AGH University of Science and Technology

Faculty of Physics and Applied Computer Science

Master thesis

Mateusz Kaleta

major: **technical physics**

Search for lepton flavour violating B meson decays at the Belle experiment

Supervisor: **Dr hab. Andrzej Bożek**

Kraków, November 2020

Aware of criminal liability for making untrue statements I decleare that the following thesis was written personally by myself and that I did not use any sources but the ones mentioned in the dissertation itself.

The subject of the master thesis and the internship by Mateusz Kaleta, student of 5th year major in technical physics

The subject of the master thesis: **Search for lepton flavour violating B meson decays at the Belle experiment**

Supervisor:	dr hab. Andrzej Bożek
Reviewer:	
A place of the internship:	Institute of Nuclear Physics Polish Academy of Sciences, Krakow, Poland

Programme of the master thesis and the internship

1. First discussion with the supervisor on realization of the thesis.
2. Collecting and studying the references relevant to the thesis topic(s).
3. The internship.
   - getting familiar with Belle (II) experiment software,
   - discussion with the supervisor focused on implemented measurement technique and obtained results,
   - preparation of the internship report.
4. Continuation of calculations concerning the thesis subject.
5. Final analysis of the results obtained, conclusions – discussion with and final approval by the thesis supervisor.
6. Typesetting the thesis.

Dean's office delivery deadline: November 2020

Supervisor's review

Reviewers's review

Contents

1. Introduction

The Standard Model (SM) of particle physics, formulated in the early 70s of of the last century, is the current best theory describing the fundamental constituents of matter and the forces acting between them. During the last half of a century, the Model has been put to the multitude of experimental tests and has passed most of them with flying colours. Despite its enormous success, the Model is not believed to be a complete theory. For instance, it does not include gravity and introduces a large number of free parameters, whose numerical values are established by experiment. Experimentally measured quark and lepton masses, together with the Cabibbo-Kobayashi-Maskawa (CKM) matrix [1, 2] elements show hierarchical structure that may point to underlying dynamics which is reflected by those parameters, but not introduced explicitly. All this makes that the Standard Model is considered to be an effective theory, rather than a complete description of Nature. Therefore, efforts have been made over the years to study extended theories, commonly refered to as the physics Beyond the Standard Model (BSM) or simply the New Physics (NP).

Experimentally, the New Physics can be hunted using two complementary approaches. The former one, known as the Energy Frontier is represented by experiments such as ATLAS and CMS conducted at CERN. Those experiments aim for the highest possible center-of-mass energy of the collision, sufficient to directly produce heavy particles predicted by many BSM models. New particles can also be created off-shell, as internal propagators. This allows to study their properties at energies much below the production threshhold, for instance in rare decays of SM particles. Such measurements require large statistics in order to push the experimental precision to the limits and represent the so-called Intensity Frontier.

In the original formulation of the Standard Model with massless neutrinos lepton flavour is a conserved quantity. Today we know this assumption does not hold in the neutrino sector, as the compelling evidence for transitions between different neutrino flavours, ν_e, ν_μ, ν_τ, known as neutrino oscillations, has been provided over the years by many experiments. Lepton Flavour Violation (LFV) in charged lepton sector can occur through processes that involve neutrino mixing [3], but these are highly suppressed and have predicted branching fractions many orders of magnitude below the experimental sensitivity of today's and future experiments.

Decays of B mesons, the lightest hadrons containing b quark, offer a unique laboratory for indirect NP searches. In particular, semileptonic B meson decays have recently drawn much attention, as many experimental hints for Lepton Flavour Universality (LFU) violation in these processes have been provided in recent years. [4–11]. Many NP models, that aim to describe these anomalies, can also lead to relatively large branching fractions for LFV B meson decays. Especially promising are processes involving heavier leptons, such as $B \to K\tau^{\pm}\mu^{\mp}$, where some of the models (e.g., [12, 13]) predict the rates for these decays to be within experimental sensitivity.

Experimentally, searches involving τ leptons typically require more advanced measurement techniques to overcome the challenges caused by missing neutrinos and lack of distincive τ signature. The lack of complete kinematic information in B meson decays to final state with τ leptons renders it impossible to completely reconstruct the B meson decay chain. Due to these difficulties, the experimental information on $B \to K\tau^{\pm}\mu^{m}p$ processes have been weak or absent to date.

$B \to K\tau^{\pm}\mu^{\mp}$ processes were first searched by the BaBAR collaboration [14, 15]. The collaboration found no evidence for these decays and set an upper limit at the level of a few times 10^{-5}. For almost ten years, these results were the only experimental limits for $B \to K\tau^{\pm}\mu^{\mp}$ processes. In 2020, LHCb has reported their first search for $B^+ \to K^+\tau^+\mu^-$ decay [16], setting an upper on its branching fraction slightly above the one obtained by BaBAR. The current 90% upper limits on $B^+ \to K^+\tau^+\mu^-$ processes are summarised in Table 1.1.

Table 1.1. Summary of the results for lepton flavour violating modes.

Experiment (year)	Decay mode	$\mathcal{B}^{\mathrm{UL}}$ (90% C.L.)
BaBar (2012) [15]	$B^+ \to K^+\tau^+\mu^-$	2.8×10^{-5}
	$B^+ \to K^+\tau^-\mu^+$	4.5×10^{-5}
	$B^+ \to K^+\tau^{\pm}\mu^{\mp}$	4.8×10^{-5}
LHCb (2020) [16]	$B^+ \to K^+\tau^+\mu^-$	3.9×10^{-5}

The aim of this thesis is concerned about the search for lepton flavour violating $B^+ \to K^+\tau^{\pm}\mu^{\mp}$ decays at the Belle experiment. The content of this thesis is organised as follows. chapter 2 presents the physics environment of the Belle experiment as well as a short description of main detector components, important from the data analysis standpoint. chapter 3 presents the measurement technique and contains the main analysis overview. The main analysis chain, applied to Monte Carlo simulated events is presented in chapter 4. chapter 5 describes the fitting procedure that will be used to extract the signal yield. A thesis summary and concluding remarks are contained in chapter 6.

2. Experimental apparatus

2.1. Experiments at the B Factories

The analyses presented in this thesis were performed with the use of simulated events dedicated for Belle experiment, that was operated in 1999-2010 at KEK (jap. Kō Enerugī Kasokuki Kenkyū Kikō, en. The High Energy Accelerator Research Organization) laboratory in Tsukuba, Japan. The Belle experiment (together with the BaBAR experiment, located at SLAC National Accelerator Laboratory in California, USA) was one of the so-called B-Factories; facilities designed to produce large number of B mesons in a clean environment. This goal is obtained by precisely tuning beam energies, so that the center-of-mass energy of the collision, $\sqrt{s}$, corresponds to the mass of the $\Upsilon(4S)$ resonance, $m_{\Upsilon(4S)} \simeq 10.58\text{GeV}$. Therefore, in every collision, a large number of $\Upsilon(4S)$ states is produced, which then quickly decay to a pair of B meson and its antiparticle:

$$e^+e^- \to \Upsilon(4S) \to B\overline{B} \tag{2.1}$$

The specific design of B factories offers many advantages that are characteristic to those devices. Among them are:

- Good signal-to-background ratio

  Figure 2.1 shows the hadronic cross-section for e^+e^- collisions in the region of $\Upsilon(nS)$ resonances. The $\Upsilon(4S)$ production cross-section at e^+e^- collisions is $\sigma(e^+e^- \to \Upsilon(4S) \simeq 1.1$ nb. The main source of background events comes from the so-called continuum processes, where e^+e^- system annihilates into a pair of light quarks, $e^+e^- \to q\overline{q}$ $(q = u, d, s, c)$, with the corresponding cross-section $\sigma(e^+e^- \to q\overline{q}) \simeq 3.4$ nb. This source of background, however, is characterised by a very unique, jet-like topology and can be efficiently suppressed in the off-line analysis.

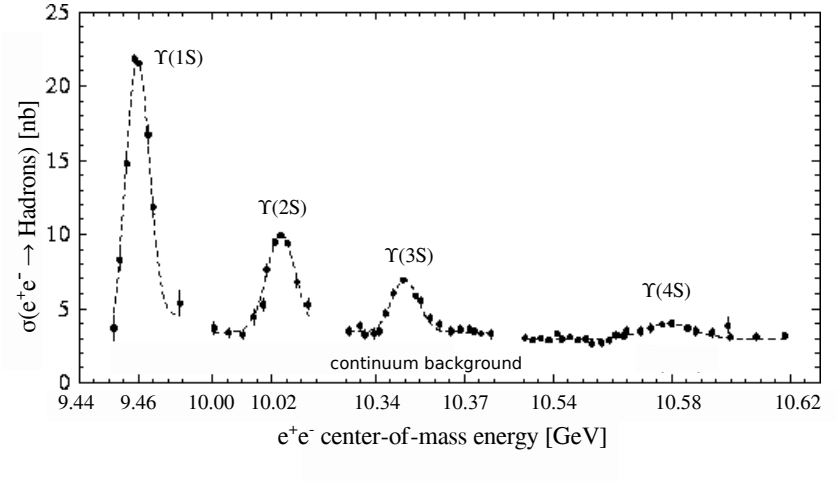


Figure 2.1. Hadronic cross-section for e^+e^- collisions as a function of e^+e^- center-of-mass energy in the region of Υ(nS) (n $\in \{1, 2, 3, 4\}$) resonances [17].

- Clean environment

  Υ(4S) is a vector state with assigned quantum numbers $J^{PC} = 1^{--}$, that can be identified as the fourth radial excitation level of the $b\bar{b}$ system. It has a mass $m_{\Upsilon(4S)} = (10.5794 \pm 0.0012)$ GeV and width $\Gamma = (20.5 \pm 2.5)$ MeV [18]. It is therefore the first (i.e., the lightest) bottomonium state that is above the threshhold for B meson pair production and decays predominantly ($\mathcal{B}(\Upsilon(4S) \to B\bar{B}) \approx 96\%$) to $B\bar{B}$ pairs. Due to limited phase space in the $\Upsilon(4S) \to B\bar{B}$ decay, $B\bar{B}$ pairs are produced exclusively, i.e., without other particles.

- Well-defined kinematical constraints

  The experimental setup of B Factories provides kinematical constraints that can be exploited to identify produced B mesons. In the center-of-mass system (CMS)[1] of the e^+e^- collision, each produced B meson has energy equal to the CMS beam energy, which corresponds to half the energy of the collision:

$$E_B = E_{beam} = \sqrt{s}/2. \tag{2.2}$$

  This allows to introduce two kinematical variables characteristic to B Factories, the beam-energy constrained mass (M_{bc}) and the energy difference (ΔE):

$$M_{bc} = \sqrt{E_{beam}^2 - (\mathbf{p}_B^{rec})^2}, \tag{2.3}$$

$$\Delta E = E_B^{rec} - E_{beam}, \tag{2.4}$$

[1] Throughout the text, all kinematical variables will be expressed in CMS, if not stated otherwise.

where $\mathbf{p}_{\mathrm{B}}^{\mathrm{rec}}$ and $\mathrm{E}_{\mathrm{B}}^{\mathrm{rec}}$ are the reconstructed momentum and energy of the B meson, respectively. Correctly reconstructed B meson candidates lie in the region $\mathrm{M_{bc}} \approx 5.279\mathrm{GeV}$, which is the nominal mass of the B meson. The experimental $\mathrm{M_{bc}}$ resolution, $\sigma_{\mathrm{M_{bc}}} \approx 3\mathrm{MeV}$, is mostly determined by the beam energy measurement precision, and is a few times better than the invariant mass of the B meson defined in a standard way. Unlike $\mathrm{M_{bc}}$, that depends only on the reconstructed momentum, $\Delta\mathrm{E}$ is by construction sensitive to particle mass hypotheses. For correctly reconstructed B candidates, where all final state particles were correctly identified, $\Delta\mathrm{E}$ takes values close to zero, which indicates that the reconstructed energy is close to the nominal one. Incorrect particle identification causes $\Delta\mathrm{E}$ distribution to shift towards negative or positive values. The typical $\Delta\mathrm{E}$ resolution varies between $\sigma_{\Delta\mathrm{E}} \approx 15\mathrm{MeV}$ and $\sigma_{\Delta\mathrm{E}} \approx 40\mathrm{MeV}$, depending on the B meson decay mode.

In case of B meson decays with one missing particle of known mass, although the complete decay reconstruction is no longer possible, some limited information about B meson flight direction is still available. In particular, the cosine of the angle between the inferred direction of the B meson momentum ($\mathbf{p}_B$) and the visible momentum of the decay ($\mathbf{p}_{vis}$) is given by:

$$\cos\angle(\mathbf{p}_{\mathrm{B}}, \mathbf{p}_{\mathrm{vis}}) = \frac{2\mathrm{E_{beam}E_{vis}} - \mathrm{m_B^2} - \mathrm{m_{vis}^2} + \mathrm{m_{mis}^2}}{2|\mathbf{p}_{\mathrm{B}}||\mathbf{p}_{\mathrm{vis}}|}, \tag{2.5}$$

where $|\mathbf{p}_B| = \sqrt{E_{beam}^2 - m_B^2}$, E_{vis} is the visible energy of the decay, $m_{vis}^2 = E_{vis}^2 - \mathbf{p}_{vis}^2$ and m_{mis} is the assumed mass of the missing particle. For correctly reconstructed signal decays, $\mathbf{p}_B$ lies on the cone around $\mathbf{p}_{vis}$ with an opening angle $2\angle(\mathbf{p}_B, \mathbf{p}_{vis})$ and values of $\cos\angle(\mathbf{p}_B, \mathbf{p}_{vis})$ are constrained to physical region $[-1, 1]$. The variable (2.5) is often used to identify semileptonic $B \to \bar{D}^{(*)}\ell^+\nu_\ell$ ($\ell = e, \mu$) decays, for which $\mathbf{p}_{vis} = \mathbf{p}_{D^{(*)}} + \mathbf{p}_\ell$, $E_{vis} = E_{D^{(*)}} + E_\ell$ and $m_{mis} = 0$.

- Vertex separation

  The $\Upsilon(4\mathrm{S}) \to \mathrm{B\bar{B}}$ process is characterised by a small Q value of 200MeV. Therefore, $\mathrm{B\bar{B}}$ pairs are produced almost at rest in the CM frame, with momenta of about 330MeV and decay lengths of 30μm, too small to be measured. Since the B Factories were designed for testing time-dependent CP violation, it was necessary to separate two B meson vertices. This was adressed by introducing asymmetric beam energies; $\mathrm{E_{e^-}} = 8$ GeV, $\mathrm{E_{e^+}} = 3.5$ GeV for electron and positron beam, respectively. In such setup, the $\Upsilon(4\mathrm{S})$ state is produced in flight, with the momentum in the laboratory frame $|\mathbf{p}_{\mathrm{lab}}| \approx (\mathrm{E_{e^-}} - \mathrm{E_{e^+}}) = 4.5\mathrm{GeV}$. This results in larger B meson momenta in that frame and larger flight distances. As a result, the distance between two B meson vertices along the beam axis becomes large enough ($\Delta\mathrm{z} \approx 200\mu$m) to be accesible experimentally.

2.2. Tagging methods

B meson decays to a final state with neutrinos are extremely challenging experimentally as they offer little to no kinematic constraints usually exploited to separate signal from background processes. To study processes with missing four-momentum, B factories have developed tagging methods that rely on the fact that $B\bar{B}$ pairs are produced exclusively in the $\Upsilon(4S) \to B\bar{B}$ process (see section 2.1). Reconstructing one of the B meson, reffered to as the "tag B" (B_{tag}) allows to:

- remove large fraction of combinatorial background from $B\bar{B}$ and continuum processes,
- obtain information about the "signal B" (B_{sig}) such as its momentum and charge.

Over the years, two different variations of tagging algorithms were exploited at Belle; the inclusive and the exclusive tagging. A brief summary of both methods is given below. A more complete description can be found in [19].

