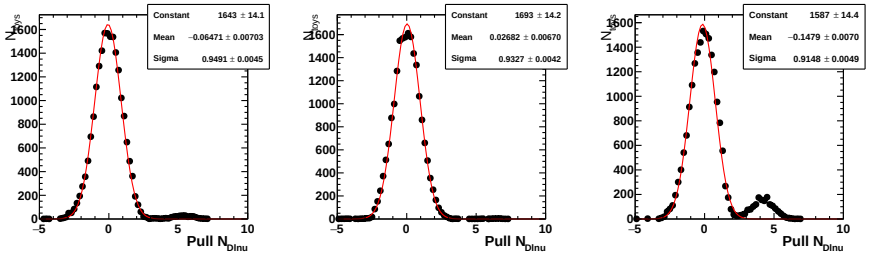


(n) bin 14



(o) bin 15

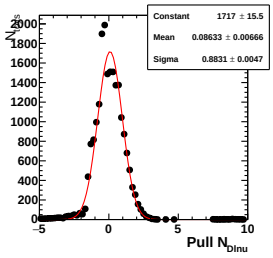
Figure A.7: Pull distribution for the rest of the $B \to X_c \ell \nu$ fit component drawn from 20000 toy MC for the two dimensional binning configuration. The mean and width of the pull are obtained from a fit to a Gaussian function.



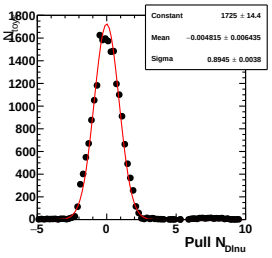
(a) bin 1

(b) bin 2

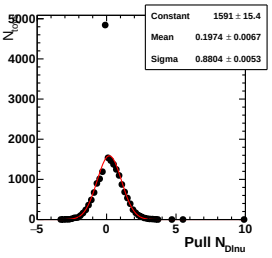
(c) bin 3



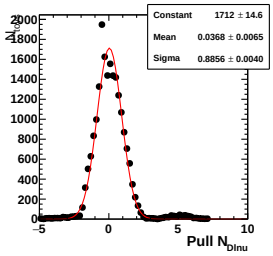
(d) bin 4



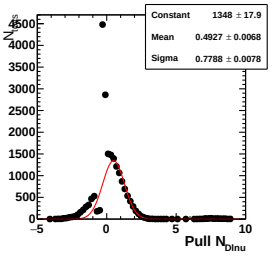
(e) bin 5



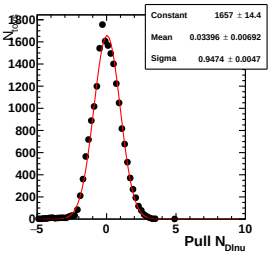
(f) bin 6



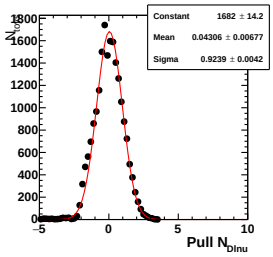
(g) bin 7



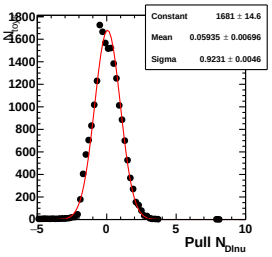
(h) bin 8



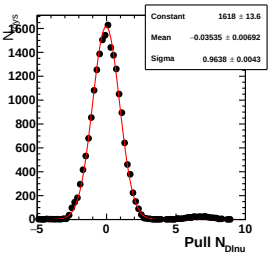
(i) bin 9



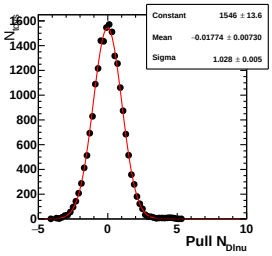
(j) bin 10



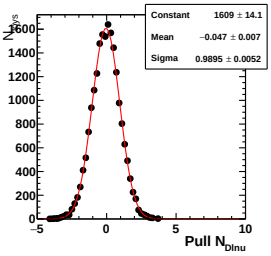
(k) bin 11



(l) bin 12

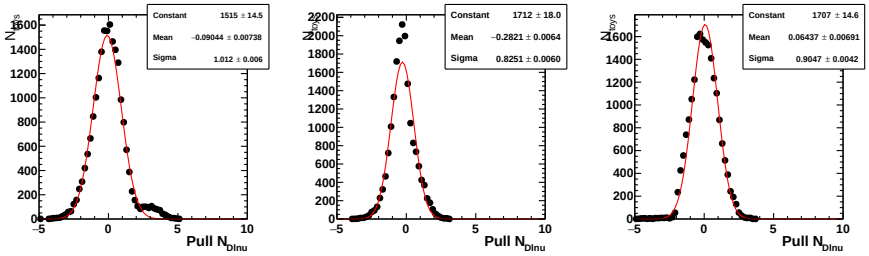


(m) bin 13



(n) bin 14

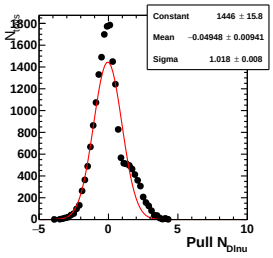
Figure A.8: Pull distribution for the $B \to D^0 \ell \nu$ fit component drawn from 20000 toy MC for the one dimensional binning configuration. The mean and width of the pull are obtained from a fit to a Gaussian function.



