Shinshu University

A Search for the Rare Leptonic Decays $B^+ \to \mu^+ \nu_\mu$ and $B^+ \to e^+ \nu_e$

Doctor Thesis

March 2007

Norihiko Satoyama

Division of Science and Technology,
Shinshu University

信州大学

荷電B中間子におけるレプトニック稀崩壊 $B^+ \to \mu^+ \nu_\mu$ および $B^+ \to e^+ \nu_e$ の探索

博士論文

2007年3月

里山　典彦

信州大学大学院　工学系研究科

Abstract

This thesis presents a search for the $B^+ \to \mu^+\nu_\mu$ and $B^+ \to e^+\nu_e$ decays following $\Upsilon(4S) \to B^+B^-$. In the Standard Model (SM) for the elementary particle physics, these decays proceed by annihilation to a virtual W boson, and these branching fractions are expected. Search for these decays, therefore, can provide sensitivity to the SM parameters and also act as a probe of physics beyond the SM. The data used in this analysis were collected with the Belle detector at KEKB e^+ (3.5 GeV) e^- (8 GeV) asymmetric-energy collider. The data sample consists of an integrated luminosity of 253 fb^{-1} accumulated at the $\Upsilon(4S)$ resonance. The data sample corresponds to 276.6×10^6 $B\bar{B}$ pairs.

As $B^+ \to \mu^+\nu_\mu$ and $B^+ \to e^+\nu_e$ decays are two-body decays, the signal charged leptons have a fixed momentum in the signal B meson rest frame, which is higher than momenta of tracks from all other B decays. The lepton with high momentum is assigned to the signal candidate lepton. All other particles are inclusively reconstructed to the companion B meson, which is the opposite side B meson of the signal side B meson. The dominant background arises from the continuum processes and semileptonic B meson decays. In order to reduce such backgrounds, we use differences of event topology between signal and background events, and require a higher momentum track for signal candidate lepton and so on.

After the evetnt selection, we can find 12 (15) events in the signal region for the muon (electron) mode. There is no significant evidence for the both modes taking into account the number of the background events we expected in the same region. We can set 90 % confidence level upper limits of the branching fractions to $\mathcal{B}(B^+ \to \mu^+\nu_\mu) < 1.7 \times 10^{-6}$ and $\mathcal{B}(B^+ \to e^+\nu_e) < 9.8 \times 10^{-7}$. These upper limits are consistent with the predictions of the SM. The upper limits are the best results to date.

Acknowledgements

First of all I would like to thank Prof. T. Takeshita and associate Prof. Y. Hasegawa for giving me the chance of join to Belle collaboration and I want to thank all of the members in HE lab. in Shinshu University.

I would like to thank every member of the Belle collaboration and KEKB who has been working hard to make the experiment run smoothly and successfully. This thesis would not have been possible without their efforts. Especially I wish to thank everyone who give me great suggestions for my study. When I started to analyze this decay mode in 2003, I has very limited knowledge about experimental particle physics. They has been always patient with me and guided me through every step of the way. I especially thanks: T. Iijima, K. Ikado, K. Okabe and T. Hokuue in Nagoya University, I never forget the days of meeting to report status at Nagoya University and via TV. I would like to thank: S. Villa, I. Nakamura, G. R. Moloney, K. Abe, Y. Sakai, T. Browder and Simon I. Eidelman, the journal paper, which is required to submit this PhD thesis, could not been completed without your great helps. I am also grateful to Y. Miyazaki for your advises to carry on my study.

Last but least, I would like to thank my wife Ryoko, my daughter Nagomi, and my parents Takashi and Yoshiko for always having faith in me and always being there supporting me.

Norihiko Satoyama

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Elementary particle physics is trying to understand fundamental particles and their interactions, the answers to some of the most fundamental questions about Nature, namely what it is made of and what holds it together. In a bit more than a century since the discovery of the electron by J. J. Thomson, the elementary particle physics has come a log way and evolved into a mature scientific field in which many experiments have to be done in collaborations of hundreds of physicists and in which theoretical calculations can take years to improve the accuracy of predictions by a few percent.

Our current knowledge on elementary particle physics is gathered into the Standard Model (SM), a result of an immense experimental and theoretical effort spanning more than fifty years. It is extremely successful in describing basically all gathered experimental data, yet there are strong indications that it is not the final answer to all the questions on the nature of elementary particles and their interactions.

A role of the experimental particle physics is to test our theoretical present knowledge: to estimate validity of our predictions and to point at problems and inconsistencies that can inspire an advance of our understanding. Many experiments have been set up around the world to the test the predictions of different segments of the SM. One of the segments that received special attention in the last few years is the so-called flavor physics, which describes quarks flavor-changing transitions by a mechanism proposed by Kobayashi and Maskawa in 1973 [2].