2.2.1. Exclusive tagging

The principle idea of the exclusive tagging method is to reconstruct B_{tag} in one of many hadronic decay channels. In such case, the efficiency of B_{tag} reconstruction is given by:

$$\epsilon_{tag} = \sum_{i}^{N} \epsilon_i \mathcal{B}_i, \tag{2.6}$$

where index i runs over all considered B_{tag} decay channels and ϵ_i, $\mathcal{B}_i$ are respectively the reconstruction efficiency and branching fraction of a particular B_{tag} decay channel. Although hadronic decays cover around 80% of B meson decay width, the individual branching fractions $\mathcal{B}_i$ are very low, typically ranging from 10^{-3} to 10^{-5}. The typical values of ϵ_i are in the order of 10%.

The overall reconstruction efficiency can thus be maximised by introducing as many as possible B_{tag} decay channels and increasing ϵ_i. At Belle, this was done through the Full Reconstruction package that provided the B_{tag} reconstruction efficiency ranging from 0.1% to 0.3% depending on the desired purity [20].

2.2.2. Inclusive tagging

To further increase the tagging efficiency, Belle developed an alternative approach, called the inclusive B_{tag} reconstruction. The main difference with respect to exlusive approach is that here B_{tag} and B_{sig} are reconstructed in reverse order. The signal side is reconstructed first and, in the second step, all remaining tracks and clusters are combined together to form the B_{tag} candidate. This way, B_{tag} is reconstructed in all possible final states, however without explicitly choosing its decay channel. Thus, the method can provide higher B_{tag} reconstruction efficiency at a cost of

lower purity. The consistency of inclusively reconstructed B_{tag} with a B meson decay is checked using standard variables; the beam-energy constrained mass (M_{tag}) and the energy difference $(\Delta E_{tag})^2$:

$$M_{tag} = \sqrt{E_{beam}^2 - (\mathbf{p}_{tag}^{rec})^2} \tag{2.7}$$

$$\Delta E_{tag} = E_{tag}^{rec} - E_{beam} \tag{2.8}$$

where $\mathbf{p}_{tag}^{rec} \equiv \sum_i \mathbf{p}_i$, $E_{tag}^{rec} \equiv \sum_i E_i$ are respectively the sum of momenta and energies of reconstructed particles that were not assigned to B_{sig}. Example distributions of M_{tag} and ΔE_{tag} for $B_{sig}^0 \to D^{*-}\pi^+$ control sample are shown in Figure 2.2.

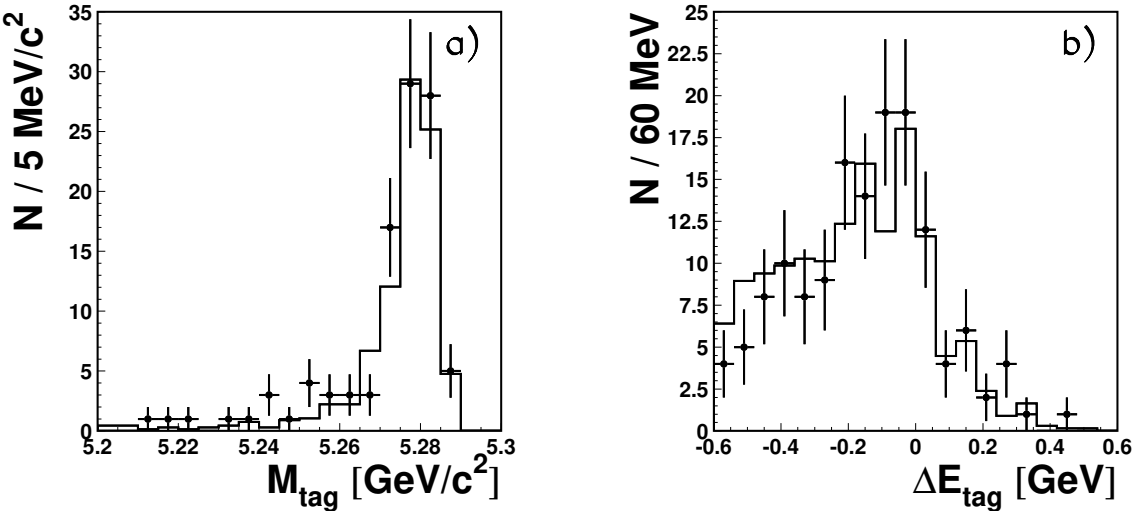


Figure 2.2. Distributions of M_{tag} in window -0.25 GeV $< \Delta E_{tag} < 0.05$ GeV and ΔE_{tag} in window $M_{tag} > 5.27$ GeV for $B_{sig}^0 \to D^{*-}\pi^+$ control sample from data (points with error bars) and MC (histograms) [21].

Since the method suffers from higher background and there are no well established tools (like Full Reconstruction) available, inclusive tagging is, in general, harder to apply and cannot be used in every analysis. In particular, the method hardly relies on a clean B_{sig} signature. Nevertheless, it benefits from higher reconstruction efficiency and can provide larger sensitivity for rare processes. In fact, the first application of the inclusive tagging at Belle [21] lead to the first observation of semitauonic $B^0 \to D^{*-}\tau^+\nu_\tau$ decay. The follow up studies [22, 23] provided the first evidence for $B^+ \to \bar{D}^0\tau^+\nu_\tau$ as well as the first D^{*-} polarization measurements in $B^0 \to D^{*-}\tau^+\nu_\tau$ decay.

[2]Variables M_{tag} and ΔE_{tag} are equivalent to variables M_{bc} and ΔE (Equation 2.3-2.4), but defined for tag side.

2.3. The KEKB accelerator

The electron and positron beams for the Belle experiment were provided by the KEKB accelerator [24, 25]. Two separate storage rings: HER (high-energy storage ring) and LER (low-energy storage ring) were continuously injected with the electron and positron beams, previously accelerated in Linear Accelerator (Linac) to their full energies[3]. The beams were brought to collision with a crossing angle of 22 mrad at the interaction point (IP) surrounded by the Belle detector. The choice of finite beam crossing angle allowed to eliminate parasitic collisions between incoming and outgoing bunches.

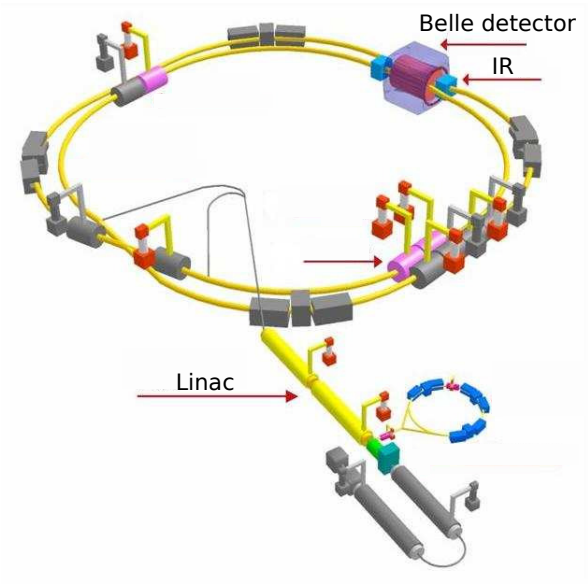


Figure 2.3. A schematic view of the KEKB accelerator.

An important feature an e^+e^- collider is its instantaneous luminosity, that can be expressed as:

$$\mathcal{L} = f\frac{N_b n_{e^-} n_{e^+}}{A} \tag{2.9}$$

where $n_{e^-}(n_{e^+})$ are the numbers of electrons(positrons) in each bunch, N_b is the number of bunches, f - the circulation frequency and A represents the cross-sectional overlapping area of the beams at the interaction point. By squishing beams to very small sizes at the interaction point ($\approx 100\mu m$ in vertical and $\approx 3\mu m$ in horizontal direction) and operating at high beam currents, KEKB was able to provide the typical luminosity of the order of 10^{34} cm^{-2}s^{-1}. In June 2009, it recorded the peak instantaneous luminosity of 2.1×10^{34} cm^{-2} s^{-1}, claiming the world record at that time.

[3]The continuous injection mode started after Liniac upgrade in 2004.

2.4. The Belle detector

The Belle detector [26] was a multifunctional spectrometer composed of several sub-detector systems. The detector layout is shown schematically in Figure 2.4. From the interaction point: the Silicon Vertex Detector (SVD), the Central Drift Chamber (CDC), the Time Of Flight Counter (TOF), the Aerogel Cherenkov Counter (ACC), the Electromagnetic Calorimeter (ECL) and the K_L^0 and Muon Detector (KLM). All inner subdetectors (SVD, CDC, TOF, ACC and ECL) were localised inside the superconducting solenoid that provided a magnetic field of 1.5 T.

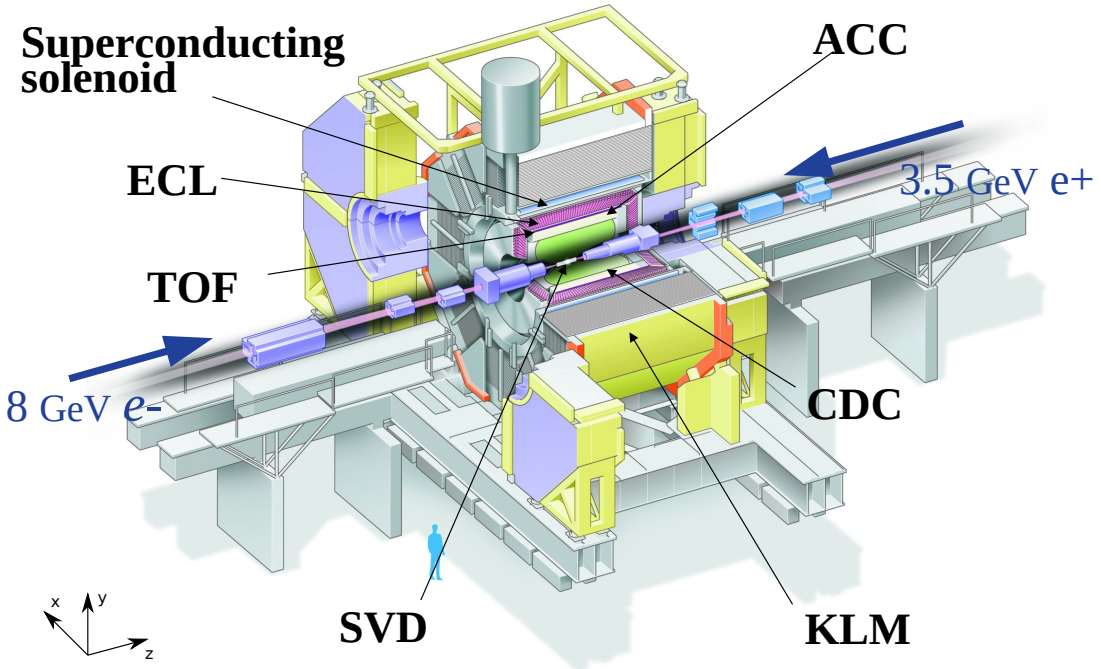


Figure 2.4. A schematic view of the Belle detector.

The coordinates of the Belle detector are represented in the right-handed coordinate system with the origin at the interaction point, where the z-axis is pointing in the direction opposite to e^+ beam and y-axis is pointing upwards. Due to a cylindrical shape of the device, it is useful to define additional coordinates: the polar angle θ measured from the positive z-axis, the azimuthal angle ϕ measured from the positive x-axis and the radial distance in xy-plane $r = \sqrt{x^2 + y^2}$.

SVD The SVD was the innermost tracking device, located in the vicinity of the interaction point. The first version of SVD (SVD1) consisted of three layers of radii 30 mm, 45 mm and 60 mm and covered the polar angle $23° < \theta < 139°$. Each layer was composed of Double-Sided Silicon Strip Detectors (DSSD). These are n-type bulk plates segmented with $p+$ and $n+$ strips on each side. The $p+$ and $n+$ strips were located in directions parallel and perpendicular to the beam line, and thus provided measurements in $r\phi$ and z direction. In 2003, the the detector was upgraded to a four-layer system (SVD2). The radii of SVD2 layers were 20 mm, 44 mm, 70 mm and 80mm and the polar angle acceptance increased to $17° < \theta < 150°$. In addition, the SVD2 was read out by a newly-developed chip, VA1TA, with improved radiation hardness needed to cope with increasing accelerator luminosity.

CDC The CDC was a large-volume device located at the central part of the Belle detector. It had cylindrical shape, with its inner radius of 8 cm and the outer radius of 88 cm. The polar acceptance of CDC was the same as for SVD2, i.e., $17° < \theta < 150°$. The detector played three important roles. First, by measuring track hit coordinates, it provided tracking information complementary to SVD. Second, it determined particle momenta from track curvature in the magnetic field. Moreover it contributed to particle identification by providing measurements of dE/dx within its gas volume. This was particularily important for low momentum tracks, which did not reach the outer sub-detectors and could only be identified by CDC alone. The detector volume was filled with the low Z gas mixture (He and C_2H_6 in equal proportions) that ensured the minimisation of multiple scattering thus providing the optimal momentum resolution.

ACC The ACC utilised the Cherenkov effect to provide particle identification information for charged hadrons (K, p, π) in momentum range from 1 GeV to 3.5 GeV. A particle travelling in a medium with a velocity larger than the phase velocity of light in that medium would cause the emmision of Cherenkov radiation. Mathematically, the condition for the Cerenkov effect to occur can be expressed as:

$$n > \frac{1}{\beta} = \sqrt{1 + \left(\frac{m}{p}\right)^2}. \tag{2.10}$$

where $\beta \equiv v/c$, n is the refractive index of a medium and v, p, m the particle velocity, momentum and mass, respectively. From the above expression it can be seen that for a given refractive index and momentum range lighter particles ($\pi^{\pm}$) would cause the Cherenkov light emission while others ($K^{\pm}$, $p^{\pm}$) would not. The ACC comprised of 960 separate modules in the barrel region and 228 in the forward region, each consisting of silica aerogel blocks with refractive indices ranging from 1.01 to 1.03 depending on the polar angle. The polar angle coverage was $33.3° < \theta < 13.6°$ for the barrel part and $13.6° < \theta < 33.4°$ in the forward endcap.

TOF The TOF system provided particle identification for kaons and pions with momenta below 1.2 GeV/c. The mass of a particle was determined by measuring the time T it needed to reach TOF according to the following relation:

$$m = p\sqrt{\left(\frac{T}{L}\right)^2 - 1}, \tag{2.11}$$

where L is the distance between the interaction point and TOF and p is the particle momentum measured in the CDC. The Time of Flight system consisted of 128 scintillator counters located at a radius $r = 1.2\ m$ from the interaction point and covered the polar angle range from 34° to 120°.

ECL The ECL was used for photon detection and electron identification. It was made up of 8736 Thallium doped Cesium Iodide CsI(Tl) crystal blocks that covered the entire tracking volume of the CDC. Electrons and photons would deposit their energy in the crystals producing electromagnetic showers and causing the CsI to scintillate. The amount of scintillation light was then read out by the system of photodiodes and provided the information about deposited energy. The energy resolution of the ECL was $\sigma_E/E = 1.3\%/\sqrt{E}$, where E is the photon evergy in GeV.

KLM The KLM identified K_L mesons and muons with momenta above 600 MeV.

3. Analysis overview

As it was mentioned in chapter 1, $B \to K\tau\mu$ processes studied in this thesis are experimentally challenging as the existence of at least one neutrino in the final state renders it impossible to fully reconstruct signal candidates. In this thesis it will be shown however, that in the experimental setup of B factories it is possible to use kinematic constraints characteristic to $B \to K\tau\mu$ processes in order to efficiently separate them from background events. A clear signal signature of B_{sig} decays allows to loosen the selection criteria for tag-side and apply inclusive tagging algorithm. The presented method is complex and contains several steps that are briefly summarised in this chapter.