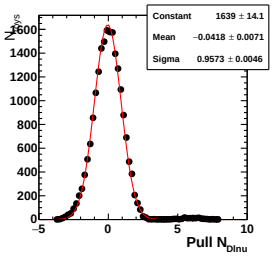
(a) bin 1

(b) bin 2

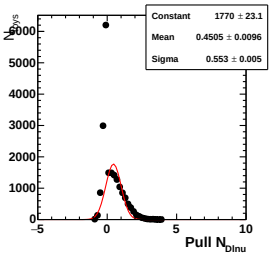
(c) bin 3



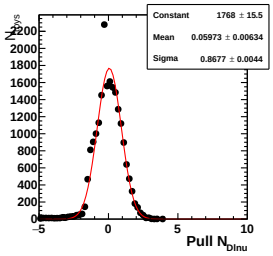
(d) bin 4



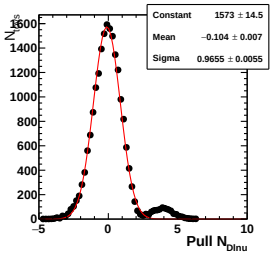
(e) bin 5



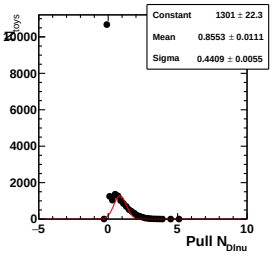
(f) bin 6



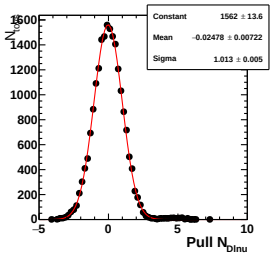
(g) bin 7



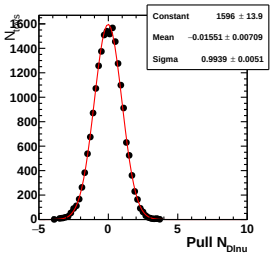
(h) bin 8



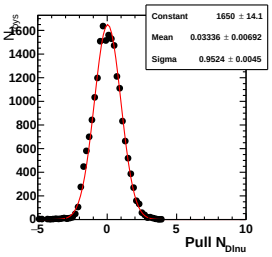
(i) bin 9



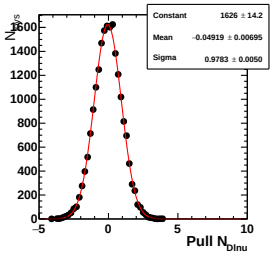
(j) bin 10



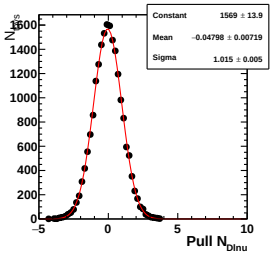
(k) bin 11



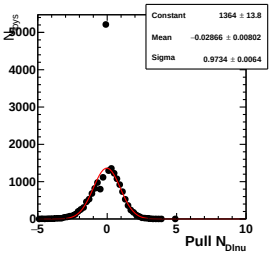
(l) bin 12



(m) bin 13



(n) bin 14



(o) bin 15

Figure A.9: Pull distribution for the $B \to D^0 \ell \nu$ fit component drawn from 20000 toy MC for the two dimensional binning configuration. The mean and width of the pull are obtained from a fit to a Gaussian function.

A.4 Fit Results and systematic uncertainties for subsamples in $B \to \pi^+\pi^-\ell\nu_\ell$ separated by lepton type and B meson charge

Bin	Cor. Y_{signal}	ϵ [10^{-4}]	$\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$ [10^{-5}]
1	5.7±3.0	9.07±1.00	$0.79 \pm 0.43\ ^{+0.06}_{-0.06}$
2	4.6±2.7	8.66±1.12	$0.67 \pm 0.40\ ^{+0.03}_{-0.03}$
3	9.5±3.6	8.99±1.04	$1.33 \pm 0.53\ ^{+0.21}_{-0.21}$
4	13.9±4.2	7.73±0.90	$2.27 \pm 0.73\ ^{+0.30}_{-0.31}$
5	39.9±6.9	9.84±0.97	$5.11 \pm 1.02\ ^{+0.35}_{-0.35}$
6	27.4±5.7	8.77±0.87	$3.94 \pm 0.91\ ^{+0.40}_{-0.40}$
7	19.2±4.7	8.36±0.82	$2.90 \pm 0.76\ ^{+0.21}_{-0.21}$
8	5.0±2.6	8.64±0.81	$0.73 \pm 0.39\ ^{+0.09}_{-0.10}$
9	3.9±2.2	8.51±0.79	$0.58 \pm 0.33\ ^{+0.04}_{-0.04}$
10	3.3±2.2	9.28±0.80	$0.45 \pm 0.30\ ^{+0.02}_{-0.02}$
11	6.6±2.9	9.68±0.81	$0.86 \pm 0.38\ ^{+0.03}_{-0.03}$
12	3.6±2.1	8.11±0.73	$0.56 \pm 0.33\ ^{+0.05}_{-0.05}$
13	3.8±2.4	9.24±0.77	$0.52 \pm 0.33\ ^{+0.04}_{-0.04}$
14	8.2±10.2	8.05±0.15	$1.28 \pm 1.60\ ^{+0.28}_{-0.27}$

Table A.2: Corrected signal yields (Y_{signal}), signal reconstruction efficiency (ϵ) and partial branching ratios ($\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$) ratios for each bin in the one dimensional configuration. These values correspond to a sample containing only electrons ($B^\pm \to \pi^+\pi^- e^\pm \nu_e$).

Bin	Cor. Y_{signal}	ϵ [10^{-4}]	$\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$ [10^{-5}]
1	0.7±1.9	6.67±0.86	$0.13 \pm 0.36\ ^{+0.04}_{-0.04}$
2	5.0±2.7	7.60±1.04	$0.83 \pm 0.46\ ^{+0.05}_{-0.05}$
3	2.3±1.8	6.64±0.89	$0.44 \pm 0.35\ ^{+0.03}_{-0.03}$
4	7.6±3.5	8.01±0.91	$1.20 \pm 0.57\ ^{+0.12}_{-0.12}$
5	48.0±7.2	8.71±0.90	$6.95 \pm 1.27\ ^{+0.84}_{-0.84}$
6	23.8±5.2	6.68±0.76	$4.49 \pm 1.11\ ^{+0.15}_{-0.15}$
7	10.1±3.5	7.92±0.80	$1.61 \pm 0.58\ ^{+0.12}_{-0.12}$
8	5.2±2.6	8.22±0.79	$0.80 \pm 0.41\ ^{+0.05}_{-0.05}$
9	4.8±2.5	9.00±0.81	$0.67 \pm 0.36\ ^{+0.02}_{-0.02}$
10	2.6±1.8	8.61±0.77	$0.38 \pm 0.27\ ^{+0.03}_{-0.03}$
11	10.1±3.5	8.34±0.75	$1.53 \pm 0.55\ ^{+0.12}_{-0.12}$
12	8.2±3.1	8.22±0.73	$1.26 \pm 0.49\ ^{+0.03}_{-0.03}$
13	4.7±2.6	7.73±0.70	$0.77 \pm 0.43\ ^{+0.06}_{-0.06}$
14	6.4±8.5	6.74±0.14	$1.20 \pm 1.59\ ^{+0.96}_{-0.77}$