The formalism of all quark flavor-changing transitions within the SM is governed by the Cabibbo-Kobayashi-Maskawa (CKM) matrix, a unitary matrix with four independent free parameters, which have to be determined by experiments. In 2001 two independent measurements observed a large CP violation in decays of B mesons, confirming that the CKM matrix is complex. The unitarity conditions of the CKM matrix can therefore be graphically represented as triangles in the complex plane. One of the triangles, which can be determined by measurements of B meson decays alone, is known as the *Unitarity Triangle*, and its determination has become the "test bed" the SM predictions.

The Unitarity Triangle can be over-determined by a variety of redundant measurements that determine different angles and sides of the Triangle. If the predictions of the SM are not describing different B meson phenomena consistently, the construction of the triangle will be unsuccessful and would be a clear indication of physics beyond the SM. To spot inconsistencies between predictions for different processes, however, the measurements have to achieve high accuracy and well understood errors.

The measurement of angle ϕ_1 in 2001, which was determined by observed CP violation in decays of B mesons, opened a new theoretically clean way of testing the SM predictions. The side of the Unitarity Triangle that lies opposite to the angle ϕ_1 is determined by the measurement of the matrix element $|V_{ub}|$, one of the smallest CKM matrix elements. While the measurement of ϕ_1 includes loops in its Feynman diagrams that are sensitive to possible new contributions of physics beyond the SM, the measurement of $|V_{ub}|$ can be determined from tree-type diagrams that are insensitive to new physics. Comparison of the measurements of ϕ_1 and $|V_{ub}|$ is therefore an excellent opportunity to test the consistency of the SM predictions.

Two e^+e^- colliders with asymmetric energies of beams (so-called B factories), KEKB and PEP-II, have been set up at KEK (High Enegy Accelerator Research Organization) and SLAC (Stanford Linear Accelerator Center) respectively, to perform precise quantitative studies of B mesons decays. They host the experiments Belle and BaBar, whose main goal is a precise measurement of CP asymmetries in B meson decays. The B mesons are produced in pairs from decays of the $\Upsilon(4S)$ resonance, and the two experiments have so far managed to collect several hundred million decays of B meson pairs. Such a large data sample enables the physicists to perform a large set of various measurements and to search rare decays from B meson, The various measurements and searches are also available to verify the SM and search for new physics. The experiments Belle and BaBar have already observed some new rare decays from B meson. However there are many expected rare decays by the SM and they have not been observed yet. This thesis describes on the analysis of two rare decays from charged B meson which have not been observed yet, $B^+ \to \mu^+\nu_\mu$ and $B^+ \to e^+\nu_e$, and the analysis described in the thesis was performed on a sample collected by the Belle detector at KEK in Japan.

The thesis is organized as follows: in the chapter 2 we first discuss the theory of the SM and the motivation for our study. In the chapter 3 we present the experimental environment and the general event reconstruction techniques used at the Belle detector. We introduce data samples analyzed in our study in the chapter 4. We describe our analysis in the chapter 5. In the section 5.1 in the analysis chapter, we present the techniques of the particle identification and particle reconstructions. In the section 5.2 the event selection is described. After the event selection, we mentioned how many the signal yields are expected and the method to extract the signal yield in the section 5.3. The number of the expected backgrounds is also described in the same section. The systematic uncertainties are described in the section 5.4. In the section 5.5, we set the upper limits on branching fractions for both modes as our analysis results. In the last chapter 6 we come to the conclusion and review the results and propose future improvements.

Chapter 2

Leptonic Decay on Charged *B* Meson System

In this chapter we review why we try to search for the leptonic decay. The search is important for our understanding of the validity of the Standard Model (SM) predictions and may give the new physics beyond the SM. The SM is firstly introduced and we explain the SM predictions for the leptonic decay on charged B meson system. A theory beyond the SM is also introduced.

2.1 Standard Model

The Standard Model (SM) is a set of gauge theories that explain how elementary particles interact with each other through basic interactions. The elementary particles are, according to their quantum-mechanical properties, separated into three groups: fermions, gauge bosons, and the predicted Higgs particle. There are twelve elementary fermions (with their twelve antiparticles) : six leptons and six quarks, which are grouped into three generations,

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix},$$
$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}.$$

The elementary particles in the SM interact through three interactions* : weak, strong and electro-magnetic, by exchanging appropriate gauge bosons pertaining to the interaction. The gauge group describing the interactions is $SU(3)_C \times SU(2)_L \times U(1)_Y$. The group $SU(3)$ denotes Quantum Chromodynamics (QCD), which governs the strong interaction among quarks, while unified electro-weak interactions are characterized by the gauge group $SU(2)_L \times U(1)_Y$.