3.1. Kinematical description of $B \to K\tau\mu$ decays

Lepton flavour violating $B \to K\tau\mu$ decays[1] are kinematically unique in that sense, that there is only one not (fully) reconstructible particle in the initial state. This is in contrast to any Standard Model B decays involving τ lepton, such as $B \to \bar{D}^{(*)}\tau^+\nu_\tau$, where τ is always associated with a neutrino (or another τ lepton) to preserve lepton flavour. In that sense, $B \to K\tau\mu$ decays are more similar to light flavour semileptonic $B \to \bar{D}^{(*)}\ell^+\nu_\ell$ ($\ell = e, \mu$) decays mentioned in chapter 2 and offer additional kinematic constraints useful to separate them from abundant background processes. In particular, one has an access to a powerful discriminating variable, $\cos\angle(\mathbf{p}_B, \mathbf{p}_{vis})$ (Equation 2.5), which takes the form of:

$$\cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{K\mu}) = \frac{2E_{beam}E_{K\mu} - m_B^2 - m_{K\mu}^2 + m_\tau^2}{2|\mathbf{p}_B||\mathbf{p}_{K\mu}|}, \tag{3.1}$$

where $\mathbf{p}_{K\mu} = \mathbf{p}_{vis}$, $E_{K\mu} = E_{vis}$ are respectively the visible momentum and energy of the decay under assumption that τ is not reconstructed, i.e., assuming $m_{mis} = m_\tau$ (cf. Equation 2.5). The superscript "sig" will be used to explicitly state that here we refer to the momentum of the signal B meson.

[1] The same considerations apply to other decays of type $B \to h\tau\ell(\ell = e, \mu)$, where h denotes the primary hadron.

In case τ decays hadronically, i.e., for $B \to K\tau(\to h\nu)\mu$ processes, where h denotes a secondary hadron (or a system of hadrons), the only undetected particle is the neutrino from τ decay. In that case, assuming that $\mathbf{p}_{vis} = \mathbf{p}_{Kh\mu}$, $E_{vis} = E_{Kh\mu}$, $m_{mis} = 0$ it is possible to calculate another angle:

$$\cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) = \frac{2E_{beam}E_{Kh\mu} - m_B^2 - m_{Kh\mu}^2}{2|\mathbf{p}_B||\mathbf{p}_{Kh\mu}|} \tag{3.2}$$

Effectively, equations 3.1, 3.2 indicate that in $\mathrm{B} \to \mathrm{K}\tau(\to \mathrm{h}\nu)\mu$ decay, $\mathbf{p}_B^{sig}$ is constrained on a cone around $\mathbf{p}_{\mathrm{K}\mu}$ and, at the same time, it lies on a cone around $\mathbf{p}_{\mathrm{Kh}\mu}$. Since both conditions (3.1, 3.2) have to be fulfilled simultaneously, it is easy to notice that there are only two possible configurations of the B meson momentum, given by the intersection of two cones (Figure 3.1).

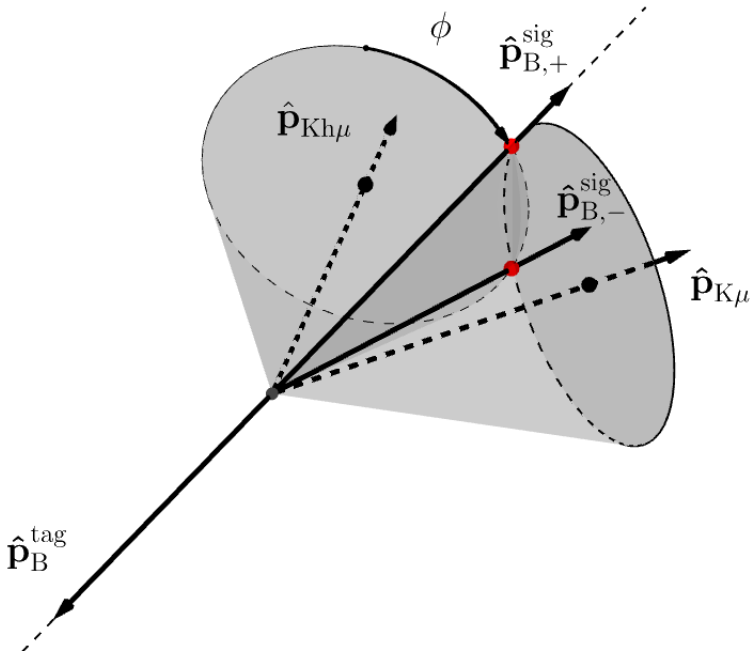


Figure 3.1. A schematic view of $B \to K\tau(\to h\nu)\mu$ decay kinematics. The signal B meson momentum, $\mathbf{p}_B^{sig}$ is constrained on a cone around $\mathbf{p}_{Kh\mu}$ and $\mathbf{p}_{K\mu}$ momenta. The intersection of two cones defines two kinematically allowed configurations of $\mathbf{p}_B^{sig}$, denoted as $\mathbf{p}_{B,\pm}^{sig}$.

Finding for the intersection points allows to analytically recover B meson momentum with a two-fold ambiguity. A detailed derivation can be found in Appendix A. Two solutions for $\mathbf{p}_B^{sig}$ are given by:

$$\mathbf{p}_{B,\pm}^{sig} = |\mathbf{p}_B|\,(\pm u\ \hat{\mathbf{a}}_1 + v\ \hat{\mathbf{a}}_2 + w\ \hat{\mathbf{a}}_3) \tag{3.3}$$

where $\hat{\mathbf{a}}_1 = \frac{\hat{\mathbf{p}}_{K\mu}\times\hat{\mathbf{p}}_{Kh\mu}}{|\hat{\mathbf{p}}_{K\mu}\times\hat{\mathbf{p}}_{Kh\mu}|}$, $\hat{\mathbf{a}}_2 = \frac{\hat{\mathbf{p}}_{K\mu} - \cos\angle(\mathbf{p}_{K\mu},\mathbf{p}_{Kh\mu})\hat{\mathbf{p}}_{Kh\mu}}{|\hat{\mathbf{p}}_{K\mu}\times\hat{\mathbf{p}}_{Kh\mu}|}$, $\hat{\mathbf{a}}_3 = \hat{\mathbf{p}}_{Kh\mu}$. Parameters u, v, w can be calculated as:

$$u = \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu})\sqrt{1-\sin^2\phi} \tag{3.4}$$

$$v = \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu})\sin\phi \tag{3.5}$$

$$w = \cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \tag{3.6}$$

where ϕ is the angle that specifies the position of $\mathbf{p}_B^{sig}$ on a cone around $\mathbf{p}_{Kh\mu}$ momentum[2]. The value of $\sin\phi$ can be expressed as:

$$\sin\phi = \frac{cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{K\mu}) - \cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu})\cos(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu})}{\sin\angle(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu})\sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu})}. \tag{3.7}$$

The great advantage of this procedure is that it allows to recover B_{sig} momentum in a way that is independent of B_{tag} reconstruction. Moreover, the variable $\sin\phi$ is also a good discriminator between signal and background. As it will be shown in subsection 4.3.3, the values of $\sin\phi$ for correctly reconstructed $B \to K\tau(\to h\nu)\mu$ decays lie in physical region $-1 < \sin\phi < 1$, with small deterioration due to finite detector resolution, while the distribution of $\sin\phi$ for background processes is much broader and extends to unphysical region $|\sin\phi| > 1$. Therefore $B \to K\tau(\to h\nu)\mu$ decays can be efficiently identified based on $\sin\phi$ distribution. Some sort of tagging is still necessary in that case in order to increase the signal-to-background ratio, but a clear signal signature allows to reconstruct B_{tag} inclusively (cf. section 2.2), which in turn increases the efficiency. The momenta of the reconstructed B_{tag} ($\mathbf{p}_B^{tag}$) and B_{sig} ($\mathbf{p}_{B,\pm}^{sig}$) provide additional useful information. In particular the variable:

$$\cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_B^{tag})^{min} \equiv min\left[\cos\angle(\mathbf{p}_{B,+}^{sig}, \mathbf{p}_B^{tag}),\ \cos\angle(\mathbf{p}_{B,-}^{sig}, \mathbf{p}_B^{tag})\right] \tag{3.8}$$

which is the smaller value of the cosine of the angle between $\mathbf{p}_B^{tag}$ and one of the solutions for $\mathbf{p}_B^{sig}$, can be used to further discriminate background events. For events where both B_{sig} and B_{tag} were correctly reconstructed, the distribution of $\cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_B^{tag})^{min}$ will peak around -1, since two B mesons are produced back-to-back in the $\Upsilon(4S) \to B\bar{B}$ decay, while for background events the distribution will be broader.

3.2. Reconstruction of B_{sig} and B_{tag}

The starting point of the analysis that relies on inclusive B_{tag} reconstruction is the choice of B_{sig} decays. In this work, the following channels were considered[3]:

$$B^+ \to K^+\tau^+\mu^-, \tag{3.9}$$

$$B^+ \to K^+\tau^-\mu^+. \tag{3.10}$$

[2]The angle ϕ was shown schematically in Figure 3.1, while its actual definition, and meaning is presented in Appendix A.

[3]Charge conjugate modes are implicitly included throughout the text.

τ lepton was reconstructed exclusively in one prong hadronic $\tau^- \to \pi^-\nu_\tau$ mode. This particular choice was a compromise between its relatively large branching fraction, $\mathcal{B}(\tau^- \to \pi^-\nu_\tau) = 10.82 \pm 0.05$ [18], and good reconstruction efficiency in comparison to other hadronic modes, such as $\tau^- \to \pi^-\pi^0\nu_\tau$. Leptonic τ modes, although are expected to have lower combinatorial background[4], were not considered because the existence of two neutrinos in the final state does not allow to apply all kinematical constraints described in section 3.1. After selecting the B_{sig} candidate, the remaining tracks and clusters are combined together to form B_{tag}. To further improve the quality of B_{tag} reconstruction and increase the background suppression, additional requirements on the total event charge as well as the number of charged and neutral particles assigned to B_{tag} are imposed. Signal-side decays are identified based on $\sin\phi$ distribution, where selecting events in region $\sin\phi \in [-1, 1]$ checks if the reconstructed B_{sig} is kinematically compatible with $B \to K\tau(\to \pi\nu_\tau)\mu$ decay.

3.3. Signal yield extraction

In the presented method of the inclusive B_{tag} reconstruction, B_{sig} candidates are found first and then it is checked if the remaining tracks and clusters are consistent with the hypothesis that they originate from the accompanying B meson (B_{tag}). In such approach, the B_{sig} decay can be observed in characteristics of the recoiling B_{tag}. In particular, the variable M_{tag} will be sensitive to the proper assignment of B_{sig} daughers to its decay. If the B_{sig} candidate was properly selected and all final state particles were correctly split between B_{sig} and B_{tag}, the distribution of M_{tag} will peak around true B meson mass. The final signal yield can thus be extracted from the fit to M_{tag} distribution, provided that the background was reduced to an acceptable level.

[4]This is related to the environment of B factories, where charged hadrons are produced more copiously than leptons.

4. Monte Carlo studies

This chapter describes the main analysis chain. The analysis procedure was established using Monte Carlo generated events. Application to experimental Belle dataset was not performed yet, in order to avoid the experimenter's bias.

4.1. Monte Carlo samples

Generic Monte Carlo To model different sources of background processes, the so-called Generic Monte Carlo samples were used. These samples were produced centrally and are made available for all collaboration members. Generic Monte Carlo contains both $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ and $e^+e^- \to q\bar{q}(q = u, d, s, c)$ processes that are divided into four categories[1]:

- $e^+e^- \to \Upsilon(4S) \to B^+B^-$,
- $e^+e^- \to \Upsilon(4S) \to B^0\bar{B}^0$,
- $e^+e^- \to q\bar{q},\ q = u, d, s,$
- $e^+e^- \to c\bar{c}$.

Experimentally known B meson decays were generated using EvtGen [27], whereas unknown B meson decays and continuum processes were reproduced with PYTHIA [28]. In addition, PHOTOS [29] package was employed to account for the final state radiation of charged particles. The detector response was performed by GEANT3 [30].

Generic Monte Carlo events are divided into subsets called streams. Each stream corresponds to the integrated luminosity of the full Belle data set ($\mathcal{L}_{int} = 711 fb^{-1}$) and is further divided into several data-taking periods (experiments). This way, the experiment-by-experiment differences, present in real data, are taken into account and are correctly reproduced in the simulation. There are ten streams of $B\bar{B}$ ($B^0\bar{B}^0$ and B^+B^-) and six streams of continuum ($e^+e^- \to q\bar{q}$ ($q = u, d, s, c$)) events available for common use. In this thesis, three streams of Generic Monte Carlo for each category will be used to evaluate background contributions, as it will be described later.

[1] Continuum processes involving c quarks often lead to charmed meson production and therefore are treated separately.

Signal Monte Carlo To estimate the selection efficiency for signal decays, dedicated Signal MC samples where one B meson (B_{sig}) decays according to a signal decay model and the other (B_{tag}) decays generically were used. Due to *a priori* unknown decay dynamics, signal decays were generated according to a uniform three-body phase space model. Since no such samples were available, they were generated by the author himself. The events were produced using the same software as the one used for Generic MC production, i.e., EvtGen (+PHOTOS) for decay modelling, followed by GEANT3 for detector response. In total, 5×10^6 events were generated for each $B^+ \to K^+\tau^\pm\mu^\mp$ decay channel.

4.2. Particle selection and identification

Charged tracks Charged tracks are required to originate from the vicinity of the interaction point. Each track has to satisfy the criteria based on distance of the closest approach (DOCA) with respect to the interaction point(IP): in the transverse direction, $dr < 1cm$ and along the beam axis $|dz| < 4cm$. In addition, each track is required to have transverse momentum in the laboratory frame, $p_T > 50MeV$, which reduces the background from low-momentum particles.

At Belle, charged hadrons ($x = K, p, \pi$) are identified based on measurements from three subdetectors: CDC, TOF and ACC. The information from all subdetectors is then combined to determine the likelihood for a track to originate from a given particle x:

$$\mathcal{L}_x = \mathcal{L}_x^{CDC} \times \mathcal{L}_x^{ACC} \times \mathcal{L}_x^{TOF}, \tag{4.1}$$

where $\mathcal{L}_x^{CDC}$, $\mathcal{L}_x^{ACC}$, $\mathcal{L}_x^{TOF}$ are the likelihoods based on information from CDC, ACC and TOF, respectively. For each track, the likelihood ratio is determined as [31]:

$$\mathcal{R}_{x|y} = \frac{\mathcal{L}_x}{\mathcal{L}_x + \mathcal{L}_y}. \tag{4.2}$$

The ratio $R_{x|y}$ is then used to separate signal particle x from background particle y. For instance, $R_{K|\pi} > 0.6$ would mean a clean kaon signature, with a small pion misidentification rate.

Electrons are identified based on measurements from ECL, CDC and ACC. The electron identification likelihood is defined as [32]:

$$\mathcal{R}_e = \frac{\mathcal{L}_e}{\mathcal{L}_e + \mathcal{L}_{\neg e}}, \tag{4.3}$$

where $\mathcal{L}_e$ ($\mathcal{L}_{\neg e}$) is the electron (non-electron) likelihood. For muon identification, the normalised muon likelihood is used [33]:

$$\mathcal{R}_\mu = \frac{\mathcal{L}_\mu}{\mathcal{L}_\mu + \mathcal{L}_\pi + \mathcal{L}_K}, \tag{4.4}$$

which is determined based on measurements from KLM.

Accepted tracks are assigned mass hypotheses based on standard Belle identification criteria (Table 4.1).

Table 4.1. Charged track selection and identification criteria.