Table A.3: Corrected signal yields (Y_{signal}), signal reconstruction efficiency (ϵ) and partial branching ratios ($\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$) for each bin in the one dimensional configuration. These values correspond to a sample containing only muons ($B^\pm \to \pi^+\pi^-\mu^\pm\nu_\mu$).

Bin	Cor. Y_{signal}	ϵ [10^{-4}]	$\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$ [10^{-5}]
1	2.0±2.1	7.98±0.94	$0.32 \pm 0.33\ ^{+0.04}_{-0.03}$
2	6.3±3.1	7.41±1.03	$1.07 \pm 0.55\ ^{+0.09}_{-0.08}$
3	4.3±2.6	6.89±0.91	$0.79 \pm 0.49\ ^{+0.09}_{-0.10}$
4	14.8±4.5	8.43±0.94	$2.21 \pm 0.72\ ^{+0.28}_{-0.28}$
5	49.7±7.4	10.50±1.00	$5.97 \pm 1.05\ ^{+0.45}_{-0.45}$
6	30.5±5.9	8.22±0.84	$4.68 \pm 1.02\ ^{+0.29}_{-0.29}$
7	15.4±4.3	8.36±0.82	$2.32 \pm 0.69\ ^{+0.19}_{-0.19}$
8	6.9±3.0	8.64±0.81	$1.01 \pm 0.45\ ^{+0.08}_{-0.08}$
9	7.5±3.1	8.65±0.79	$1.09 \pm 0.46\ ^{+0.06}_{-0.06}$
10	5.2±2.3	8.57±0.77	$0.76 \pm 0.35\ ^{+0.04}_{-0.04}$
11	10.4±3.7	9.97±0.82	$1.32 \pm 0.48\ ^{+0.07}_{-0.07}$
12	4.7±2.3	8.13±0.73	$0.73 \pm 0.36\ ^{+0.04}_{-0.05}$
13	6.3±3.0	8.27±0.73	$0.96 \pm 0.47\ ^{+0.03}_{-0.03}$
14	0.0±12.6	7.25±0.15	$0.00 \pm 2.19\ ^{+0.00}_{-0.00}$

Table A.4: Corrected signal yields (Y_{signal}), signal reconstruction efficiency (ϵ) and partial branching ratios ($\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$) for each bin in the one dimensional configuration. These values correspond to a sample containing only negatively charged B mesons ($B^- \to \pi^+\pi^-\ell^-\bar{\nu}_\ell$).

Bin	Cor. Y_{signal}	ϵ [10^{-4}]	$\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$ [10^{-5}]
1	4.8±2.9	7.86±0.93	$0.77 \pm 0.47\ ^{+0.05}_{-0.06}$
2	3.6±2.4	8.99±1.14	$0.50 \pm 0.34\ ^{+0.04}_{-0.04}$
3	6.2±2.8	8.74±1.02	$0.89 \pm 0.42\ ^{+0.09}_{-0.09}$
4	9.1±3.6	7.32±0.88	$1.57 \pm 0.65\ ^{+0.20}_{-0.20}$
5	39.1±6.8	8.24±0.88	$5.99 \pm 1.22\ ^{+0.64}_{-0.64}$
6	20.1±5.0	7.39±0.80	$3.43 \pm 0.93\ ^{+0.20}_{-0.20}$
7	13.7±4.2	8.07±0.80	$2.13 \pm 0.68\ ^{+0.17}_{-0.17}$
8	3.6±2.5	8.38±0.80	$0.54 \pm 0.38\ ^{+0.06}_{-0.06}$
9	2.0±1.8	9.00±0.81	$0.28 \pm 0.25\ ^{+0.02}_{-0.02}$
10	0.0±2.0	9.46±0.81	$0.00 \pm 0.27\ ^{+0.00}_{-0.00}$
11	5.4±2.8	8.19±0.74	$0.83 \pm 0.44\ ^{+0.06}_{-0.06}$
12	7.7±3.1	8.33±0.74	$1.17 \pm 0.48\ ^{+0.08}_{-0.08}$
13	2.8±2.0	8.82±0.75	$0.40 \pm 0.29\ ^{+0.02}_{-0.02}$
14	10.3±9.9	7.56±0.15	$1.72 \pm 1.65\ ^{+0.27}_{-0.27}$

Table A.5: Corrected signal yields (Y_{signal}), signal reconstruction efficiency (ϵ) and partial branching ratios ($\Delta\mathcal{B}(B^+ \to \pi^+\pi^-\ell^+\nu_\ell)$) for each bin in the one dimensional configuration. These values correspond to a sample containing only positively charged B mesons ($B^+ \to \pi^+\pi^-\ell^+\nu_\ell$).