The SM is a result of a joint effort of theoretical and experimental physicists over the last 50 years. Its predictions are continuously confronted by new data and experimental methods. Until recently, all the measured results could be described, within theoretical and experimental errors,

*A unified theory including gravitational interaction has not been achieved yet. Since the gravitational interaction is much weaker than the other three at elementary particle level, its omission does not effect the applicability of the SM predictions to phenomena at the energies obtainable at accelerators today.

by the SM predictions. Nevertheless, physicists expect that the SM is not the final theory and they will eventually observe physical processes which need to be described by theories beyond the SM. Recently, the neutrino oscillations have been experimentally confirmed, and shown that neutrinos are not massless particles: to include this, the SM needs to be extended. Other conceptual problems, for example the so-called gauge hierarchy problem, a large number of free parameters of the SM and some cosmological observations hint at the possibility of physical processes that cannot be satisfactorily explained and described by the SM.

There is a wide range of proposed elementary particle processes in which the contributions beyond the SM can arise, and there are important tests of the SM predictions. A set of tests is currently performed in the weak decays of heavy mesons, of which the search for rare decays plays an important part.

2.2 Leptonic Decay of Charged B Meson

The leptonic decay of a charged B belongs to a weak decay. One way to test the SM predictions is to look at the weak interaction. The weak interaction is described within the SM with an exchange of $W^{\pm}$ and Z^0 bosons. Both quarks and leptons are affected by the weak interaction, and it is the only interaction of neutrinos. Weak decays are also the only ones to depend on quark flavor. The leptonic decay on a meson system also indicates the annihilation of two quarks, of which the meson consists, into a $W^{\pm}$ boson and pair creation of a lepton and a neutrino from the $W^{\pm}$ boson. Figure 2.1 shows the Feynman diagram for the leptonic decay of the charged B meson based on the SM prediction.

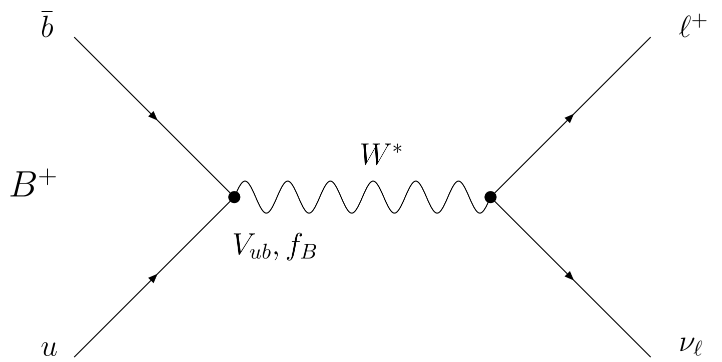


Figure 2.1: Feynman diagram for the leptonic decay on the charged B_d meson based on SM prediction.

The amplitude for the Feynman diagram is of the form

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} V_{ub}[u\gamma_u(1-\gamma^5)\bar{b}][\ell\gamma_u(1-\gamma^5)\nu] \tag{2.1}$$

where G_F is the Fermi coupling constant, V_{ub} is one of the Cabibbo-Kobayashi-Maskawa matrix elements and $u, \bar{b}, \ell$ and ν correspond to the wave functions of themselves. The standard V-A

term induces $B^+ \to \ell^+\nu$ decay via axial-vector current, with

$$< 0|u\gamma_u\gamma^5\bar{b}|B^+ >= i f_B p_B^\mu. \tag{2.2}$$

Equation 2.1 is given as

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} V_{ub} f_B m_\ell [\ell\gamma_u(1-\gamma^5)\nu], \tag{2.3}$$

where this amplitude is proportional to lepton mass (m_ℓ) (helicity suppression). Then the branching fraction is calculated from Equation 2.3 as

$$\mathcal{B}(B^+ \to \ell^+\nu_\ell) = \frac{G_F^2 m_B m_\ell^2}{8\pi}\left(1 - \frac{m_\ell^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B, \tag{2.4}$$

where m_B is B meson mass, $|V_{ub}|$ is the magnitude of one of the Cabibbo-Kobayashi-Maskawa matrix elements, and τ_B is the B^+ lifetime. The V_{ub} is introduced in the following sections.

2.2.1 CKM Matrix

V_{ub} is one of the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements. The theory of the CKM is introduced from this section.