Candidate	Selection
Charged tracks	dr $<$ 1cm $\wedge$ |dz| $<$ 4cm
$e^\mp$	$\mathcal{R}_e > 0.9 \wedge \mathcal{R}_\mu < 0.98 \wedge \mathcal{R}_{K\|\pi} < 0.98$
$\mu^\mp$	$\mathcal{R}_\mu > 0.9 \wedge \mathcal{R}_e < 0.98 \wedge \mathcal{R}_{K\|\pi} < 0.98$
$K^\pm$	$\mathcal{R}_{K\|\pi} > 0.6 \wedge \mathcal{R}_\mu < 0.98 \wedge \mathcal{R}_e < 0.98$
$p^\pm$	$\mathcal{R}_{p\|\pi} > 0.6 \wedge \mathcal{R}_{p\|K} > 0.6 \wedge \mathcal{R}_\mu < 0.98 \wedge \mathcal{R}_e < 0.98$

π^0 mesons π^0 candidates are reconstructed from photon pairs having invariant mass in range $118MeV < m_{\gamma\gamma} < 150MeV$. To reduce the combinatorial background, photons originating from π^0 are required to have energy above $50MeV$ in barrel part of the ECL and above $100MeV$ in endcaps. For each pair, a mass-constrained vertex fit is perfomed which improves the reconstructed momentum resolution. Among π^0 candidates that shared a common γ daughter, the one with with the lowest χ^2 value of the vertex fit is selected.

Photons Photons that are not associated with π^0 are accepted based on energy requirements that range from $100MeV$ to $200MeV$, depending on their polar angle in the ECL (Table 4.2).

Table 4.2. Photon selection criteria.

ECL Region	Selection
$-0.75 < \cos\theta \leq 0.5$	$E_\gamma > 100$MeV
$0.5 < \cos\theta \leq 0.6$	$E_\gamma > 160$MeV
$0.6 < \cos\theta \leq 0.7$	$E_\gamma > 180$MeV
$0.7 < \cos\theta \leq 0.85$	$E_\gamma > 200$MeV

The above selection is optimised to reject soft photons originating from beam background and secondary interactions in the electromagnetic calorimeter. The sum of energies of clusters that do not fulfill the photon selection requirements, called the residual energy E_{res}, is calculated and used as a discriminating variable later on.

4.3. Event reconstruction

4.3.1. Preselection

As it was mentioned in section 3.2, the event reconstruction starts with selecting B_{sig} candidates. Each event is required to contain three tracks: a K^+ (K^- for charge conjugate modes) candidate, a muon candidate and a pion candidate with charge opposite to that muon. From now on a $K^+\mu^\pm\pi^\mp$ triple will be reffered to as the B_{sig} candidate. At this point it is allowed to have multiple B_{sig} candidates per event, while after the final selection only one candidate will be kept. For each B_{sig} candidate the event is searched for the remaining tracks and clusters that are not assigned to this B_{sig} and those are assumed to originate from the B_{tag} decay. The sum of momenta and energies of those is used to calculate variables M_{tag} and ΔE_{tag} (cf. section 2.2). At this stage, all candidates are required to pass the following preselection requirements:

- $\mathrm{M_{tag}} > 5.2$ GeV,
- -0.5 GeV $< \Delta\mathrm{E_{tag}} < 0.3$ GeV,
- Normalised second Fox-Wolfram moment [34]: $\mathrm{R_2} < 0.4$.

The purpose of this selection is to reduce data sample to a manageable level by rejecting events that are not useful for further analysis. The first two requirements are loose contraints on $\mathrm{B_{tag}}$ reconstruction quality. The third requirement rejects around 30% of continuum background, while retaining 98% of $\mathrm{B\bar{B}}$ events. The signal efficiency[2] and the number of reconstructed background events after the preselection are reported in Table 4.3. At this stage, the expected background levels are still large. Further background suppression will be obtained by exploiting properties characteristic to B_{tag} and B_{sig} decays, which will be described in next subsections.

Table 4.3. The signal efficiency and the number of background events that pass the preselection criteria. The number of background events was normalised to the integrated luminosity in full Belle dataset.

Decay mode	$N_{background}$				Signal efficiency [%]
	B^+B^-	$B^0\bar{B}^0$	$e^+e^- \to q\bar{q}$ (q $\in u, d, s$)	$e^+e^- \to c\bar{c}$	
$B^+ \to K^+\tau^+\mu^-$	3.89×10^6	3.53×10^6	1.72×10^6	2.64×10^6	4.86%
$B^+ \to K^+\tau^-\mu^+$	4.76×10^6	4.39×10^6	1.53×10^6	2.79×10^6	4.86%

[2]Here, the signal efficiency is defined as the ratio between the number of reconstructed signal events that pass the selection and the number of generated events in Signal Monte Carlo sample. Only the events with correctly reconstructed (true-matched) B_{sig} candidate are considered.

4.3.2. B_{tag} characteristics

The distributions of M_{tag} and ΔE_{tag} after the preselection, for B_{sig} candidates reconstructed in Signal and Generic Monte Carlo samples are shown in Figure 4.1. For background evaluation, three streams of Generic Monte Carlo samples were used, which were then scaled down by a factor of 1/3 to normalise them to the integrated luminosity recorded in the full Belle dataset. All background processes are divided into four categories (see section 4.1) representing different background contributions and drawn additively in correct proportion. The histogram for signal candidates (correctly selected B_{sig} candidates reconstructed in Signal Monte Carlo sample), normalised to the same area as the total background histogram was shown for comparison.

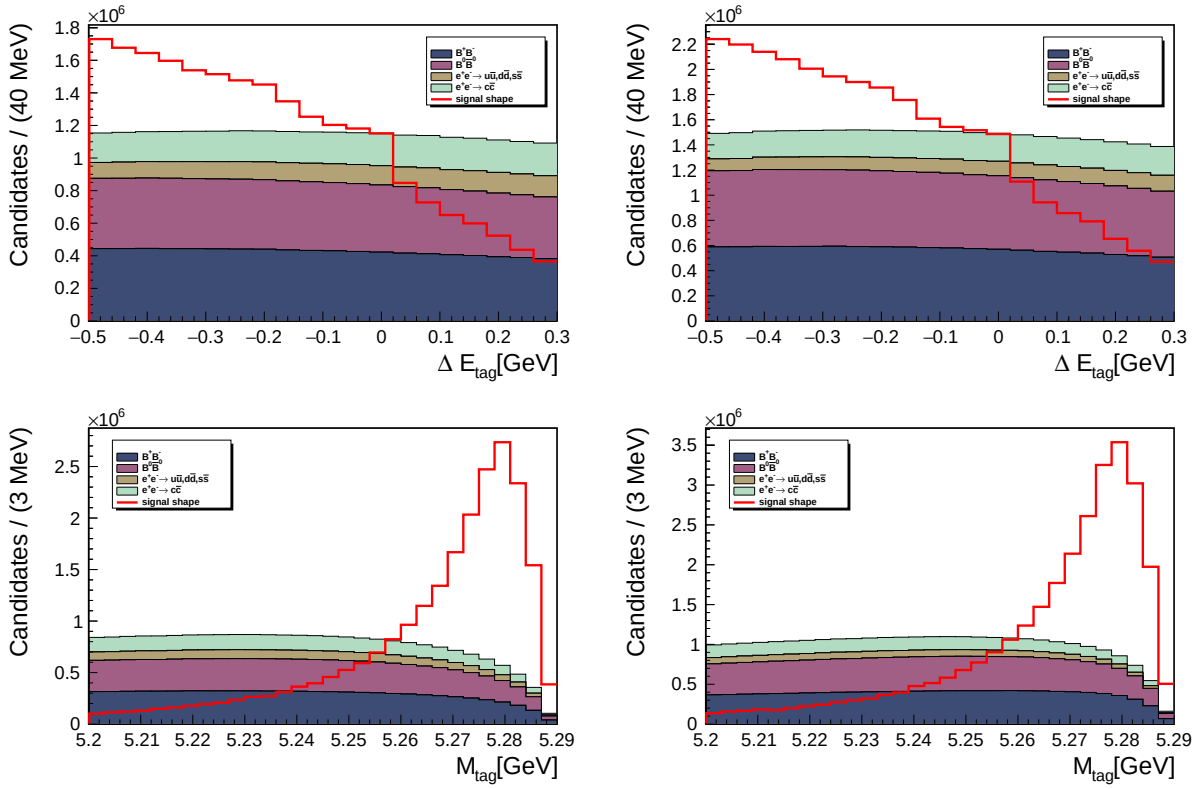


Figure 4.1. Distributions of ΔE_{tag} and M_{tag} after the preselection for $B^+ \to K^+\tau^+\mu^-$ (left) and $B^+ \to K^+\tau^-\mu^+$ (right) channel.

As it can be seen, the characteristics of the inclusively reconstructed B_{tag} are different between signal and background. For signal candidates, where B_{sig} is properly selected, all the remaining tracks and clusters should ideally originate from B_{tag} produced the $\Upsilon(4S) \to B_{sig}B_{tag}$ decay. In that case, the distributions of ΔE_{tag} and M_{tag} should show narrow structures around $\Delta E_{tag} \approx 0$ and $M_{tag} \approx 5.279\ GeV$, indicating a correctly reconstructed B_{tag} decay. In fact, for events with fully reconstructed B_{tag}, distributions of M_{tag} and ΔE_{tag} are expected to be similar to those obtained by exclusive reconstruction, however with deteriorated resolution due to averaging over many different B_{tag} decay modes.

In practice, even if B_{sig} is properly selected, distributions of $\mathrm{M_{tag}}$ and $\Delta\mathrm{E_{tag}}$ will contain events where not all final state particles were detected and assigned to B_{tag}. The distribution of $\Delta\mathrm{E_{tag}}$ is asymmetric and shows an excess of events in region $\Delta\mathrm{E_{tag}} < 0$, which is caused by the finite detection efficiency and semileptonic transitions in $\mathrm{B_{tag}}$ decay chain. The source of events with positive values of $\Delta\mathrm{E_{tag}}$, much above the expected resolution, are spurious showers in the electromagnetic calorimeter caused by the beam background or secondary interactions of hadrons. In contrast to $\Delta\mathrm{E_{tag}}$ distribution, that is very broad and asymetric, the distribution of $\mathrm{M_{tag}}$ shows a more pronounced peak around true B meson mass. This is because $\mathrm{M_{tag}}$ is not sensitive to loosing very slow particles or adding additional clusters to $\mathrm{B_{tag}}$. Such contributions add linearly to $\Delta\mathrm{E_{tag}}$, but tend to average in the vector sum $\mathbf{p}_{\mathrm{tag}}^{\mathrm{rec}} = \sum \mathbf{p}_{\mathrm{i}}$, used to evaluate $\mathrm{M_{tag}}$ (Equation 2.7, Equation 2.8).

For background processes, the particles assigned to B_{sig} candidate ($K^+\mu^\pm\pi^\mp$ triple) do not necessarily originate from a single B meson decay[3], but are rather a random combination of tracks originating from different sources; a mixture of B and $\bar{B}$ decay products for $B\bar{B}$ background or quark fragmentation products in case of continuum processes. In that case, also the tracks and clusters assigned to B_{tag} do not originate from its decay (at least not all of them) and therefore the distributions of ΔE_{tag} and M_{tag} will show a combinatorial shape; with no obvious peaking structure that would indicate a correctly reconstructed B meson decays.

The quality of the $\mathrm{B_{tag}}$ reconstruction can be improved by rejecting events in which not all particles were detected and those that receive a sizeable contribution from spurious showers. In further selection, events were accepted if they satisfied the following requirements:

A.1 The total event charge: $\mathrm{Q_{tot}} = 0$,

A.2 The numer of leptons on tag side: $\mathrm{N_{leptons}^{tag}} = 0$,

A.3 The number of neutral particles on tag side: $\mathrm{N_{\gamma}^{tag}} + \mathrm{N_{\pi^0}^{tag}} < 5$,

A.4 Residual energy in the ECL (section 4.2): $\mathrm{E_{res}} < 0.6\mathrm{GeV}$.

The first requirement rejects events in which some of the charged particles were not detected. The second requirement rejects $\mathrm{B_{tag}}$ decays involving semileptonic transitions, that suffer from incomplete kinematic information. The last two selections remove events with a large number of spurious showers in the ECL. A set of criteria $\mathrm{A.1} - \mathrm{A.4}$ will be reffered to as the "B_{tag} selection".

[3]This is not entirely true as some B meson decays can produce final states that are identical to signal. For example, a semileptonic decay $B^+ \to \bar{D}^0\mu^+\bar{\nu}_\mu$ followed by $\bar{D}^0 \to K^+\pi^-$ has final state identical to $B^+ \to K^+\tau^-(\to \pi^-\nu_\tau)\mu^+$ signal decay. If the visible $B^+ \to \bar{D}^0(\to K^+\pi^-)\mu^+\bar{\nu}_\mu$ decay producs ($K^+\pi^-\mu^+$ triple) are treated as the B_{sig} candidate, and the remaining particles, originating from accompanying (hadronic) B meson decay are all reconstructed, then the distributions of ΔE_{tag} and M_{tag} will show peaking structures, similar to those for signal decays. This type of background will be discussed separately in later subsections.

The result of applying requirements $A.1 - A.4$ on B_{tag} reconstruction quality is illustrated in Figure 4.2, that shows two-dimensional distributions of $M_{tag} - \Delta E_{tag}$ for candidates reconstructed in Signal Monte Carlo sample. As it can be seen, rejecting events in which some of the particles were not detected, or were falsely added to B_{tag} improves the resolution in ΔE_{tag}, M_{tag} variables. As a result, most of the reconstructed signal candidates are localised within small window: $-0.25\text{GeV} < \Delta\text{E}_{\text{tag}} < 0.05\text{GeV} \wedge \text{M}_{\text{tag}} > 5.27\text{GeV}$. The region $-0.25\text{GeV} < \Delta\text{E}_{\text{tag}} < 0.05\text{GeV} \wedge \text{M}_{\text{tag}} > 5.27\text{GeV}$ will be called the signal region henceforth.

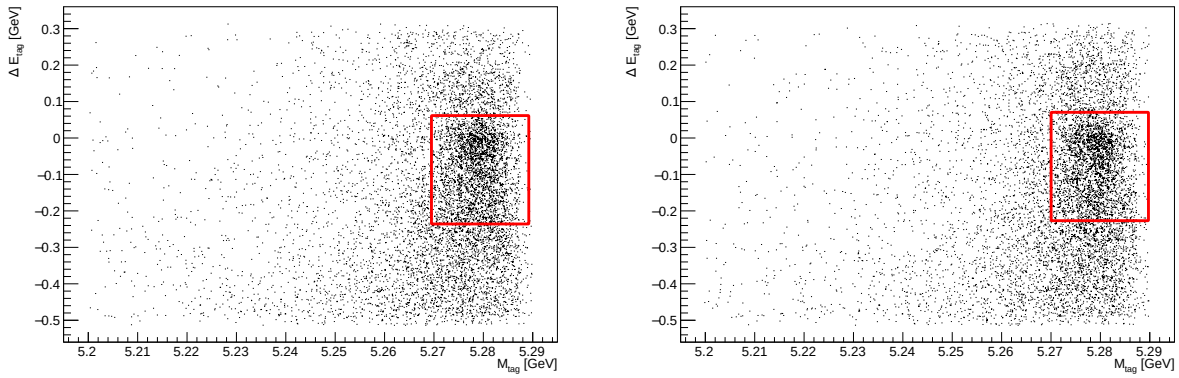


Figure 4.2. Distributions of $M_{tag} - \Delta E_{tag}$ for candidates reconstructed in Signal Monte Carlo sample after selection $A.1 - A.4$, for $B^+ \to K^+\tau^+\mu^-$ (left) and $B^+ \to K^+\tau^-\mu^+$ (right) channel. The signal region $-0.25GeV < \Delta E_{tag} < 0.05GeV \wedge M_{tag} > 5.27GeV$ was indicated by a red rectangle.

Figure 4.3 shows one-dimensional projections of ΔE_{tag} in window $M_{tag} > 5.27GeV$ and M_{tag} in window $-0.25GeV < \Delta E_{tag} < 0.05GeV$ (in other words: one variable was plotted in signal window of the other variable). For signal candidates, the distribution of ΔE_{tag} shows a clear peak around $\Delta E_{tag} \approx 0$, which comes from fully reconstructed B_{tag} decays. A less pronounced structure around $\Delta E_{tag} \approx -0.15GeV$ originates from candidates with undetected slow particles, that do not significantly deteriorate the resolution of M_{tag}. For signal candidates lying in region $-0.25GeV < \Delta E_{tag} < 0.05GeV$, the distribution of M_{tag} shows a clear peak around true B meson mass, with about 80% of the candidates constrained in window $M_{tag} > 5.27GeV$.