Source	bin 1	bin 2	bin 3	bin 4	bin 5	bin 6	bin 7	bin 8	bin 9	bin 10	bin 11	bin 12	bin 13	bin 14
FF $B \to D^{(*)}\ell\nu$	$^{+0.11}_{-0.13}$	$^{+0.09}_{-0.08}$	$^{+0.06}_{-0.08}$	$^{+0.11}_{-0.16}$	$^{+0.13}_{-0.08}$	$^{+0.04}_{-0.05}$	$^{+0.04}_{-0.13}$	$^{+0.19}_{-0.18}$	$^{+0.24}_{-0.70}$	$^{+98.88}_{-103.91}$	$^{+0.14}_{-0.18}$	$^{+0.14}_{-0.18}$	$^{+0.61}_{-0.20}$	$^{+2.36}_{-1.10}$
FF $B \to D^{**}\ell\nu$	$^{+0.20}_{-0.20}$	$^{+0.54}_{-0.54}$	$^{+0.22}_{-0.22}$	$^{+0.70}_{-0.70}$	$^{+0.09}_{-0.09}$	$^{+0.09}_{-0.09}$	$^{+0.45}_{-0.45}$	$^{+1.08}_{-1.08}$	$^{+1.95}_{-1.95}$	$^{+53.85}_{-53.85}$	$^{+0.42}_{-0.42}$	$^{+0.10}_{-0.10}$	$^{+0.27}_{-0.27}$	$^{+1.96}_{-1.96}$
Shapes $B \to X_u\ell\nu$	±2.41	±1.44	±0.36	±0.82	±0.42	±0.57	±0.49	±1.05	±0.49	±187.05	±0.22	±0.14	±0.12	±0.81
$\mathcal{B}(B \to D^{(*)}\ell\nu)$	$^{+0.14}_{-0.07}$	$^{+0.08}_{-0.08}$	$^{+0.27}_{-0.09}$	$^{+0.18}_{-0.63}$	$^{+0.21}_{-0.15}$	$^{+0.07}_{-0.04}$	$^{+0.12}_{-0.13}$	$^{+0.37}_{-0.48}$	$^{+0.49}_{-1.04}$	$^{+166.94}_{-170.93}$	$^{+0.67}_{-0.44}$	$^{+0.25}_{-0.22}$	$^{+0.18}_{-0.31}$	$^{+5.26}_{-6.58}$
$\mathcal{B}(B \to D^{**}\ell\nu)$	$^{+0.11}_{-0.18}$	$^{+0.11}_{-0.11}$	$^{+0.59}_{-0.24}$	$^{+0.22}_{-0.22}$	$^{+0.19}_{-0.14}$	$^{+0.08}_{-0.08}$	$^{+0.15}_{-0.14}$	$^{+0.42}_{-0.34}$	$^{+0.57}_{-0.66}$	$^{+198.65}_{-135.42}$	$^{+0.35}_{-0.06}$	$^{+0.26}_{-0.23}$	$^{+0.29}_{-0.42}$	$^{+0.79}_{-1.04}$
$\mathcal{B}(B \to X_u\ell\nu)$	$^{+5.14}_{-6.04}$	$^{+2.53}_{-2.78}$	$^{+1.71}_{-1.88}$	$^{+2.58}_{-2.85}$	$^{+1.38}_{-1.08}$	$^{+2.17}_{-2.25}$	$^{+2.68}_{-3.31}$	$^{+4.07}_{-4.14}$	$^{+0.38}_{-0.28}$	$^{+218.15}_{-189.82}$	$^{+0.28}_{-0.61}$	$^{+0.21}_{-0.25}$	$^{+0.91}_{-1.04}$	$^{+0.76}_{-1.50}$
Continuum	$^{+0.66}_{-0.05}$	$^{+0.55}_{-0.37}$	$^{+0.03}_{-0.08}$	$^{+1.27}_{-0.57}$	$^{+0.03}_{-0.05}$	$^{+0.28}_{-0.11}$	$^{+0.37}_{-0.30}$	$^{+0.21}_{-0.14}$	$^{+0.22}_{-0.11}$	$^{+99.50}_{-99.96}$	$^{+0.01}_{-0.07}$	$^{+0.53}_{-0.30}$	$^{+1.56}_{-0.86}$	$^{+6.03}_{-4.58}$
Rare	±1.38	±1.94	±0.97	±1.68	±0.26	±0.62	±0.41	±0.59	±0.79	±0.00	±0.93	±0.94	±2.94	±2.07
Sec. Leptons	$^{+0.03}_{-0.13}$	$^{+0.19}_{-0.13}$	$^{+0.08}_{-0.01}$	$^{+0.31}_{-0.44}$	$^{+0.01}_{-0.07}$	$^{+0.03}_{-0.03}$	$^{+0.14}_{-0.21}$	$^{+0.36}_{-0.12}$	$^{+1.02}_{-0.69}$	$^{+91.49}_{-87.30}$	$^{+0.14}_{-0.16}$	$^{+0.08}_{-0.26}$	$^{+0.06}_{-0.28}$	$^{+0.97}_{-3.09}$
Fake leptons	$^{+0.23}_{-0.16}$	$^{+0.07}_{-0.05}$	$^{+0.06}_{-0.01}$	$^{+0.32}_{-0.17}$	$^{+0.07}_{-0.00}$	$^{+0.02}_{-0.14}$	$^{+0.07}_{-0.09}$	$^{+0.72}_{-0.70}$	$^{+0.81}_{-0.66}$	$^{+48.55}_{-92.50}$	$^{+0.60}_{-0.66}$	$^{+0.10}_{-0.15}$	$^{+0.35}_{-0.08}$	$^{+1.27}_{-1.29}$
ℓID	±1.86	±1.88	±1.79	±1.80	±1.92	±1.93	±1.89	±1.85	±1.96	±1.87	±1.91	±1.92	±1.91	±2.05
πID	±0.94	±0.96	±0.91	±1.00	±0.98	±1.01	±1.01	±1.01	±1.03	±1.01	±1.03	±1.04	±0.99	±1.19
FSR	±0.34	±0.54	±1.04	±0.65	±0.23	±0.26	±0.20	±0.34	±0.24	±0.00	±0.19	±0.17	±0.64	±1.81
Signal model	±1.63	±5.93	±9.59	±11.95	±10.40	±4.80	±6.93	±9.70	±3.74	±5.10	±5.97	±5.95	±0.88	±11.64
Fit procedure	±0.30	±3.20	±1.16	±0.24	±0.33	±0.15	±0.25	±0.77	±1.54	±83.30	±1.27	±0.75	±1.82	±4.02
Semitotal	$^{+6.47}_{-7.18}$	$^{+7.94}_{-8.01}$	$^{+10.14}_{-10.16}$	$^{+12.65}_{-12.67}$	$^{+10.73}_{-10.70}$	$^{+5.78}_{-5.81}$	$^{+7.79}_{-8.03}$	$^{+10.93}_{-10.95}$	$^{+5.34}_{-5.38}$	$^{+436.10}_{-406.71}$	$^{+6.64}_{-6.65}$	$^{+6.49}_{-6.48}$	$^{+4.66}_{-4.48}$	$^{+15.58}_{-15.79}$

Table A.6: Relative systematic uncertainties in percent for the fits in $M_{\pi\pi}$ bins, for a sample containing only negatively charged B mesons ($B^- \to \pi^+\pi^-\ell^-\nu_\ell$).