The transformation property under the electroweak gauge group $SU(2)_L \times U(1)_Y$ is different for left and right-handed fermions. The right-handed components of the leptons and quarks are singlets under the weak symmetry $SU(2)_L$, while the left-handed components transform as weak $SU(2)_L$ doublets:

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L \begin{pmatrix} c \\ s' \end{pmatrix}_L \begin{pmatrix} t \\ b' \end{pmatrix}_L . \tag{2.5}$$

The quark mass states are not eigen-states of the weak interaction, therefore the states coupled in the doublets need to be rotated into the weak eigen-state frame, where the rotated states are denoted with a prime (see Equation 2.5). This rotation was first proposed by Cabibbo in 1963 [1] for the case of three quarks that were known at that time, and was later generalized for three quark generations with six quark flavors by Kobayashi and Maskawa (1973) [2], by the introduction of the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix.

The model was proposed when only three quarks were known and was able to predict the existence of six quarks. A large CP violation is arisen in B meson decays [3].

2.2.2 The Origin of the CKM Matrix

The elementary particles in the SM are massless, since mass terms in the Lagrangian break the local gauge invariance. But it was shown that by introducing scalar Higgs fields the particles can, after spontaneous symmetry breaking (SSB), acquire mass by coupling with the Higgs fields. The derivation follows the steps described in Ref. [4].

The mass of a fermion is obtained from the Yukawa coupling between a fermionic fields (e, ν, u or d) and Higgs field (ϕ):

$$\mathcal{L}_Y = -C_{ij}^e(\bar{\ell}_{iL}\phi)e'_{jR} - C_{ij}^u(\bar{q}_{iL}\phi^c)u'_{jR} - C_{ij}^d(\bar{q}_{iL}\phi)d'_{jR} + h.c., \tag{2.6}$$

where u' and d' represent vectors of all up-type and down-type quarks, e is one of the charged leptons, and ℓ and q represent one of the leptons and one of the quarks, respectively. The indices i and j denote the generation of the quark or the lepton, and subscripts L and R denote the left-handed and right-handed particle fields, respectively. The coefficients C_{ij} are three 3×3 matrices that determine the strength of the Yukawa couplings between fermions and Higgs fields (f represents either a charged lepton, an up-type quark or a down-type quark) and can be arbitrary complex matrices. After SSB with weak isospin doublet Higgs fields, the Higgs doublet can be written as follows:

$$\phi^{(c)} \to \frac{1}{\sqrt{2}}(v + H)\chi^{(c)}, \chi = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{2.7}$$

where the Higgs field is split into its vacuum expectation value v and the remaining Higgs field H, which obtaines its mass in the process of SSB. Inserting Equation 2.7 into Equation 2.6 we obtain the following form of the Yukawa part of the Lagrangian:

$$\mathcal{L}_Y = -\left(1 + \frac{H}{v}\right)(\bar{e}'_L \mathcal{M}'_e e'_R + \bar{u}'_L \mathcal{M}'_u u'_R + \bar{d}'_L \mathcal{M}'_d d'_R + h.c.). \tag{2.8}$$

The non-diagonal mass matrices are directly connected to the Yukawa coupling coefficients in Equation 2.6:

$$\mathcal{M}'_f = \frac{v}{\sqrt{2}} C^f_{ij}. \tag{2.9}$$

Since the matrices representing the Yukawa coupling constants C^f_{ij} can be arbitrary, the mass matrices are by default neither diagonal nor symmetric. The absence of right-handed neutrinos results in a diagonalized mass matrix for leptons ($\mathcal{M}'_e$), which means that the lepton fields in the electroweak Lagrangian have also definite mass. This is not the case for quark fields: the quark fields u' and d' in the Yukawa Lagrangian in Equation 2.6 do not have definite mass. To obtain the physical states with definite mass, we preform a unitary transformation using unitary matrices S and T to diagonalize the quark mass matrices $\mathcal{M}'_q$:

$$\mathcal{M}'_q = S^\dagger_q \mathcal{M}_q S_q T_q. \tag{2.10}$$

The matrices S_q transform the gauge (interaction) quark eigen-states ψ'_q into the mass eigen-states ψ_f:

$$\psi_{qL} \equiv S_q \psi'_{qL} \tag{2.11}$$

$$\psi_{qR} \equiv S_q T_q \psi'_{qR}. \tag{2.12}$$

The fact that the interaction quark eigen-states are not the same as the mass eigen-states has important consequences on the electroweak interactions, which can be derived from the Lagrangian term:

$$\mathcal{L} = \bar{\Psi}_L i\gamma^u D^L_\mu \Psi_L + \bar{\Psi}_R i\gamma^u D^R_\mu \Psi_R. \tag{2.13}$$

After explicitly writing the covariant derivatives D^L_μ and D^R_μ, we obtain three types of electroweak interactions, weak charged, weak neutral and electromagnetic interactions. The weak neutral and electromagnetic interactions are not flavor-changing, therefore they have the same form in both physical and interaction bases.