The procedure described above, when applied to Signal Monte Carlo, allows to select a relatively clean sample of events where one B meson (B_{tag}) decays hadronically and is properly reconstructed. Applying requirements $A.1 - A.4$ also changes the composition of background. As it can be seen in Figure 4.3, the distributions of ΔE_{tag} and M_{tag} for background processes, although still dominated by combinatorial effects, also contain events where B_{tag} is properly reconstructed. An excess of candidates around $\Delta E_{tag} \approx 0 \wedge M_{tag} \approx 5.279GeV$ originates from $B\bar{B}$ events in which one B meson decayed to the same final state as signal (for example, the

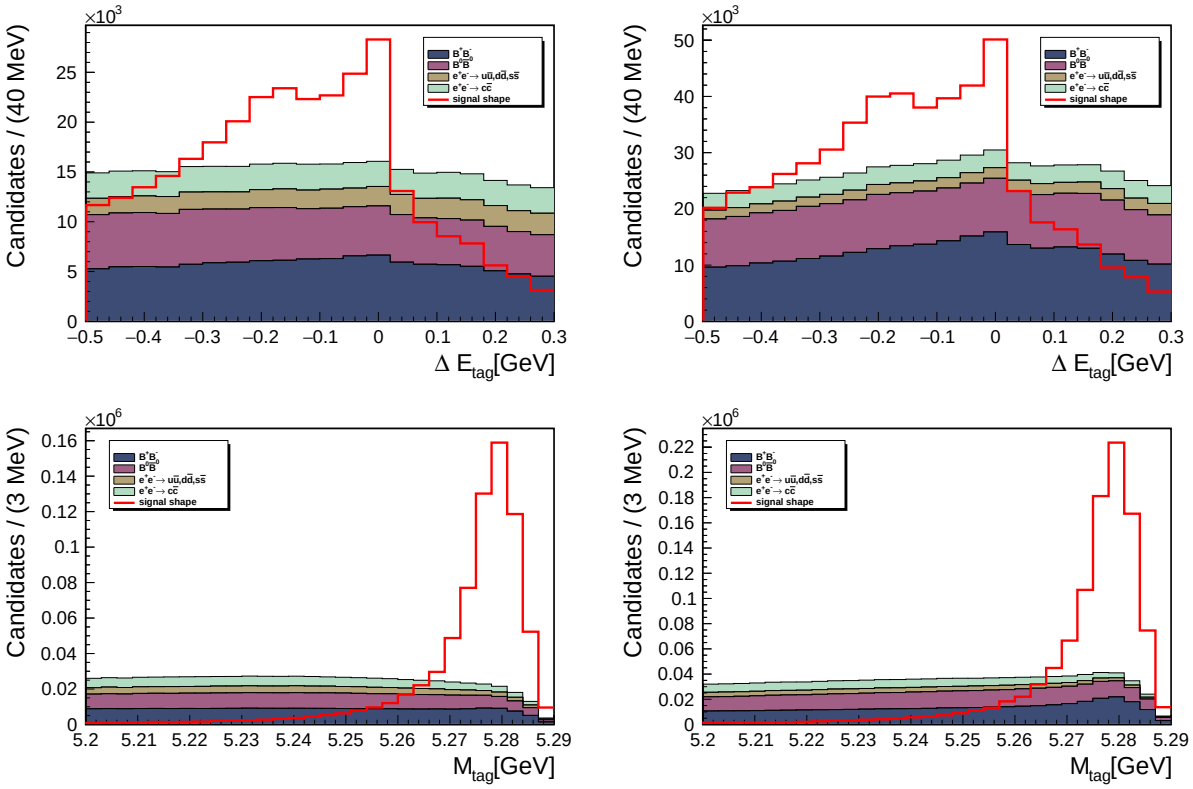


Figure 4.3. Distributions of ΔE_{tag} in window $M_{tag} > 5.27GeV$ and M_{tag} in window $-0.25GeV < \Delta E_{tag} < 0.05GeV$ for $B^+ \to K^+\tau^+\mu^-$ (left) and $B^+ \to K^+\tau^-\mu^+$ (right) channel. All candidates pass the requirements A.1 – A.4.

$B^+ \to \bar{D}^0(\to K^+\pi^-)\mu^+$ decay mentioned before), while the other decayed hadronically and was fully reconstructed. This type of background cannot be removed based on B_{tag} properties and will be suppressed by exploiting observables characteristic to B_{sig} decay.

4.3.3. B_{sig} characteristics

Kinematic constraints The distributions of $\sin\phi$ for candidates that pass selection A.1 − A.4 and lie in signal region $-0.25\text{GeV} < \Delta E_{tag} < 0.05\text{GeV} \wedge M_{tag} > 5.27$ GeV are shown in Figure 4.4. For correctly reconstructed signal decays, values of $\sin\phi$ lie in physical region $-1 < \sin\phi < 1$, however with small deterioration due to a finite detector resolution. The distribution for background processes is much broader and extends to unphysical region $|\sin\phi| > 1$.

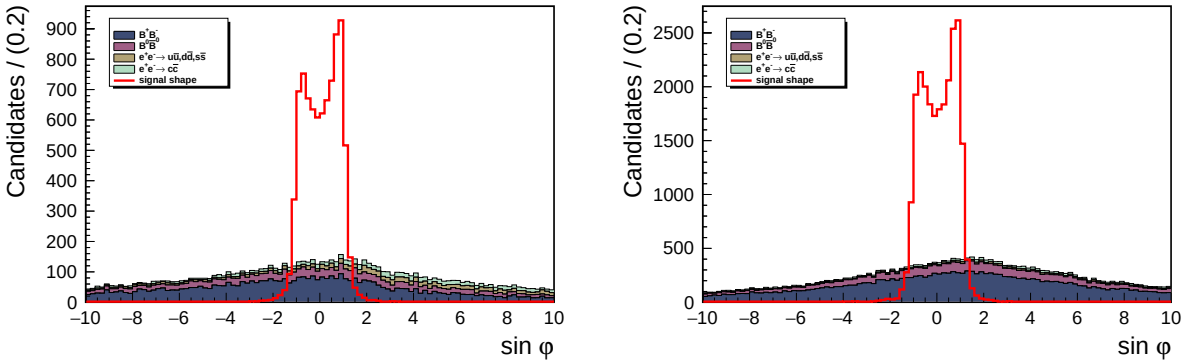


Figure 4.4. Distributions of $\sin\phi$ for $B^+ \to K^+\tau^+\mu^-$ (left) and $B^+ \to K^+\tau^-\mu^+$ (right) modes. All candidates satisfy requirements $A.1 - A.4$ and lie in signal window: $-0.25GeV < \Delta E_{tag} < 0.05GeV \wedge M_{tag} > 5.27\ GeV$.

Based on these considerations, the further analysis was constrained to region:

B.1 $-1 < \sin\phi < 1$

The selection described above is around 80% efficient for signal decays and removes about 98% of background events Table 4.4.

Table 4.4. Relative selection efficiency for $|\sin\phi| < 1$ requirement, calculated with respect to the number of reconstructed events that pass selection A.1 − A.4 and lie in signal window: $-0.25\text{GeV} < \Delta E_{tag} < 0.05\text{GeV} \wedge M_{tag} > 5.27$ GeV.

	Relative selection efficiency				
Decay mode	B^+B^-	$B^0\bar{B}^0$	$e^+e^- \to q\bar{q}$ (q $\in u, d, s$)	$e^+e^- \to c\bar{c}$	Signal
$B^+ \to K^+\tau^+\mu^-$	1.75%	0.78%	1.27%	0.95%	80.24%
$B^+ \to K^+\tau^+\mu^-$	2.42%	1.17%	1.25%	0.85%	80.31%

$B\bar{B}$ background After imposing requirements on $\sin\phi$, the dominant background is from $B\bar{B}$ events. The dominant background depends on the relative charge of the kaon and muon:

- For $B^+ \to K^+\tau^-\mu^+$ candidates, the dominant background is from semileptonic B decays of type $B \to \bar{D}(\to K^+\pi^- X)\mu^+\nu$, where X refers to additional particles that are either not reconstructed or are added to B_{tag}.
- For $B^+ \to K^+\tau^+\mu^-$ candidates, the dominant background comes from B decays followed by semileptonic D decays. For example, a $\bar{D}^0 \to K^+\mu^-\nu_\mu$ decay, together with an additional π^+ can mimick signal.

To remove this source of background, the selection on $m_{K\pi}$; the invariant mass of the kaon candidate and the oppositely-charged B_{sig} daughter for which pion mass hypothesis is assumed, was imposed. The distributions of $m_{K\pi}$ are shown in Figure 4.5. The distributions of $m_{\mu\mu}$ are shown in Figure 4.5.

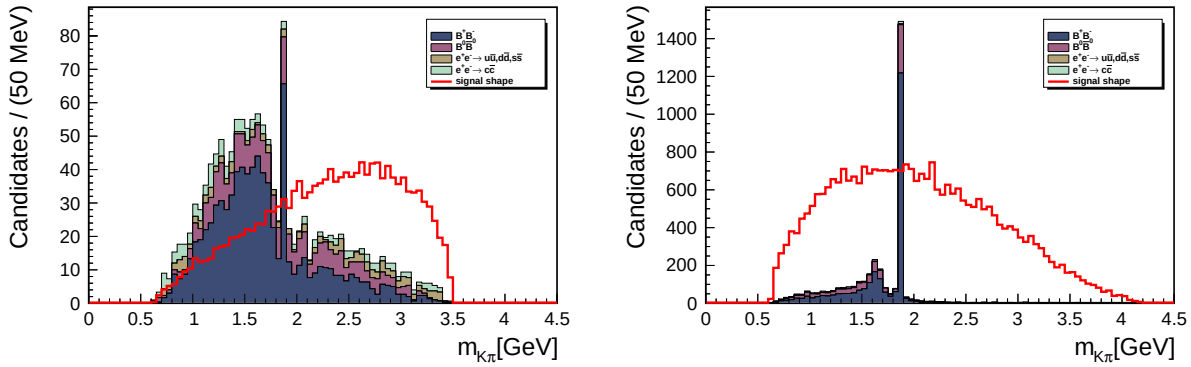


Figure 4.5. Distributions of $m_{K\pi}$ for $B^+ \to K^+\tau^+\mu^-$ (left) and $B^+ \to K^+\tau^-\mu^+$ (right) modes. All candidates satisfy requirements $A.1 - A.4$ and lie in $B_{tag}(-0.25GeV < \Delta E_{tag} < 0.05GeV \wedge M_{tag} > 5.27\ GeV)$ and B_{sig} $(-1 < \sin\phi < 1)$ signal regions. The normalisation of signal histograms is arbitrary.

To remove this type of background the following selection was imposed in the further analysis:

B.2 $m_{K\pi} > 1.95GeV$

This requirement removes background arising from mentioned processes at a relative cost of about 0% signal efficiency.

Another source of an irreducible $B\bar{B}$ background comes from B decays involving charmonium resonances; $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ and $B^+ \to K^+\psi(2S)(\to \mu^+\mu^-)$. If one of the muons is misidentified as a pion, such processes can mimick signal decays.

This source of background was rejected by imposing a selection on the variable $m_{\mu\mu}$, which is the invariant mass of the $\pi\mu$ pair, where pion was assigned with a muon mass hypothesis.

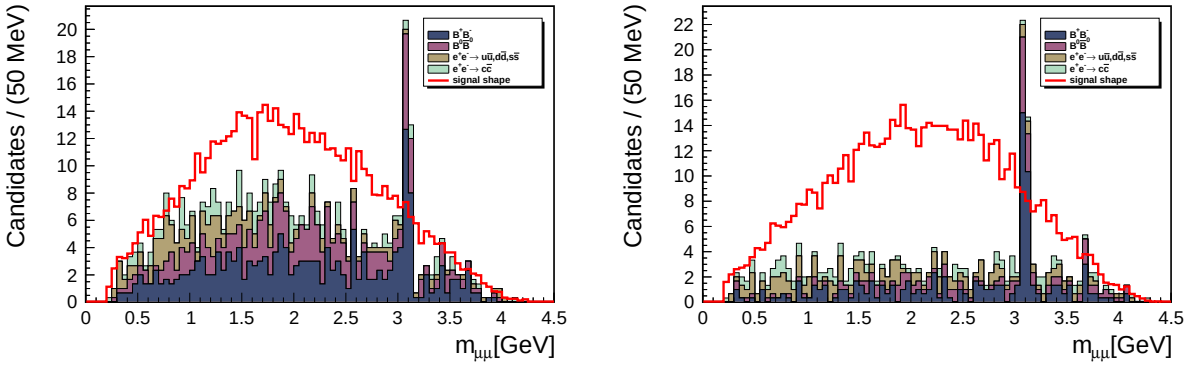


Figure 4.6. Distributions of $m_{\mu\mu}$ for $B^+ \to K^+\tau^+\mu^-$ (left) and $B^+ \to K^+\tau^-\mu^+$ (right) modes. All candidates satisfy requirements $A.1 - A.4$, lie in $B_{tag}(-0.25GeV < \Delta E_{tag} < 0.05GeV \wedge M_{tag} > 5.27\ GeV)$ and B_{sig} $(-1 < \sin\phi < 1)$ signal regions and pass $m_{K\pi} > 1.95\ GeV$ selection. The normalisation of signal histograms is arbitrary.

In further selection, the following requirements were imposed:

B.3 $m_{\mu\mu} \notin [3.03, 3.14]GeV$

B.4 $m_{\mu\mu} \notin [3.60, 3.75]GeV$.

These selections remove background arising from $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$, and $B^+ \to K^+\psi(2S)(\to \mu^+\mu^-)$ processes at a minimal cost of signal efficiency.

4.4. The best candidate selection

As it was mentioned, in some events more than one B_{sig} candidate satisfying the selection criteria can be found. After the final selection, more than one candidate is reconstructed in 0.5% (0.6%) of events from $B^+ \to K^+\tau^+\mu^-$ ($B^+ \to K^+\tau^-\mu^+$) signal MC samples. In such events, only "the best" (i.e., the one with a highest probability to be the true one) candidate is kept, while the others are rejected. In case of multiple candidates per event, the one with the highest confidence level of the vertex fit of the B_{tag} decay was chosen. This procedure selects the correct (true) candidate in about 90% of events with multiple B_{sig} entries reconstructed in Signal Monte Carlo. The results of applying the best candidate selection on M_{tag} spectrum is illustrated in Figure 4.7.

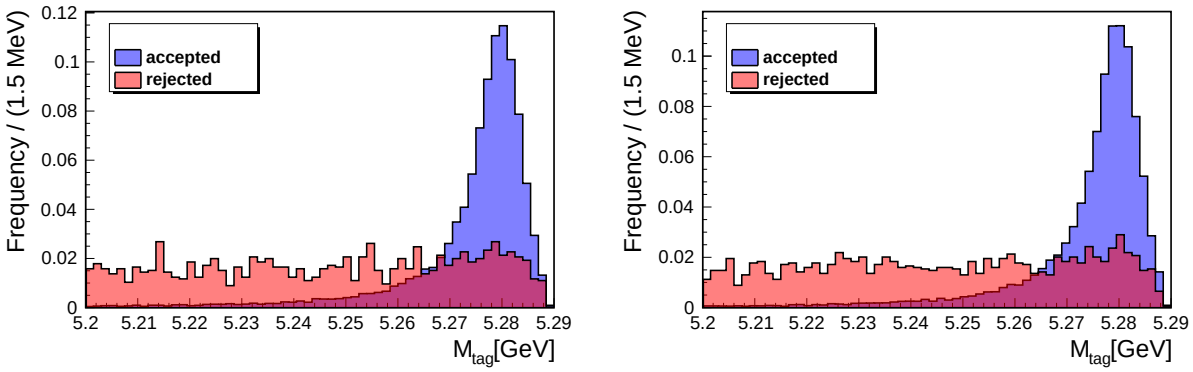


Figure 4.7. Distributions of M_{tag} for events with multiple B_{sig} candidates reconstructed in Signal Monte Carlo samples, for $B^+ \to K^+\tau^+\mu^-$ (left) and $B^+ \to K^+\tau^-\mu^+$ (right) modes. Blue histograms represent accepted (best), red histograms- rejected candidates. The events satisfy requirements A.1 − A.4 ∧ B.1 − B.4 and lie in window $-0.25\text{GeV} < \Delta\text{E}_{\text{tag}} < 0.05\text{GeV}$. Histograms are normalised to unit area.