Source	bin 1	bin 2	bin 3	bin 4	bin 5	bin 6	bin 7	bin 8	bin 9	bin 10	bin 11	bin 12	bin 13	bin 14
FF $B \to D^{(*)}\ell\nu$	$^{+0.21}_{-0.07}$	$^{+0.25}_{-0.33}$	$^{+0.51}_{-0.40}$	$^{+0.11}_{-0.21}$	$^{+0.04}_{-0.09}$	$^{+0.08}_{-0.10}$	$^{+0.08}_{-0.07}$	$^{+0.17}_{-0.22}$	$^{+0.09}_{-0.08}$	$^{+0.14}_{-0.25}$	$^{+0.15}_{-0.09}$	$^{+0.18}_{-0.24}$	$^{+0.22}_{-0.78}$	$^{+22.74}_{-10.61}$
FF $B \to D^{**}\ell\nu$	$^{+0.45}_{-0.45}$	$^{+0.55}_{-0.55}$	$^{+0.65}_{-0.65}$	$^{+0.32}_{-0.32}$	$^{+0.27}_{-0.27}$	$^{+0.09}_{-0.09}$	$^{+0.25}_{-0.25}$	$^{+0.19}_{-0.19}$	$^{+2.93}_{-2.93}$	$^{+0.13}_{-0.13}$	$^{+0.33}_{-0.33}$	$^{+0.14}_{-0.14}$	$^{+0.38}_{-0.38}$	$^{+19.84}_{-19.84}$
Shapes $B \to X_u\ell\nu$	±3.52	±2.23	±1.42	±1.88	±0.16	±0.49	±0.55	±0.52	±0.23	±0.32	±0.30	±0.17	±0.45	±54.64
$\mathcal{B}(B \to D^{(*)}\ell\nu)$	$^{+0.45}_{-1.00}$	$^{+0.69}_{-0.63}$	$^{+0.58}_{-0.39}$	$^{+0.32}_{-0.09}$	$^{+0.09}_{-0.12}$	$^{+0.08}_{-0.10}$	$^{+0.12}_{-0.14}$	$^{+0.37}_{-0.14}$	$^{+0.21}_{-0.23}$	$^{+0.27}_{-0.34}$	$^{+0.53}_{-0.09}$	$^{+0.40}_{-0.73}$	$^{+0.33}_{-0.38}$	$^{+100.98}_{-113.79}$
$\mathcal{B}(B \to D^{**}\ell\nu)$	$^{+0.76}_{-0.79}$	$^{+0.54}_{-0.64}$	$^{+0.60}_{-0.75}$	$^{+0.17}_{-0.10}$	$^{+0.07}_{-0.06}$	$^{+0.14}_{-0.08}$	$^{+0.27}_{-0.07}$	$^{+0.39}_{-0.14}$	$^{+1.09}_{-0.83}$	$^{+0.32}_{-0.33}$	$^{+0.24}_{-0.22}$	$^{+0.77}_{-0.34}$	$^{+0.36}_{-0.59}$	$^{+24.29}_{-98.97}$
$\mathcal{B}(B \to X_u\ell\nu)$	$^{+9.96}_{-9.52}$	$^{+3.98}_{-3.46}$	$^{+6.16}_{-8.20}$	$^{+3.91}_{-4.32}$	$^{+0.70}_{-0.69}$	$^{+1.70}_{-1.46}$	$^{+1.24}_{-1.74}$	$^{+3.16}_{-3.97}$	$^{+2.10}_{-2.08}$	$^{+0.38}_{-0.08}$	$^{+0.38}_{-0.80}$	$^{+0.28}_{-0.96}$	$^{+0.45}_{-0.31}$	$^{+15.53}_{-22.28}$
Continuum	$^{+1.64}_{-0.34}$	$^{+0.94}_{-0.06}$	$^{+0.15}_{-0.39}$	$^{+0.08}_{-0.02}$	$^{+0.27}_{-0.11}$	$^{+0.04}_{-0.02}$	$^{+0.12}_{-0.14}$	$^{+2.88}_{-1.39}$	$^{+0.07}_{-0.09}$	$^{+0.72}_{-0.45}$	$^{+0.21}_{-0.08}$	$^{+0.20}_{-0.16}$	$^{+0.33}_{-0.43}$	$^{+44.13}_{-8.56}$
Rare	±0.92	±1.18	±1.23	±0.91	±0.31	±0.35	±1.25	±0.84	±0.78	±1.59	±0.96	±0.44	±2.23	±0.01
Sec. Leptons	$^{+0.20}_{-0.39}$	$^{+0.36}_{-0.18}$	$^{+0.20}_{-0.18}$	$^{+0.03}_{-0.09}$	$^{+0.02}_{-0.04}$	$^{+0.01}_{-0.06}$	$^{+0.04}_{-0.11}$	$^{+0.09}_{-0.01}$	$^{+0.02}_{-0.08}$	$^{+0.15}_{-0.11}$	$^{+0.04}_{-0.14}$	$^{+0.13}_{-0.14}$	$^{+0.20}_{-0.26}$	$^{+27.82}_{-11.80}$
Fake leptons	$^{+0.45}_{-0.73}$	$^{+0.05}_{-0.06}$	$^{+0.01}_{-0.24}$	$^{+0.16}_{-0.18}$	$^{+0.01}_{-0.04}$	$^{+0.03}_{-0.11}$	$^{+0.11}_{-0.11}$	$^{+1.11}_{-1.08}$	$^{+0.02}_{-0.11}$	$^{+0.07}_{-0.07}$	$^{+0.30}_{-0.52}$	$^{+0.02}_{-0.08}$	$^{+0.19}_{-0.11}$	$^{+37.87}_{-37.58}$
ℓID	±1.83	±1.92	±2.02	±1.94	±1.93	±1.86	±1.97	±1.82	±1.83	±1.83	±1.86	±1.85	±1.91	±2.00
πID	±1.03	±1.01	±0.99	±1.00	±0.98	±0.99	±1.02	±0.98	±0.99	±1.01	±1.06	±1.01	±1.05	±1.18
FSR	±0.18	±0.24	±0.25	±0.18	±0.06	±0.07	±0.25	±0.17	±0.16	±0.32	±0.19	±0.09	±0.45	±0.05
Signal model	±1.63	±5.93	±9.28	±11.70	±7.13	±5.56	±7.42	±6.68	±2.99	±4.40	±4.30	±5.65	±0.09	±8.77
Fit procedure	±0.73	±0.51	±0.27	±0.25	±0.21	±0.50	±0.54	±0.83	±0.08	±0.88	±0.30	±0.05	±0.79	±72.00
Semitotal	$^{+11.14}_{-10.68}$	$^{+8.03}_{-7.73}$	$^{+11.59}_{-12.79}$	$^{+12.72}_{-12.84}$	$^{+7.51}_{-7.50}$	$^{+6.24}_{-6.18}$	$^{+7.99}_{-8.08}$	$^{+8.39}_{-8.34}$	$^{+5.31}_{-5.26}$	$^{+5.30}_{-5.27}$	$^{+5.00}_{-5.03}$	$^{+6.13}_{-6.19}$	$^{+3.40}_{-3.52}$	$^{+156.03}_{-183.36}$