The weak charged interaction, which plays the most important role in semileptonic decays, on the other hand has a different form in the two bases. The corresponding term int the Lagrangian of the weak charged interaction is of the form:

$$\mathcal{L}^{w.c.} = -\frac{g}{\sqrt{2}}(J^{\mu\dagger}W_\mu + J^\mu W_\mu^\dagger), \tag{2.14}$$

where the weak charged current J^μ is coupled to a charged massive boson field W_μ and the strength of the interaction is determined by the coupling constant g.

The quark contribution to this charged current $J^\mu_{w.c.}$ is:

$$J^\dagger_{w.c.} = \bar{u}'_L\gamma_\mu d'_L = \bar{u}_L\gamma_\mu S_\mu S_d^\dagger d_L = \bar{u}_L\gamma_\mu V_{CKM} d_L. \tag{2.15}$$

We define the Cabibbo-Kobayashi-Maskawa matrix $V_{CKM} \equiv S_u S_d^\dagger$, which is a unitary matrix introduced by Kobayashi and Maskawa in 1973 [2] and rotating the down-type quark states, while leaving the up-type quarks unchanged: $d' = V_{CKM} d$.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{2.16}$$

so that the charged current can be written as:

$$J^\dagger_{c.c.} = (\bar{u}\bar{c}\bar{t})_L\,\gamma_\mu \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L . \tag{2.17}$$

The weak charged interaction involves a change of quark flavor between the up-type and down-type quarks, and the V_{CKM} matrix elements determine the strength of the coupling of up-type quarks to down-type quarks. The probability for a flavor transition of the i-th generation up-type quark to a j-th generation down-type quark is proportional to the CKM matrix element squared, $|V_{ij}|^2$.

2.2.3 Parametrization of the CKM Matrix

The CKM matrix is in general a complex $n \times n$ matrix, where n is the number of generations of elementary particles. In the case of three generations there are 18 parameters, but due to unitarity conditions only nine of them are independent, and further five phases can be removed by appropriate rotations of the quark fields, therefore 4 independent parameters remains. The CKM matrix can thus be parameterized with four parameters (three real angles and one complex phase). These four parameters are free parameters of the SM.

The standard parameterization [5] of the matrix is given by:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -s_{23}c_{12} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{2.18}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$ where $i, j = 1, 2, 3$ label the quark generation and δ is the phase. All of the c_{ij} and s_{ij} can be chosen to be positive and δ may vary in the range $0 \leq \delta \leq 2\pi$.

One of the more common and illustrative parameterizations is the Wolfenstein parameterization [6], which takes into account the hierarchical structure of the sizes of CKM matrix elements:

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4). \tag{2.19}$$

It is an expansion in powers of $\lambda \equiv |V_{us}| = 0.2200 \pm 0.0026$ [7]. $A = 0.85 \pm 0.09$ and λ are known to high precision, while $\varrho(\equiv A\lambda^3\rho + O(\lambda^4))$, and η (Equation 2.21) are not well determined yet. If we define

$$s_{12} = \lambda; \quad s_{23} = A\lambda^2; \quad s_{13}e^{i\delta} = A\lambda^3(\varrho - i\eta), \tag{2.20}$$

it follows that

$$\varrho = \frac{s_{13}}{s_{12}s_{23}} \cos\delta, \quad \eta = \frac{s_{13}}{s_{12}s_{23}} \sin\delta. \tag{2.21}$$

We can write the CKM matrix parameterization that is correct to $O(\lambda^7)$ [8]:

$$\hat{V}_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} - \frac{1}{8}\lambda^4 & \lambda + O(\lambda^7) & A\lambda^3(\varrho - i\eta) \\ -\lambda + \frac{1}{2}A^2\lambda^5[1 - 2(\varrho + i\eta)] & 1 - \frac{\lambda^2}{2} - \frac{1}{8}\lambda^4(1 + 4A^2) & A\lambda^2 + O(\lambda^8) \\ A\lambda^3(1 - \bar{\varrho} - i\bar{\eta}) & -A\lambda^2 + \frac{1}{2}A\lambda^4[1 + 2(\varrho + i\eta)] & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix}, \tag{2.22}$$

where we have, by including the corrections of the order of λ^2, defined two parameters:

$$\bar{\varrho} = \varrho\left(1 - \frac{\lambda^2}{2}\right), \quad \bar{\eta} = \eta\left(1 - \frac{\lambda^2}{2}\right). \tag{2.23}$$