The selection on $\cos\left(\mathbf{p}_B^{sig}, \mathbf{p}_B^{tag}\right)^{min}$ was optimised by maximizing the Punzi figure of merit (FOM) [35]:

$$FOM_{Punzi} = \frac{\epsilon(t)}{a/2 + \sqrt{B(t)}}, \tag{4.5}$$

where a is the desired confidence level (CL) in units of standard deviations corresponding to a one-sided Gaussian test, while $\epsilon(t)$ and $B(t)$ are respectively the signal efficiency and the number of remaining background events as a function of a selection t. A 90% CL ($a = 1.64$) was chosen for this optimisation. The advantage of using this particular figure of merit is that it does not rely on the *a priori* unknown branching fraction of signal processes.

5. Signal extraction method

The ultimate goal of this analysis is to extract the signal yields of searched $B^+ \to K^+\tau^\pm\mu^\mp$ processes from experimental Belle dataset. In the method presented in this thesis, the signal decays will be observed indirectly; in characteristics of the recoiling B_{tag}. Applying all selection criteria, described in chapter 4, results in flat M_{tag} distribution for all background components, while the distribution for signal events shows a prominent structure around true B meson mass, $M_{tag} \approx 5.279GeV$. Exploiting these shape differences allows to use M_{tag} distribution to separate signal from background events. The signal yield will be extracted by an unbinned extended likelihood (UEML) fit to M_{tag} distribution, for events that pass the final selection: $\mathrm{A.1} - \mathrm{A.4} \wedge \mathrm{B.1} - \mathrm{B.4}$ and lie in $-0.25 < \Delta\mathrm{E_{tag}} < 0.05\mathrm{GeV}$ window.

5.1. Fitting procedure

As it was mentioned, in this analysis, the signal yields will be extracted by an unbinned extended maximum likelihood (UEML) fit to M_{tag} distribution. This technique has become popular in High Energy Physics, especially for rare signal searches, as it correctly takes into account the Poisson distribution of the number of observed events. The likelihood function was constructed as follows:

$$\mathcal{L} = \frac{(N_{sig} + N_{bkg})^N}{N!} e^{-(N_{sig}+N_{bkg})} \prod_{i=1}^{N} \left[\frac{N_{sig}}{N_{sig} + N_{bkg}} f_s(M^i_{tag}) + \frac{N_{bkg}}{N_{sig} + N_{bkg}} f_b(M^i_{tag}) \right], \quad (5.1)$$

where index i runs over all fitted events, N. f_s and f_b are the probability density functions (PDFs) for signal and background components, whose shape parameters were fixed from fits to Signal and Generic Monte Carlo samples, as it will be described in section 5.2. Free parameters of the fit are the number of signal and background events; N_{sig} and N_{bkg}. The goal of the UEML method is to find the values of N_{sig} and N_{bkg} that maximise the likelihood, or in other words, that best describe the observed data.

5.2. Probability density functions for signal and background

The advantage of using M_{tag} distribution for signal extraction is its known parametrisation for both signal and background components:

- Signal component, f_s, was described using the Crystal Ball function [36]:

$$CB(x;J,A,M,W) = \begin{cases} C\frac{\left(\frac{J}{|A|}\right)^J e^{-\frac{A^2}{2}}}{\left(\frac{J}{|A|}-|A|-\frac{x-M}{W}\right)^J} & \text{for } (x-M)/W < -|A| \\ \\ Ce^{-\frac{1}{2}\left(\frac{x-M}{W}\right)^2} & \text{for } (x-M)/W > -|A|, \end{cases} \tag{5.2}$$

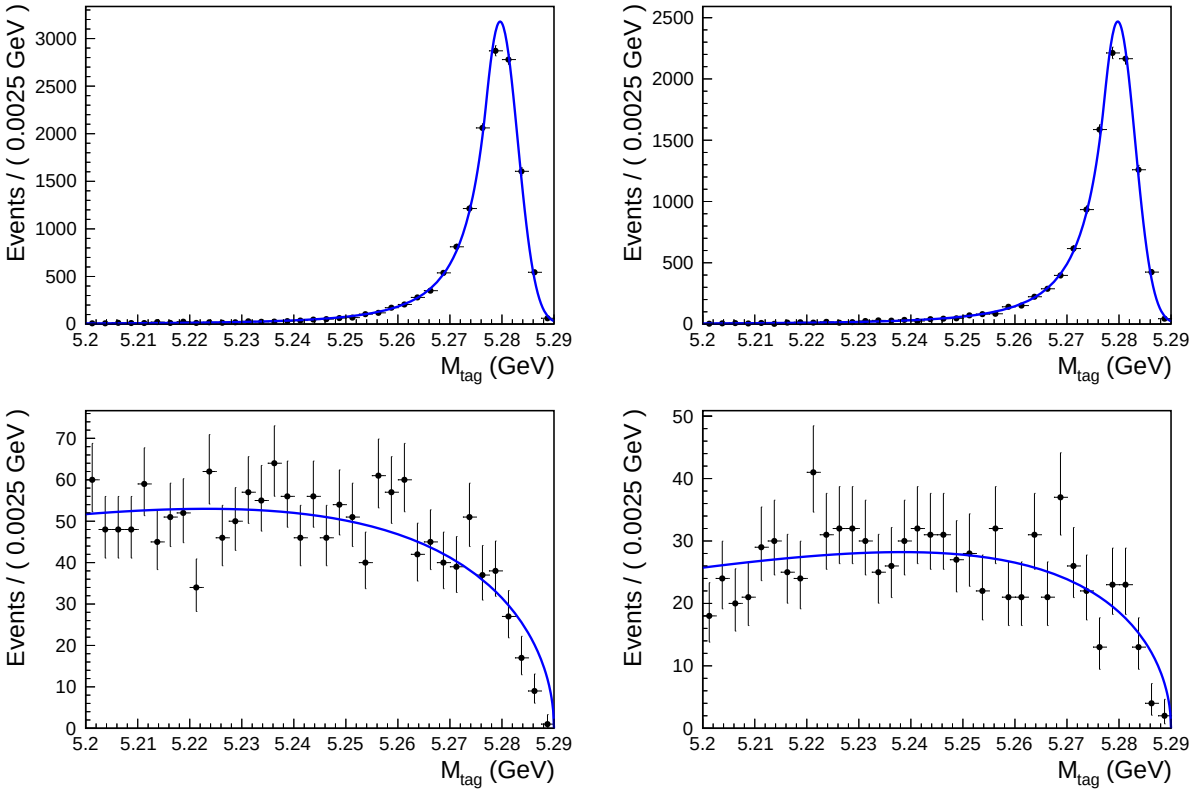
  where C is the normalisation factor and J, A, M, W are the shape parameters. The Crystal Ball function is effectively a modified Gaussian curve, that is often used to model B meson decays to final states with a large number of neutral particles. It consist of a Gaussian core and a power-low tail below a certain threshold. Parameters M and W represent respectively the mean and the width of the Gaussian part, while A and J describe the shape of the power-low tail.

- To model the background component, f_b, the Argus function was used [37]:

$$AR(x;\xi) = D\sqrt{1-\left(\frac{x}{E_{beam}}\right)^2}\; e^{\xi\left[1-\left(\frac{x}{E_{beam}}\right)^2\right]} \tag{5.3}$$

  where D is the normalisation factor and the parameter ξ controls the shape of the distribution.

The shape parameters of the Crystal Ball and Argus functions were determined by fitting the M_{tag} distributions for events reconstructed respectively in Signal and Generic Monte Carlo samples. The fitted distributions are shown in Figure 5.1. The extracted parameters are reported in Table 5.1 and Table 5.2.



(a) $B^+ \to K^+\tau^+(\to \pi^+\bar{\nu}_\tau)\mu^-$ (b) $B^+ \to K^+\tau^-(\to \pi^-\nu_\tau)\mu^+$

Figure 5.1. Distributions of M_{tag} for events that pass the final selection ($A.1 - A.4 \wedge B.1 - B.4$) and lie in $-0.25GeV < \Delta E_{tag} < 0.05GeV$ window. (top) signal fitted with the Crystal Ball function, (bottom) background fitted with the Argus function. Three streams of Generic Monte Carlo were used for evaluating background shape.

Table 5.1. Crystal Ball function shape parameters for individual decay modes obtained from signal MC samples.

Decay channel	J	A	M [GeV]	W [MeV]
$B^+ \to K^+\tau^+\mu^-$	3.03 ± 0.15	0.792 ± 0.033	5.2796 ± 0.0001	3.384 ± 0.071
$B^+ \to K^+\tau^+\mu^-$	3.08 ± 0.16	0.763 ± 0.029	5.2797 ± 0.0001	3.325 ± 0.058

Table 5.2. Argus function shape parameter for individual decay modes obtained from generic MC samples.

Decay mode	ξ
$B^+ \to K^+\tau^+\mu^-$	-19.4 ± 2.7
$B^+ \to K^+\tau^-\mu^+$	-25.2 ± 3.7

5.3. Fit validation

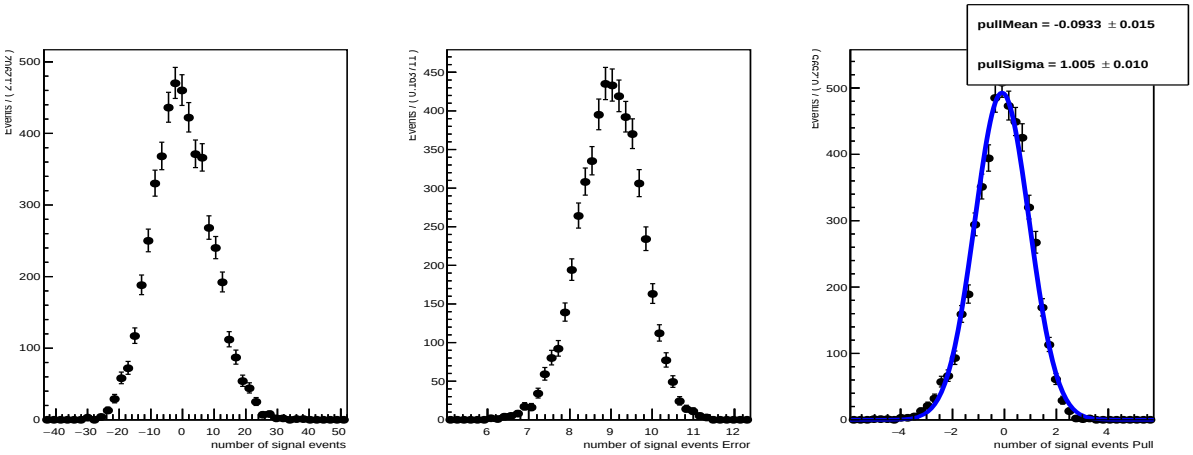
In order to verify the stability of the fitting algorithm, several tests were performed. The procedure relies on generating pseudo data sets- the so-called Toy MC experiments. In each Toy MC experiment, events are generated according to their PDF shapes, established in section 5.2, with the number of signal and background events being Poisson distributed around the expected (input) values. Generated events are then fitted using the same PDF parameters. The procedure allows to verify if the signal and background yields, N_{sig} and N_{bkg}, extracted from the fit, are consistent with the true (generated) values and if the uncertainties of those are properly estimated.

Robustness of the fit was examined by generating 5000 Toy MC experiments. The procedure was repeated several times for different number of input signal events, varying from 0 to 20, while the input number of background events was set to the expected number of background events in data.

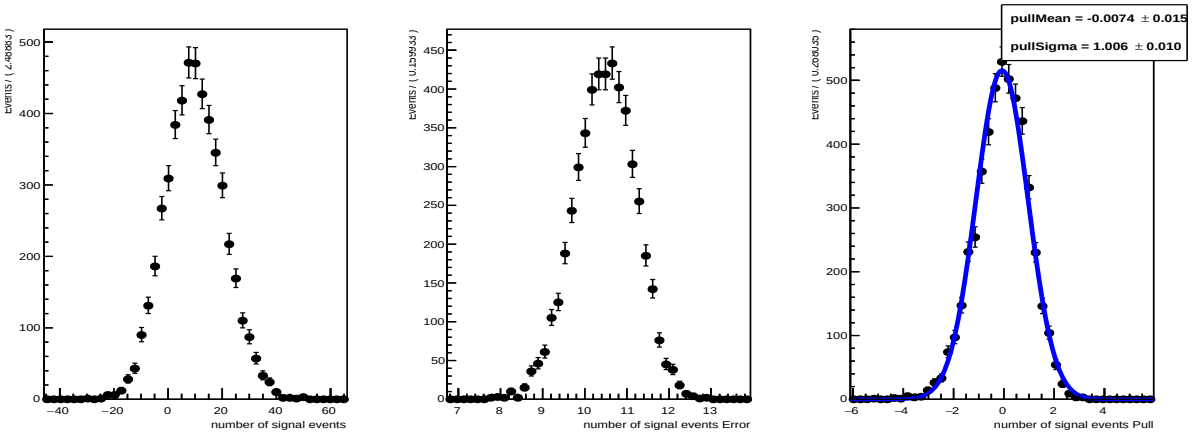
The results of the Toy MC studies are illustrated in Figure 5.2 and Figure 5.3 that show the distributions of the extracted number of signal events (N_{sig}^{fit}), the uncertainty of the extracted number of signal events (σ^{fit}) and the pull distributions defined as the difference between the fitted and expected number of events in units of uncertainty:

$$N_{sig}^{pull} = \frac{N_{sig}^{fit} - N_{sig}^{exp}}{\sigma^{fit}}. \tag{5.4}$$

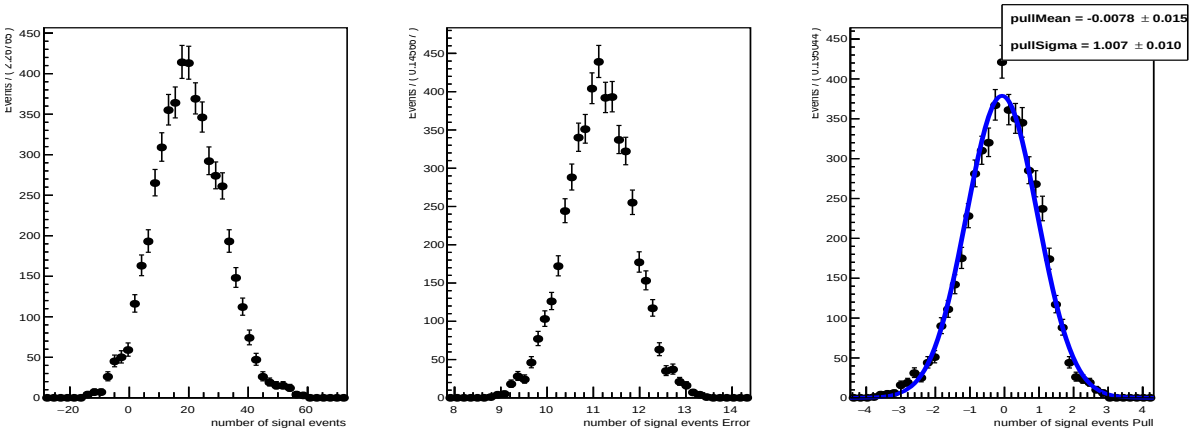
Three different cases for input number of signal events $N_{sig}^{gen} = 0, 10, 20$ were shown for illustration. The mean value and σ of the pull distributions, being consistent with 0 and 1, indicate that the there is no bias in fitting procedure and the fit uncertainty of is properly estimated.