Table A.7: Relative systematic uncertainties in percent for the fits in $M_{\pi\pi}$ bins, for a sample containing only positively charged B mesons ($B^+ \to \pi^+\pi^-\ell^+\nu_\ell$).

Source	bin 1	bin 2	bin 3	bin 4	bin 5	bin 6	bin 7	bin 8	bin 9	bin 10	bin 11	bin 12	bin 13	bin 14
FF $B \to D^{(*)}\ell\nu$	$^{+0.10}_{-0.18}$	$^{+0.21}_{-0.16}$	$^{+0.11}_{-0.68}$	$^{+0.13}_{-0.15}$	$^{+0.04}_{-0.05}$	$^{+0.14}_{-0.12}$	$^{+0.05}_{-0.19}$	$^{+0.14}_{-0.24}$	$^{+0.07}_{-0.04}$	$^{+0.27}_{-0.29}$	$^{+0.08}_{-0.08}$	$^{+0.22}_{-0.21}$	$^{+0.27}_{-0.31}$	$^{+0.16}_{-2.18}$
FF $B \to D^{**}\ell\nu$	$^{+0.59}_{-0.59}$	$^{+0.35}_{-0.35}$	$^{+0.41}_{-0.41}$	$^{+0.12}_{-0.12}$	$^{+0.09}_{-0.09}$	$^{+0.54}_{-0.54}$	$^{+0.47}_{-0.47}$	$^{+0.34}_{-0.34}$	$^{+4.72}_{-4.72}$	$^{+0.36}_{-0.36}$	$^{+0.16}_{-0.16}$	$^{+0.26}_{-0.26}$	$^{+0.28}_{-0.28}$	$^{+1.79}_{-1.79}$
Shapes $B \to X_u\ell\nu$	±1.65	±0.77	±0.37	±0.59	±0.21	±0.63	±0.31	±1.62	±0.26	±4.39	±0.33	±0.11	±0.26	±0.49
$\mathcal{B}(B \to D^{(*)}\ell\nu)$	$^{+0.16}_{-0.01}$	$^{+0.01}_{-0.02}$	$^{+0.35}_{-0.08}$	$^{+0.13}_{-0.12}$	$^{+0.04}_{-0.07}$	$^{+0.13}_{-0.13}$	$^{+0.06}_{-0.06}$	$^{+0.22}_{-0.17}$	$^{+0.15}_{-0.14}$	$^{+0.35}_{-0.36}$	$^{+0.18}_{-0.32}$	$^{+0.12}_{-0.14}$	$^{+0.28}_{-0.34}$	$^{+8.29}_{-8.44}$
$\mathcal{B}(B \to D^{**}\ell\nu)$	$^{+0.25}_{-0.21}$	$^{+0.07}_{-0.24}$	$^{+0.36}_{-0.31}$	$^{+0.19}_{-0.13}$	$^{+0.05}_{-0.06}$	$^{+0.10}_{-0.13}$	$^{+0.17}_{-0.19}$	$^{+0.18}_{-0.10}$	$^{+1.01}_{-1.12}$	$^{+0.40}_{-0.35}$	$^{+0.26}_{-0.30}$	$^{+0.36}_{-0.05}$	$^{+0.45}_{-0.52}$	$^{+0.33}_{-0.38}$
$\mathcal{B}(B \to X_u\ell\nu)$	$^{+5.14}_{-5.58}$	$^{+3.36}_{-3.30}$	$^{+2.65}_{-2.80}$	$^{+2.92}_{-3.46}$	$^{+0.64}_{-0.78}$	$^{+1.97}_{-2.05}$	$^{+2.19}_{-2.72}$	$^{+6.29}_{-9.20}$	$^{+0.98}_{-1.72}$	$^{+1.13}_{-0.71}$	$^{+1.48}_{-1.65}$	$^{+0.33}_{-0.28}$	$^{+1.53}_{-1.82}$	$^{+0.24}_{-0.36}$
Continuum	$^{+0.04}_{-0.02}$	$^{+0.43}_{-0.02}$	$^{+0.04}_{-0.02}$	$^{+0.45}_{-0.05}$	$^{+0.07}_{-0.04}$	$^{+0.00}_{-0.01}$	$^{+0.55}_{-0.28}$	$^{+2.19}_{-0.61}$	$^{+0.02}_{-0.04}$	$^{+0.04}_{-0.24}$	$^{+0.05}_{-0.10}$	$^{+0.05}_{-0.01}$	$^{+0.37}_{-0.41}$	$^{+5.03}_{-3.38}$
Rare	±0.05	±0.18	±0.53	±0.30	±0.05	±0.06	±0.01	±0.21	±0.54	±0.19	±0.90	±0.76	±1.87	±0.08
Sec. Leptons	$^{+0.02}_{-0.09}$	$^{+0.03}_{-0.12}$	$^{+0.04}_{-0.04}$	$^{+0.12}_{-0.16}$	$^{+0.10}_{-0.12}$	$^{+0.05}_{-0.05}$	$^{+0.08}_{-0.00}$	$^{+0.08}_{-0.04}$	$^{+0.04}_{-0.04}$	$^{+0.21}_{-0.56}$	$^{+0.18}_{-0.10}$	$^{+0.23}_{-0.01}$	$^{+0.03}_{-0.22}$	$^{+1.04}_{-1.19}$
Fake leptons	$^{+0.38}_{-0.80}$	$^{+0.09}_{-0.07}$	$^{+0.03}_{-0.25}$	$^{+0.11}_{-0.02}$	$^{+0.01}_{-0.02}$	$^{+0.25}_{-0.03}$	$^{+0.04}_{-0.06}$	$^{+1.70}_{-1.58}$	$^{+0.03}_{-0.04}$	$^{+0.19}_{-0.09}$	$^{+0.19}_{-0.41}$	$^{+0.21}_{-0.02}$	$^{+0.22}_{-0.26}$	$^{+1.41}_{-1.36}$
ℓID	±1.77	±1.86	±1.75	±1.75	±1.77	±1.77	±1.76	±1.78	±1.78	±1.74	±1.81	±1.86	±1.92	±2.07
πID	±0.98	±0.94	±0.95	±0.94	±0.93	±0.94	±0.97	±0.99	±0.98	±0.95	±1.02	±1.00	±0.98	±1.16
FSR	±0.48	±0.30	±0.20	±0.02	±0.03	±0.01	±0.28	±0.04	±0.19	±0.19	±0.38	±0.86	±0.44	±0.44
Signal model	±5.10	±1.94	±15.24	±12.88	±6.58	±9.72	±6.45	±10.59	±3.20	±0.04	±1.80	±7.91	±5.94	±17.45
Fit procedure	±0.25	±0.03	±0.29	±0.20	±0.06	±0.39	±0.49	±0.35	±0.47	±0.58	±0.28	±0.56	±0.45	±7.29
Semitotal	$^{+7.75}_{-8.08}$	$^{+4.53}_{-4.47}$	$^{+15.63}_{-15.67}$	$^{+13.39}_{-13.51}$	$^{+6.91}_{-6.93}$	$^{+10.17}_{-10.18}$	$^{+7.17}_{-7.34}$	$^{+12.91}_{-14.38}$	$^{+6.27}_{-6.44}$	$^{+5.05}_{-5.00}$	$^{+3.33}_{-3.44}$	$^{+8.31}_{-8.30}$	$^{+6.84}_{-6.93}$	$^{+21.55}_{-21.40}$