2.2.4 Unitarity Conditions of the CKM Matrix

The CKM matrix V_{CKM} is unitary by construction, $V_{CKM}V^{\dagger}_{CKM} = I$, which leads to the following relations amongst its elements:

$$\sum_i V_{ij}V^*_{ik} = \delta_{jk}. \tag{2.24}$$

Since the matrix elements of V_{CKM} are in general complex, the unitarity conditions for different rows ($j \neq k$) can be illustrated as triangles in the complex plane. The triangle formed from the unitarity relation imposed on the first and third columns has specail significance since it is one of the few such triangles with sides of roughly the same length ($O(\lambda^3)$). The relation is given by

$$V_{ud}V^*_{ub} + V_{cd}V^*_{cb} + V_{td}V^*_{tb} = 0, \tag{2.25}$$

and determines the so called *Unitarity Triangle*. For convenience, we normalize one of the sides by dividing the relation in Equation 2.25 with $|V_{cd}V^*_{cb}|$ and choose a phase convention so that $V_{cd}V^*_{cb}$ is real. The vertices along the normalized side are fixed at $(0, 0)$ and $(0, 1)$, while the remaining vertex has the coordinates $(\bar{\varrho}, \bar{\eta})$, and needs to be determined by experiments (See Figure 2.2).

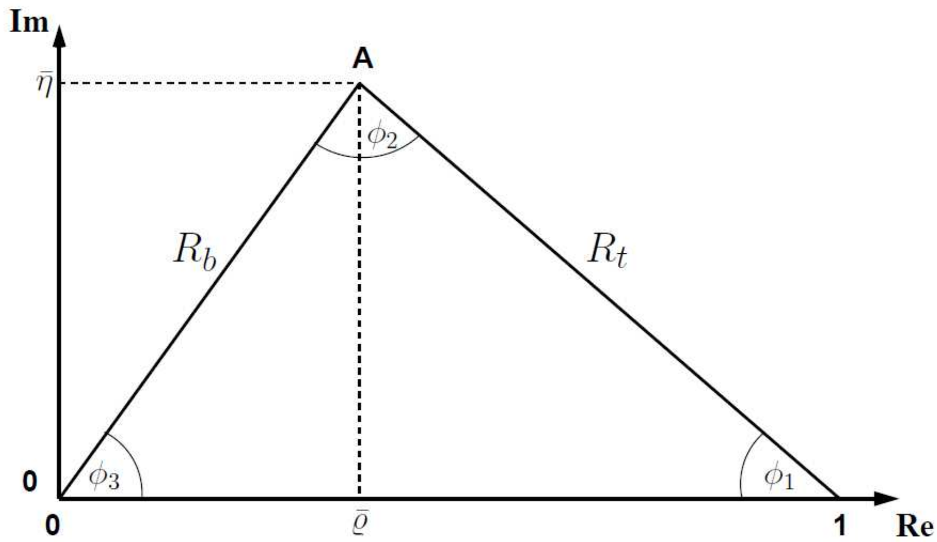


Figure 2.2: The rescaled Unitarity Triangle

The angles and side-lengths of the Unitarity Triangle are given by:

$$\phi_1 \equiv \left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right]; \phi_2 \equiv \left[-\frac{V_{ud}V_{ub}^*}{V_{td}V_{tb}^*} \right]; \phi_3 \equiv \left[-\frac{V_{cd}V_{cb}^*}{V_{ud}V_{ub}^*} \right] \equiv \pi - \phi_1 - \phi_2; \tag{2.26}$$

$$R_b \equiv \frac{|V_{ud}V_{ub}^*|}{|V_{cd}V_{cb}^*|} = \sqrt{\bar{\varrho}^2 + \bar{\eta}^2} = \left(1 - \frac{\lambda^2}{2}\right) \frac{1}{\lambda} \left| \frac{V_{ub}}{V_{cb}} \right|; \tag{2.27}$$

$$R_t \equiv \frac{|V_{td}V_{tb}^*|}{|V_{cd}V_{cb}^*|} = \sqrt{(1 - \bar{\varrho})^2 + \bar{\eta}^2} = \frac{1}{\lambda} \left| \frac{V_{td}}{V_{cb}} \right|. \tag{2.28}$$

2.2.5 Constraining the Unitarity Triangle

There are several processes that can determine sides or angles of the Unitarity Triangle. Since both $|V_{cb}|$ and $|V_{ub}|$ are present in Equation 2.25, decays of B mesons play an important role in these determinations. While the angle ϕ_1 can be measured from the time-dependent CP violation in decays like $B^0 \to J/\psi K_S$, the side R_b is determined by the ration of $|V_{ub}|/|V_{cb}|$, which can be measured in semileptonic B decays.

There is one important difference in the two measurements: the loops in box and penguin diagrams in the ϕ_1 determination make it sensitive to contributions from possible new particles that would appear in the loop, whereas the measurements of matrix elements $|V_{ub}|$ and $|V_{cb}|$ are determined from tree-amplitude diagrams, which are practically insensitive to such contributions. A comparison of such measurements can, when the theoretical and experimental precision allows, show a possible departure from unitarity, hinting at the existence of physics that is not predicted by the SM (See Figure 2.3).