(a) $N_{sig}^{gen} = 0$

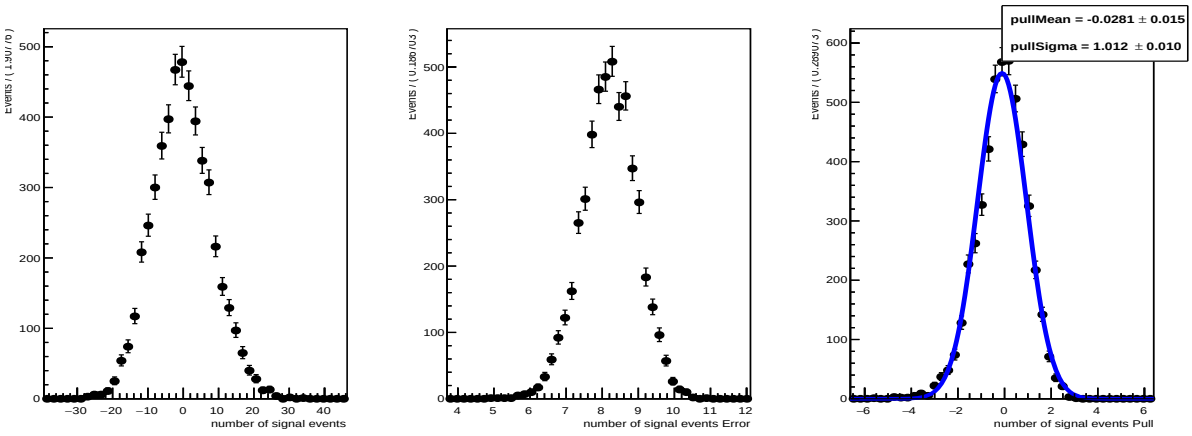


(b) $N_{sig}^{gen} = 10$

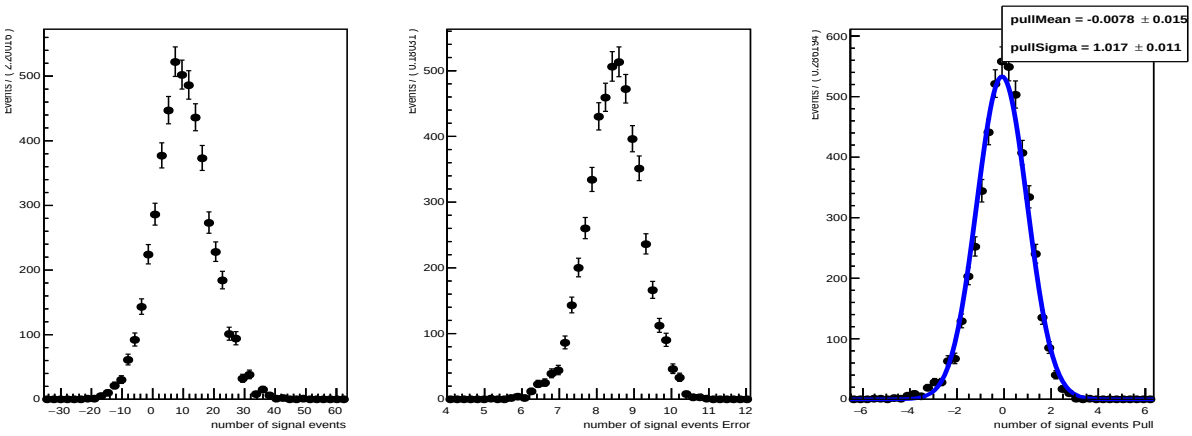


(c) $N_{sig}^{gen} = 20$

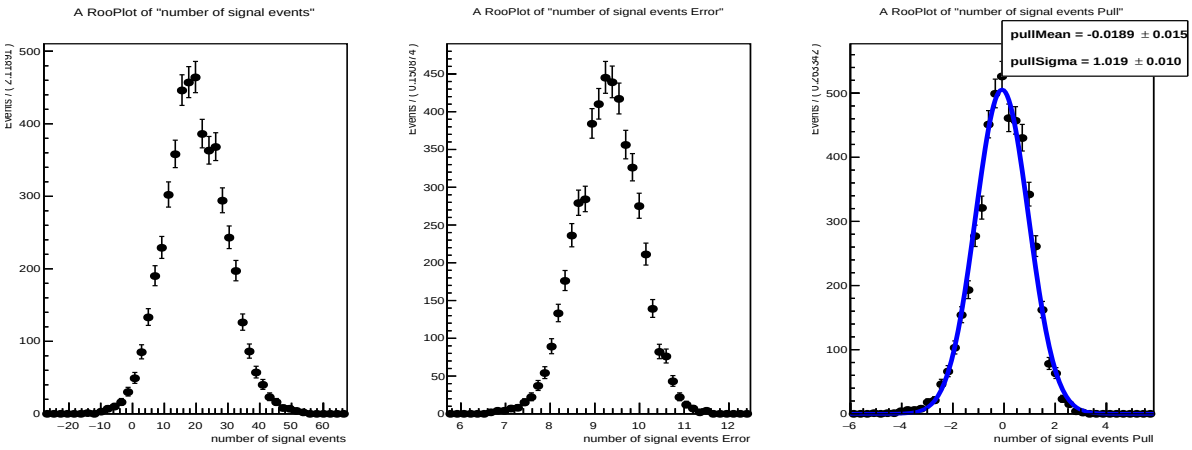
Figure 5.2. Distributions of N_{sig}^{fit} (left), σ^{fit} (center) and $(N_{sig} - N_{sig}^{exp})/\sigma^{fit}$ (right) for $B^+ \to K^+\tau^+(\to \pi^+\bar{\nu}_\tau)\mu^-$ mode obtained from 5000 toy MC experiments.



(a) $N_{sig}^{gen} = 0$



(b) $N_{sig}^{gen} = 10$

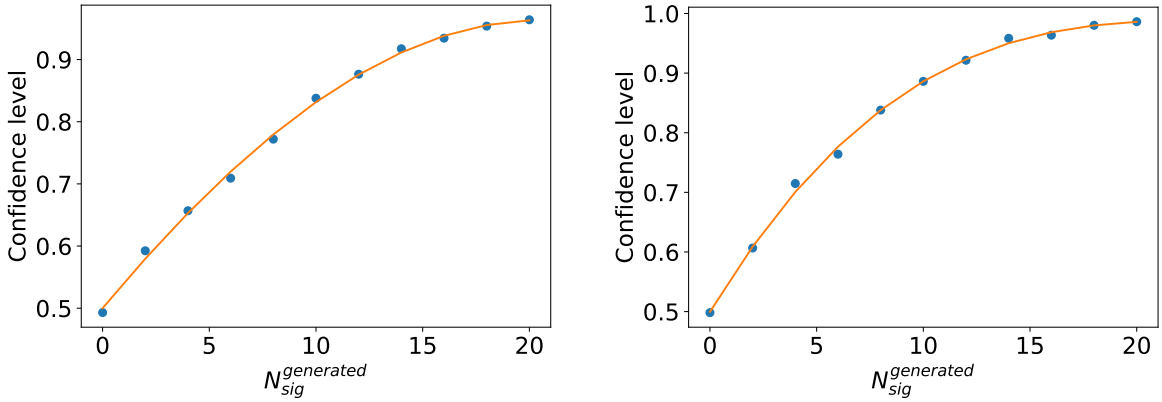


(c) $N_{sig}^{gen} = 20$

Figure 5.3. Distributions of N_{sig}^{fit} (left), σ^{fit} (center) and $(N_{sig} - N_{sig}^{exp})/\sigma^{fit}$ (right) for $B^+ \to K^+\tau^-(\to \pi^-\bar{\nu}_\tau)\mu^+$ mode obtained from 5000 toy MC experiments.

5.4. Upper limit estimation

Toy MC ensembles were also used to calculate the expected upper limit (UL) on the number of signal events with the frequentist method. In this method, a scan through a range of possible signal yields is done by generating 5000 Toy experiments with different number of signal events, as described in section 5.3. For each value of input signal events, the fraction of experiments with the fitted signal yield larger than zero was calculated. These fractions, which correspond to confidence levels (CLs) are plotted as a function of input signal events in Figure 5.4. The input number of signal events for which 90% of experiments give a signal yield larger than zero is taken as the 90% CL upper limit [1].



(a) $B^+ \to K^+\tau^+(\to \pi^+\bar{\nu}_\tau)\mu^-$ (b) $B^+ \to K^+\tau^-(\to \pi^-\bar{\nu}_\tau)\mu^+$

Figure 5.4. Confidence level vs $N_{sig}^{generated}$ obtained from 5000 toy MC experiments.

Upper limit on signal branching fraction was calculated using the following formula:

$$\mathcal{B}^{U.L} = \frac{N_{sig}^{U.L.}}{N_{B\bar{B}} \times \mathcal{B}(\tau \to \pi\nu) \times \epsilon} \tag{5.5}$$

where $N_{B\bar{B}} = 7.7 \times 10^8$ is the number of B$\bar{\text{B}}$ pairs produced in full Belle dataset, $\mathcal{B}(\tau \to \pi\nu) = 10.82\%$ is the branching fraction for reconstructed τ decay and ϵ is the reconstruction efficiency evaluated from signal Monte Carlo separately for each B decay mode ($B^+ \to K^+\tau^+\mu^-$ and $B^+ \to K^+\tau^-\mu^+$). The estimated upper limits are summarised in Table 5.3.

Table 5.3. Upper limit estimates for signal branching fraction.

Decay mode	ϵ	$N_{sig}^{U.L.}$	$\mathcal{B}^{U.L}$
$B^+ \to K^+\tau^+(\to \pi^+\bar{\nu}_\tau)\mu^-$	2.8×10^{-3}	13.1	5.6×10^{-5}
$B^+ \to K^+\tau^-(\to \pi^-\nu_\tau)\mu^+$	2.2×10^{-3}	10.9	6.0×10^{-5}

[1] This way, the quoted upper limit is calculated under assumption of no observed signal.

6. Summary

In this thesis, the novel analysis method for the search of $B^+ \to K^+\tau^\pm\mu^\mp$ processes that can be applied in B factory environment was presented. The distinct kinematics of signal decays was exploited to construct observables that can be used to identify $B^+ \to K^+\tau^\pm\mu^\mp$ processes, in a way that, in principle, does not require tag-side reconstruction at all. A loose (inclusive) B_{tag} reconstruction was applied only to improve the signal-to-background ratio and for signal yield extraction.

The event reconstruction and selection was optimised based on Monte Carlo simulated events, dedicated for Belle experiment. The expected 90% CL upper limits on signal branching fractions (assuming phase space signal decay model), calculated under assumption of no observed signal in data, were found to be $\mathcal{B}(B^+ \to K^+\tau^+\mu^-) = 5.6\times10^{-5}$ and $\mathcal{B}(B^+ \to K^+\tau^-\mu^+) = 6.0\times10^{-5}$, where only statistical uncertainties were considered. These values, although should be treated with care as they do not include systematic uncertainties, are comparable with the previous measurements done at BaBAR and LHCb (Table 1.1). It is worth mentioning that the previous measurements were done using inclusive "one-prong" τ decays, while the studies presented in this thesis were performed only with a single $\tau^- \to \pi^-\nu_\tau$ mode, which effectively reduces the available dataset 8 times (the branching fraction for τ decays to a "one-prong" final state is approximately 85%).

The expected sensitivities can be further improved by applying more advanced selection techniques. An interesting approach would be to replace some of the rectangular selections by multivariate algorithm(s), that would take advantage of the correlations between variables used for the selection. These studies, although interesting in the long run, require more work and were not yet addressed in this thesis. Introducing too many variables or complicated, non-linear selectors can largely increase the eventual systematic uncertainty of the measurement, especially in case of unreliable Monte Carlo simulation. Therefore, at this early stage of the analysis, rather than aiming for the best upper limit estimate, the stress was put on verifying the new measurement technique, understanding the main sources of background and describing them properly.

A. B momentum recovery - derivation

A sequential decay $\mathrm{B_{sig}} \to \mathrm{K}\tau(\to \mathrm{h}\nu)\mu$ satisfies the following four-momentum conservation rules:

$$\begin{cases} \mathrm{p_B^{sig}} = \mathrm{p_K} + \mathrm{p}_\tau + \mathrm{p}_\mu & \text{(A.1)} \\ \mathrm{p}_\tau = \mathrm{p}_\nu + \mathrm{p_h} & \text{(A.2)} \end{cases}$$

Using A.1, A.2 and recalling that $\mathrm{m}^2 \equiv \mathrm{p}^2$, the following relations can be written:

$$\begin{cases} \mathrm{m}_\nu^2 = \mathrm{p}_\nu^2 = (\mathrm{p_B^{sig}} - \mathrm{p_{Kh\mu}})^2 & \text{(A.3)} \\ \mathrm{m}_\tau^2 = \mathrm{p}_\tau^2 = (\mathrm{p_B^{sig}} - \mathrm{p_{K\mu}})^2 & \text{(A.4)} \end{cases}$$

where $\mathrm{m_B}$, m_τ are the masses of B and τ, respectively and the shorthand $\mathrm{p_{Kh(\mu)}} \equiv \mathrm{p_K} + \mathrm{p_h}$ $(+\mathrm{p}_\mu)$ was introduced to simplify the notation. Evaluating the right-hand side of A.3, A.4 one gets:

$$\begin{cases} \mathrm{m}_\nu^2 = -2\mathrm{E_B E_{Kh\mu}} + \mathrm{m_B^2} + \mathrm{m_{Kh\mu}^2} + 2|\mathbf{p}_\mathrm{B}||\mathbf{p}_\mathrm{Kh\mu}| \cos\angle(\mathbf{p_B^{sig}}, \mathbf{p}_\mathrm{Kh\mu}) & \text{(A.5)} \\ \mathrm{m}_\tau^2 = -2\mathrm{E_B E_{K\mu}} + \mathrm{m_B^2} + \mathrm{m_{K\mu}^2} + 2|\mathbf{p}_\mathrm{B}||\mathbf{p}_\mathrm{K\mu}| \cos\angle(\mathbf{p_B^{sig}}, \mathbf{p}_\mathrm{K\mu}) & \text{(A.6)} \end{cases}$$

Under assumption that $\mathrm{E_B} = \mathrm{E_{beam}}$ and $\mathrm{m}_\nu = 0$ it is possible to express $\cos\angle(\mathbf{p_B^{sig}}, \mathbf{p}_\mathrm{Kh\mu})$ and $\cos\angle(\mathbf{p_B^{sig}}, \mathbf{p}_\mathrm{K\mu})$ using measured quantities:

$$\begin{cases} \cos\angle(\mathbf{p_B^{sig}}, \mathbf{p}_\mathrm{Kh\mu}) = \dfrac{2\mathrm{E_{beam}E_{Kh\mu}} - \mathrm{m_B^2} - \mathrm{m_{Kh\mu}^2}}{2|\mathbf{p}_\mathrm{B}||\mathbf{p}_\mathrm{Kh\mu}|} & \text{(A.7)} \\ \cos\angle(\mathbf{p_B^{sig}}, \mathbf{p}_\mathrm{K\mu}) = \dfrac{2\mathrm{E_{beam}E_{K\mu}} - \mathrm{m_B^2} - \mathrm{m_{K\mu}^2} + \mathrm{m}_\tau^2}{2|\mathbf{p}_\mathrm{B}||\mathbf{p}_\mathrm{K\mu}|} & \text{(A.8)} \end{cases}$$

As it was mentioned in chapter 3, the system of equations A.7, A.8 indicate that in $B \to K\tau(\to h\nu)\mu$ decay, the B meson momentum is simultaneously constrained on cones around $\mathbf{p}_{K\mu}$ and $\mathbf{p}_{Kh\mu}$. Two kinematically allowed configurations of $\mathbf{p}_B^{sig}$ are given by the intersection of two cones (Figure A.1).