Table A.8: Relative systematic uncertainties in percent for the fits in $M_{\pi\pi}$ bins, for a sample containing only electrons ($B^{\pm} \to \pi^+\pi^- e^{\pm}\nu_\ell$).

Source	bin 1	bin 2	bin 3	bin 4	bin 5	bin 6	bin 7	bin 8	bin 9	bin 10	bin 11	bin 12	bin 13	bin 14
FF $B \to D^{(*)}\ell\nu$	$^{+0.71}_{-1.20}$	$^{+0.17}_{-0.24}$	$^{+0.26}_{-0.19}$	$^{+0.29}_{-0.13}$	$^{+0.02}_{-0.06}$	$^{+0.10}_{-0.10}$	$^{+0.46}_{-0.15}$	$^{+0.18}_{-0.16}$	$^{+0.16}_{-0.45}$	$^{+0.27}_{-0.20}$	$^{+0.09}_{-0.04}$	$^{+0.09}_{-0.18}$	$^{+0.29}_{-0.62}$	$^{+4.91}_{-7.16}$
FF $B \to D^{**}\ell\nu$	$^{+2.32}_{-2.32}$	$^{+0.29}_{-0.29}$	$^{+0.16}_{-0.16}$	$^{+1.00}_{-1.00}$	$^{+0.22}_{-0.22}$	$^{+0.09}_{-0.09}$	$^{+0.08}_{-0.08}$	$^{+0.66}_{-0.66}$	$^{+0.50}_{-0.50}$	$^{+0.19}_{-0.19}$	$^{+0.09}_{-0.09}$	$^{+0.28}_{-0.28}$	$^{+0.51}_{-0.51}$	$^{+12.58}_{-12.58}$
Shapes $B \to X_u\ell\nu$	±10.36	±1.46	±0.93	±2.03	±0.17	±0.25	±0.59	±0.84	±0.42	±0.17	±0.03	±0.10	±0.08	±5.19
$\mathcal{B}(B \to D^{(*)}\ell\nu)$	$^{+0.62}_{-1.53}$	$^{+0.22}_{-0.20}$	$^{+0.16}_{-0.42}$	$^{+0.65}_{-0.19}$	$^{+0.08}_{-0.09}$	$^{+0.08}_{-0.07}$	$^{+0.37}_{-0.16}$	$^{+0.22}_{-0.18}$	$^{+0.26}_{-0.30}$	$^{+0.51}_{-0.51}$	$^{+0.11}_{-0.15}$	$^{+0.25}_{-0.12}$	$^{+0.11}_{-0.14}$	$^{+36.18}_{-39.05}$
$\mathcal{B}(B \to D^{**}\ell\nu)$	$^{+0.65}_{-1.02}$	$^{+0.29}_{-0.19}$	$^{+0.27}_{-0.21}$	$^{+0.40}_{-0.20}$	$^{+0.07}_{-0.05}$	$^{+0.09}_{-0.16}$	$^{+0.19}_{-0.36}$	$^{+0.23}_{-0.33}$	$^{+0.57}_{-0.26}$	$^{+0.61}_{-0.23}$	$^{+0.09}_{-0.01}$	$^{+0.19}_{-0.16}$	$^{+0.14}_{-0.65}$	$^{+1.59}_{-1.61}$
$\mathcal{B}(B \to X_u\ell\nu)$	$^{+24.94}_{-24.97}$	$^{+3.60}_{-4.24}$	$^{+3.43}_{-3.73}$	$^{+3.63}_{-3.95}$	$^{+1.03}_{-1.08}$	$^{+1.27}_{-1.45}$	$^{+1.13}_{-0.93}$	$^{+1.53}_{-1.97}$	$^{+0.30}_{-1.05}$	$^{+0.12}_{-0.10}$	$^{+0.04}_{-0.15}$	$^{+0.28}_{-0.43}$	$^{+0.40}_{-0.31}$	$^{+8.03}_{-8.64}$
Continuum	$^{+3.86}_{-1.52}$	$^{+0.14}_{-0.38}$	$^{+0.35}_{-0.11}$	$^{+1.32}_{-0.61}$	$^{+0.23}_{-0.06}$	$^{+0.19}_{-0.12}$	$^{+0.46}_{-0.25}$	$^{+1.63}_{-0.88}$	$^{+0.20}_{-0.14}$	$^{+0.94}_{-0.51}$	$^{+0.03}_{-0.02}$	$^{+0.46}_{-0.14}$	$^{+1.22}_{-0.16}$	$^{+66.67}_{-41.57}$
Rare	±5.37	±2.68	±1.95	±2.76	±0.46	±1.03	±2.23	±1.50	±1.05	±6.15	±0.94	±0.73	±3.24	±5.74
Sec. Leptons	$^{+0.87}_{-0.50}$	$^{+0.04}_{-0.27}$	$^{+0.08}_{-0.05}$	$^{+0.01}_{-0.53}$	$^{+0.02}_{-0.00}$	$^{+0.05}_{-0.06}$	$^{+0.16}_{-0.03}$	$^{+0.02}_{-0.44}$	$^{+0.03}_{-0.19}$	$^{+0.37}_{-0.10}$	$^{+0.01}_{-0.04}$	$^{+0.25}_{-0.35}$	$^{+0.11}_{-0.04}$	$^{+4.63}_{-14.36}$