In 2001 the prediction of a large CP violation in the decays of B mesons was confirmed by independent measurements [10, 11], which now set the allowed range of the angle ϕ_1 of the

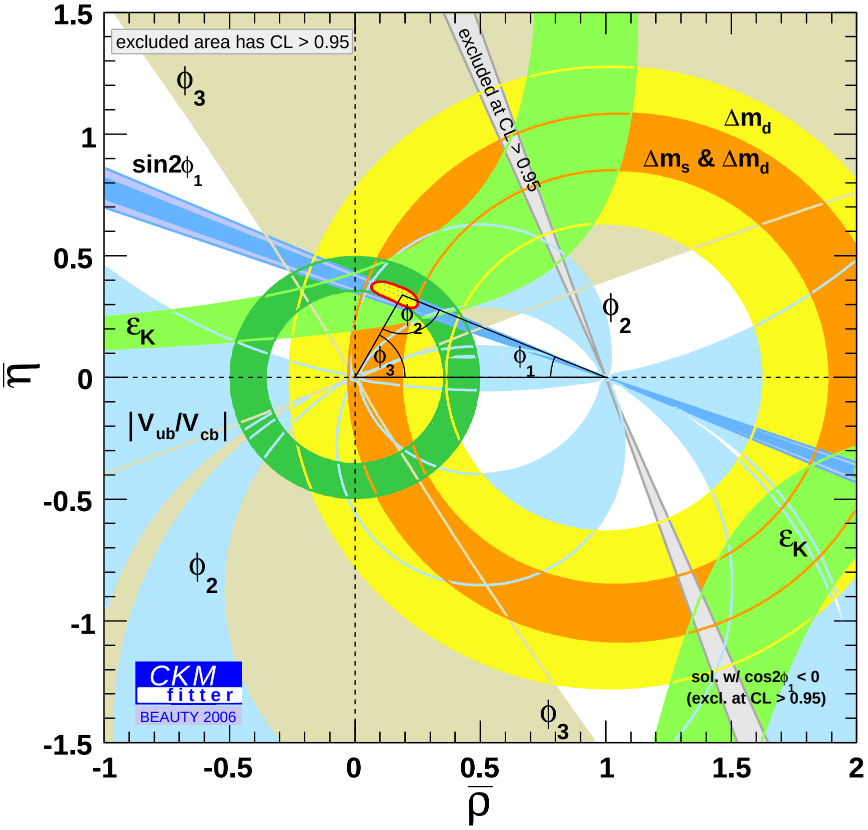


Figure 2.3: Schematic view of the Unitarity Triangle and determination of its upper vertex, by intersecting areas defined by measurements of different quantities. The constraints are obtained in measurements of the following processes: ϵ_K in CP violation of K mesons, Δm_d and Δm_s from $B\bar{B}$ and $B_s\bar{B}_s$ oscillations, respectively, $\sin(2\phi_1)$ in CP-violating B decays like $B \to J/\psi K_S$, and $|V_{ub}|/|V_{cb}|$ in the B meson semileptonic decays. Different measurements agree within current accuracy and their intervals intersect in a common area (shaded red). From Ref. [9].

Triangle (see Figure 2.3). The measurement of the angle ϕ_1 is done by observing the time-dependent asymmetries between the decays of B and $\bar{B}$ mesons to a common final state. The asymmetries arise due to the interference between the amplitudes for the direct decay and for a decay with the mixing of the B meson. The decay whose asymmetry is the most accuracy predicted by the theory $B^0 \to J/\psi K_S$. It is a decay with relatively large branching fraction in which only a single CKM phase appears in the leading decay amplitudes [10]. The average of all measurements for $\sin 2\phi_1$ is 0.686 ± 0.032 [12], with a total error of less than 5 %. It is therefore important to determine the side opposite to ϕ_1 by an accurate measurements of the ratio $|V_{ub}|/|V_{cb}|$.

2.2.6 Magnitude of V_{ub}

The determination of $|V_{ub}|$ from inclusive $B \to X_u \ell \nu$ decays suffers from large $B \to X_c \ell \nu$ backgrounds, where X_c and X_u stand for the hadronic part including c and u quack, respectively. In most regions of phase space where the charm background is kinematically forbidden the hadronic physics affects the detemination via unknown nonperturbative functions, so-called shape functions. At leading order in $\Gamma_{\rm QCD}/m_b$ there is only one shape function, which can be extracted from the photon energy spectrum in $B \to X_s \gamma$ [13, 14] and applied to several spectra in $B \to X_u \ell \nu$. The subleading shape functions are modeled in the current calculations. It is possible to choose phase space cuts in order that the rate does notdepend on the shape function [15].