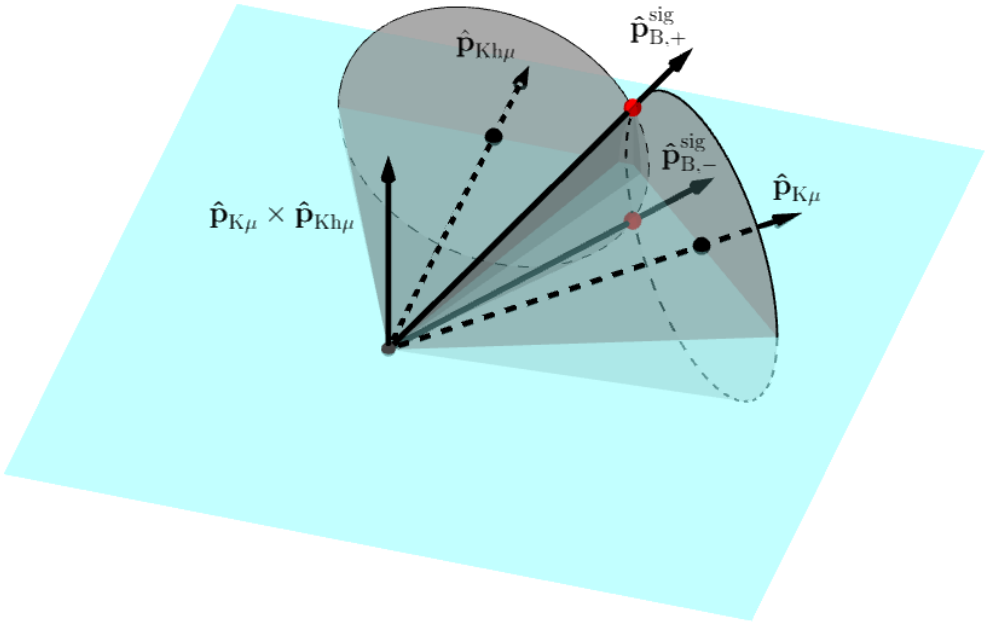


Figure A.1. Schematic view of B meson momentum determination procedure. The signal B meson momentum, $\mathbf{p}_B^{sig}$ is constrained on a cone around $\mathbf{p}_{Kh\mu}$ and $\mathbf{p}_{K\mu}$ momenta. The intersection of two cones defines two kinematically allowed configurations of $\mathbf{p}_B^{sig}$, symmetric with respect to the plane spanned by $\hat{\mathbf{p}}_{\mathrm{K}\mu}$ and $\hat{\mathbf{p}}_{\mathrm{Kh}\mu}$ vectors, that were denoted as $\mathbf{p}_{B,\pm}^{sig}$.

To find the solutions for $\mathbf{p}_{B,\pm}^{sig}$ it is useful to first introduce two bases:

$$\mathcal{A} = \{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \hat{\mathbf{a}}_3\} = \{\frac{\hat{\mathbf{p}}_{K\mu} \times \hat{\mathbf{p}}_{Kh\mu}}{|\hat{\mathbf{p}}_{K\mu} \times \hat{\mathbf{p}}_{Kh\mu}|}, \frac{\hat{\mathbf{p}}_{K\mu} - cos\angle(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu})\hat{\mathbf{p}}_{Kh\mu}}{|\hat{\mathbf{p}}_{K\mu} \times \hat{\mathbf{p}}_{Kh\mu}|}, \hat{\mathbf{p}}_{Kh\mu}\} \tag{A.9}$$

$$\mathcal{B} = \{\hat{\mathbf{b}}_1, \hat{\mathbf{b}}_2, \hat{\mathbf{b}}_3\} = \{\frac{\hat{\mathbf{p}}_{K\mu} \times \hat{\mathbf{p}}_{Kh\mu}}{|\hat{\mathbf{p}}_{K\mu} \times \hat{\mathbf{p}}_{Kh\mu}|}, \frac{cos\angle(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu})\hat{\mathbf{p}}_{K\mu} - \hat{\mathbf{p}}_{Kh\mu}}{|\hat{\mathbf{p}}_{K\mu} \times \hat{\mathbf{p}}_{Kh\mu}|}, \hat{\mathbf{p}}_{K\mu}\} \tag{A.10}$$

where $\hat{\mathbf{a}}_3 \equiv \hat{\mathbf{p}}_{\mathrm{Kh}\mu}$, $\hat{\mathbf{b}}_3 \equiv \hat{\mathbf{p}}_{\mathrm{K}\mu}$; $\hat{\mathbf{a}}_1$ and $\hat{\mathbf{b}}_1$ are colinear and orthogonal to $\mathbf{p}_{\mathrm{K}\mu} - \mathbf{p}_{\mathrm{Kh}\mu}$ plane while $\hat{\mathbf{a}}_2$ and $\hat{\mathbf{b}}_2$ were chosen so that both bases are orthonormal and have right-handed orientation. Such choice allows to conveniently parametrise two cones, shown in Figure A.1, as:

$$[\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}(\phi)]_{\mathcal{A}} = |\mathbf{p}_{\mathrm{B}}| \begin{bmatrix} \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu})\cos\phi \\ \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu})\sin\phi \\ \cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \end{bmatrix}, \quad \phi \in [0, 2\pi]; \tag{A.11}$$

$$[\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}(\phi')]_{\mathcal{B}} = |\mathbf{p}_{\mathrm{B}}| \begin{bmatrix} \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{K\mu}) \cos\phi' \\ \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{K\mu}) \sin\phi' \\ \cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{K\mu}) \end{bmatrix}, \quad \phi' \in [0, 2\pi]. \tag{A.12}$$

Effectively, Equation A.11 (Equation A.12) represents the coordinates of the B meson momentum, as constrained by Equation A.7 (Equation A.8), with respect to basis $\mathcal{A}$ ($\mathcal{B}$). To calculate the intersection, the coordinates A.11, A.12 should be represented in the same basis. This can be done, for example, by moving the coordinates A.11 from $\mathcal{A}$ to $\mathcal{B}$ with the following transformation:

$$[\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}(\phi)]_{\mathcal{B}} = \hat{\mathcal{T}}_{\mathcal{A}\to\mathcal{B}} [\mathbf{p}_B^{sig}(\phi)]_{\mathcal{A}}$$

$$= |\mathbf{p}_B| \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\angle(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu}) & -\sin\angle(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu}) \\ 0 & \sin\angle(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu}) & \cos\angle(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu}) \end{bmatrix} \begin{bmatrix} \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \cos\phi \\ \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \sin\phi \\ \cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \end{bmatrix}$$

$$= |\mathbf{p}_B| \begin{bmatrix} \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \cos\phi \\ \cos\angle(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu}) \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \sin\phi - \sin\angle(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu}) \cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \\ \sin\angle(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu}) \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \sin\phi + \cos\angle(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu}) \cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \end{bmatrix} \tag{A.13}$$

where $\hat{\mathcal{T}}_{\mathcal{A}\to\mathcal{B}}$ is the transformation matrix, that specifies the change of coordinates of $\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}(\phi)$ under the change of basis from $\mathcal{A}$ to $\mathcal{B}$. Requiring that:

$$[\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}(\phi)]_{\mathcal{B}} = [\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}(\phi')]_{\mathcal{B}} \tag{A.14}$$

allows to constrain the value of $\sin\phi$:

$$\sin\phi = \frac{\cos\angle(\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}, \mathbf{p}_{\mathrm{K}\mu}) - \cos\angle(\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}, \mathbf{p}_{\mathrm{Kh}\mu}) \cos(\mathbf{p}_{\mathrm{K}\mu}, \mathbf{p}_{\mathrm{Kh}\mu})}{\sin\angle(\mathbf{p}_{\mathrm{K}\mu}, \mathbf{p}_{\mathrm{Kh}\mu}) \sin\angle(\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}, \mathbf{p}_{\mathrm{Kh}\mu})}. \tag{A.15}$$

Once the value of $\sin\phi$ is fixed, it is possible to express the coordinates of $\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}$ with respect to basis $\mathcal{A}$ with a two-fold ambiguity [1]:

$$[\mathbf{p}_{\mathrm{B},\pm}^{\mathrm{sig}}]_{\mathcal{A}} = |\mathbf{p}_{\mathrm{B}}| \begin{bmatrix} \pm\sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \sqrt{1 - \sin^2\phi} \\ \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \sin\phi \\ \cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu}) \end{bmatrix}. \tag{A.16}$$

[1] Equivalently, it is possible to constrain the value of $\sin\phi'$ and express the coordinates of $\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}$ with respect to basis $\mathcal{B}$ (Equation A.12). The choice of basis does not change the final result ($\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}$) so one is free to choose between $\mathcal{A}$ or $\mathcal{B}$.

Finally, $\mathbf{p}_{B,\pm}^{sig}$ can be expressed in more elegant form, as:

$$\mathbf{p}_{B,\pm}^{sig} = |\mathbf{p}_B| \left(\pm u\ \hat{\mathbf{a}}_1 + v\ \hat{\mathbf{a}}_2 + w\ \hat{\mathbf{a}}_3\right) \tag{A.17}$$

where $\{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \hat{\mathbf{a}}_3\} = \{\frac{\hat{\mathbf{p}}_{K\mu} \times \hat{\mathbf{p}}_{Kh\mu}}{|\hat{\mathbf{p}}_{K\mu} \times \hat{\mathbf{p}}_{Kh\mu}|}, \frac{\hat{\mathbf{p}}_{K\mu} - \cos\angle(\mathbf{p}_{K\mu}, \mathbf{p}_{Kh\mu})\hat{\mathbf{p}}_{Kh\mu}}{|\hat{\mathbf{p}}_{K\mu} \times \hat{\mathbf{p}}_{Kh\mu}|}, \hat{\mathbf{p}}_{Kh\mu}\}$ are basis vectors introduced in Equation A.9, while parameters $u = \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu})\sqrt{1 - \sin^2\phi}$, $v = \sin\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu})\sin\phi$, $w = \cos\angle(\mathbf{p}_B^{sig}, \mathbf{p}_{Kh\mu})$ represent the coordinates of $\mathbf{p}_{B,\pm}^{sig}$ with respect to basis $\mathcal{A} = \{\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \hat{\mathbf{a}}_3\}$, as given by Equation A.16.

With the procedure described above it is possible to analytically find two kinematically allowed configurations of $\mathbf{p}_{\mathrm{B}}^{\mathrm{sig}}$ that are symmetric with respect to the plane spanned by $\mathbf{p}_{\mathrm{K}\mu}$, $\mathbf{p}_{\mathrm{Kh}\mu}$ vectors (Figure A.1). Although the remaining two-fold ambiguity cannot be resolved experimentally, the recovered momentum provides a useful information about B_{sig} flight direction that can be further used do discriminate background events.

Bibliography

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

[2]M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction", Prog. Theor. Phys. **49**, 652–657 (1973).

[3]X.-G. He, G. Valencia, and Y. Wang, "Lepton flavor violating tau and B decays and heavy neutrinos", Phys. Rev. D **70**, 113011 (2004).

[4]J. Lees et al., "Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons", Phys. Rev. D **88**, 072012 (2013).

[5]M. Huschle et al., "Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle", Phys. Rev. D **92**, 072014 (2015).

[6]R. Aaij et al., "Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$", Phys. Rev. Lett. **115**, [Erratum: Phys.Rev.Lett. 115, 159901 (2015)], 111803 (2015).

[7]R. Aaij et al., "Test of Lepton Flavor Universality by the measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays", Phys. Rev. D **97**, 072013 (2018).

[8]R. Aaij et al., "Measurement of the ratio of the $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ branching fractions using three-prong τ-lepton decays", Phys. Rev. Lett. **120**, 171802 (2018).

[9]R. Aaij et al., "Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays", Phys. Rev. Lett. **113**, 151601 (2014).

[10]R. Aaij et al., "Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays", JHEP **08**, 055 (2017).

[11]R. Aaij et al., "Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays", Phys. Rev. Lett. **122**, 191801 (2019).

[12]L. Calibbi, A. Crivellin, and T. Li, "Model of vector leptoquarks in view of the B-physics anomalies", Phys. Rev. D **98**, 115002 (2018).

[13]C. Cornella, J. Fuentes-Martin, and G. Isidori, "Revisiting the vector leptoquark explanation of the B-physics anomalies", JHEP **07**, 168 (2019).

[14]B. Aubert et al., "Search for the decay $B^+ \to K^+\tau^\mp\mu^\pm$", Phys. Rev. Lett. **99**, 201801 (2007).

[15]J. Lees et al., "A search for the decay modes $B^\pm \to h^\pm\tau^\pm\ell$", Phys. Rev. D **86**, 012004 (2012).

[16]R. Aaij et al., "Search for the lepton flavour violating decay $B^+ \to K^+\mu^-\tau^+$ using B^{*0}_{s2} decays", JHEP **06**, 129 (2020).

[17]T. E. Browder and K. Honscheid, "B mesons", Prog. Part. Nucl. Phys. **35**, edited by A. Faessler, 81–220 (1995).

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

[19]A. Bevan et al., "The Physics of the B Factories", Eur. Phys. J. C **74**, 3026 (2014).

[20]M. Feindt et al., "A Hierarchical NeuroBayes-based Algorithm for Full Reconstruction of B Mesons at B Factories", Nucl. Instrum. Meth. A **654**, 432–440 (2011).

[21]A. Matyja et al., "Observation of $B^0 \to D^{*-}\tau^+\nu_\tau$ decay at Belle", Phys. Rev. Lett. **99**, 191807 (2007).

[22]A. Bozek et al., "Observation of $B^+ \to \bar{D}^{*0}\tau^+\nu_\tau$ and Evidence for $B^+ \to \bar{D}^0\tau^+\nu_\tau$ at Belle", Phys. Rev. D **82**, 072005 (2010).

[23]A. Abdesselam et al., "Measurement of the D^{*-} polarization in the decay $B^0 \to D^{*-}\tau^+\nu_\tau$", in 10th International Workshop on the CKM Unitarity Triangle (Mar. 2019), arXiv:1903.03102 [hep-ex].

[24]S. Kurokawa and E. Kikutani, "Overview of the KEKB accelerators", Nucl. Instrum. Meth. A **499**, 1–7 (2003).

[25]T. Abe et al., "Achievements of KEKB", PTEP **2013**, 03A001 (2013).

[26]A. Abashian et al., "The Belle Detector", Nucl. Instrum. Meth. A **479**, 117–232 (2002).

[27]D. Lange, "The EvtGen particle decay simulation package", Nucl. Instrum. Meth. A **462**, edited by S. Erhan, P. Schlein, and Y. Rozen, 152–155 (2001).

[28]T. Sjostrand et al., "High-energy physics event generation with PYTHIA 6.1", Comput. Phys. Commun. **135**, 238–259 (2001).

[29]E. Barberio and Z. Was, "PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0", Comput. Phys. Commun. **79**, 291–308 (1994).

[30]R. Brun et al., "GEANT3", CERN Report DD-EE-84-1 (1987).

[31]E. Nakano, "Belle PID", Nucl. Instrum. Meth. A **494**, edited by G. Fedotovich and B. Khazin, 402–408 (2002).

[32]K. Hanagaki et al., "Electron identification in Belle", Nucl. Instrum. Meth. A **485**, 490–503 (2002).

[33]A Abashian et al., "Muon identification in the Belle experiment at KEKB", Nucl. Instrum. Meth. A **491**, 69–82 (2002).

[34]G. C. Fox and S. Wolfram, "Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes", Phys. Rev. Lett. **41**, 1581 (1978).

[35]G. Punzi, "Sensitivity of searches for new signals and its optimization", eConf **C030908**, MODT002 (2003), arXiv:physics/0308063.

[36]T. Skwarnicki, "A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances", PhD thesis (Cracow, INP, 1986).

[37]H. Albrecht et al., "Search for $b \to s\gamma$ in Exclusive Decays of B Mesons", Phys. Lett. B **229**, 304–308 (1989).