Fake leptons	$^{+3.35}_{-0.70}$	$^{+0.23}_{-0.02}$	$^{+0.01}_{-0.18}$	$^{+0.46}_{-1.16}$	$^{+0.05}_{-0.08}$	$^{+0.13}_{-0.04}$	$^{+0.09}_{-0.05}$	$^{+0.33}_{-0.23}$	$^{+0.42}_{-0.19}$	$^{+0.61}_{-0.30}$	$^{+0.11}_{-0.13}$	$^{+0.02}_{-0.07}$	$^{+0.28}_{-0.39}$	$^{+13.58}_{-14.86}$
ℓID	±1.94	±1.96	±2.08	±2.04	±2.13	±2.07	±2.14	±1.93	±2.04	±1.99	±2.01	±1.95	±1.94	±1.98
πID	±0.99	±1.05	±0.95	±1.08	±1.04	±1.06	±1.06	±1.00	±1.03	±1.08	±1.08	±1.05	±1.08	±1.23
FSR	±0.02	±0.07	±0.11	±0.03	±0.05	±0.20	±0.27	±0.02	±0.13	±0.81	±0.36	±0.34	±0.16	±0.82
Signal model	±2.00	±3.33	±5.02	±8.46	±11.73	±1.63	±6.83	±5.10	±0.47	±2.62	±7.54	±0.49	±7.13	±1.61
Fit procedure	±2.00	±0.17	±0.12	±0.13	±0.19	±0.45	±0.40	±0.34	±2.13	±0.73	±0.51	±0.00	±0.00	±9.34
Semitotal	$^{+28.36}_{-28.03}$	$^{+6.22}_{-6.62}$	$^{+6.87}_{-7.03}$	$^{+10.27}_{-10.36}$	$^{+12.03}_{-12.03}$	$^{+3.34}_{-3.41}$	$^{+7.73}_{-7.68}$	$^{+6.28}_{-6.27}$	$^{+3.50}_{-3.62}$	$^{+7.29}_{-7.20}$	$^{+7.96}_{-7.96}$	$^{+2.52}_{-2.51}$	$^{+8.27}_{-8.23}$	$^{+79.78}_{-64.21}$

Table A.9: Relative systematic uncertainties in percent for the fits in $M_{\pi\pi}$ bins, for a sample containing only muons ($B^\pm \to \pi^+\pi^-\mu^\pm\nu_\ell$).

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to my supervisor Ariane Frey for giving me the opportunity to do research on B physics in her group. I also would like to thank the members of the CKM working from the Belle collaboration, especially Christoph Schwanda and Kenkiki Miyabayashi for their insightful comments and for refereeing the progress of this analysis. For technical support on topics related to the Belle software, and in some occasions statistics, I also would like to thank Yoshi Sakai and Shohei Nishida.

Special thanks also goes to Jochen Dingfelder for agreeing to co-referee this work, and to the members of my evaluation committee Aida Achkar, Wolfram Kollatchny, Stan Lai and Steffen Schumann. I also would like to thank Christoph Hanhart and Stefan Ropertz for providing me the numerical solution of the scalar form factor that I used in one part of the analysis.

I also would like to thank Chepe, Uwe, Andrew, Damián and Juan Carlos for their help in proofreading this thesis.

Apart from the work on the thesis, I also would like to thank the members of my group for organizing activities in our spare time, I really enjoyed go to Bowling, racing, to the cinema with you guys, Elisabeth, the two Philipps, Ulf, Benjamin.

Last but not least, I would like to thank my parents and siblings for their continuous support and love, and Jasmin for always being there for me.