The measurements of both hadronic and leptonic systems are important for an effective choice of phase space. A different approach is to extend the measurements deeper into the $B \to X_c \ell \nu$ region to reduce the theoretical uncertainties. Analyses of the electron-energy endpoint for CLEO [16], BaBar [17] and Belle [18] quote $B \to X_u e \bar{\nu}$ partial rates for $|\vec{p}_e| \geq 2.0\,{\rm GeV}/c$ and $1.9\,{\rm GeV}/c$, which are well below the charm endpoint. The large and pure $B\bar{B}$ samples at the B factories permit the selection of $B \to X_u \ell \nu$ decays in events where the other B is fully reconstructed [19]. With this full-reconstruction tag method the four-momenta of both the leptonic and hadronic systems can be measured. It also gives access to a wider kinematic region due to improved signal purity.

To extract $|V_{ub}|$ from an exclusive channel, the form factors have to be known. Experimentally, better signal-to-background ratios are offset by smaller yields. The $B \to \pi \ell \nu$ branching fraction is now known to 8 %. The first unquenched lattice QCD calculations of the $B \to \pi \ell \nu$ form factor for $q^2 > 16\,{\rm GeV}^2$ appeared recently [20, 21]. Light-cone QCD sum rules are applicable for $q^2 < 14\,{\rm GeV}^2$ [22] and yield somewhat smaller $|V_{ub}|$, $(3.3^{+0.6}_{-0.4}) \times 10^{-3}$. The theoretical uncertainties in extracting $|V_{ub}|$ from inclusive and exclusive decays are different. A combination of the determinations is quoted by the V_{cb} and V_{ub} mini-review as [23],

$$|V_{ub}| = (4.31 \pm 0.30) \times 10^{-3}, \tag{2.29}$$

which is dominated by the inclusive measurement. In the previous edition of the RPP [24] the average was reported as $|V_{ub}| = (3.67 \pm 0.47) \times 10^{-3}$, with an uncertainty around 13 %. The new average is 17 % larger, somewhat above the range favored by the measurement of $\sin 2\phi_1$ discussed below.

2.2.7 SM Prediction

We have already mentioned the leptonic branching fraction in the SM. The Equation 2.4 is put again

$$\mathcal{B}(B^+ \to \ell^+ \nu_\ell) = \frac{G_F^2 m_B m_\ell^2}{8\pi} \left(1 - \frac{m_\ell^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B.$$

One of the Cabibbo-Kobayashi-Maskawa matrix elements, V_{ub} have been also mentioned. They are summarized by the HPACD collaboration, then $|V_{ub}| = (4.39 \pm 0.33) \times 10^{-3}$ is determined from inclusive charmless semileptonic B decay data [12]. The decay constant of B mesons (f_B) is 0.216 ± 0.022 GeV [12]. The lifetime of B mesons (τ_B) is 1.643 ± 0.010 ps [12]. Then we can get the branching fractions for the decays $B^+ \to \mu^+ \nu_\mu$ and $B^+ \to e^+ \nu_e$ in the SM as

$$\mathcal{B}(B^+ \to \mu^+ \nu_\mu) = (4.7 \pm 0.7) \times 10^{-7} \quad (2.30)$$

$$\mathcal{B}(B^+ \to e^+ \nu_e) = (1.1 \pm 0.2) \times 10^{-11} \quad (2.31)$$

2.2.8 Recent Result of Other Leptonic Decays

The recent result of the leptonic decay of charged B meson $B^+ \to \tau^+ \nu_\tau$ have been published in Ref. [25] by the Belle collaboration. The result is the first evidence of the leptonic decay in B meson system and the first direct determination of the B decay constant. The branching fraction of $B^+ \to \tau^+ \nu_\tau$ is measured as

$$\mathcal{B}(B^+ \to \tau^+ \nu_\tau) = \left(1.79^{+0.56}_{-0.49}(\text{stat})^{+0.46}_{-0.51}(\text{syst})\right) \times 10^{-4}, \quad (2.32)$$

and the B decay constant is

$$f_B = 0.229^{+0.036}_{-0.031}(\text{stat})^{+0.034}_{-0.037}(\text{syst}) \text{ GeV}. \quad (2.33)$$

Other leptonic decays, $B^+ \to \mu^+ \nu_\mu$ and $B^+ \to e^+ \nu_e$ have not yet been observed. The most stringent current upper limits at 90 % confidence level for these modes are $\mathcal{B}(B^+ \to \mu^+ \nu_\mu) < 6.6 \times 10^{-6}$ [26] and $\mathcal{B}(B^+ \to e^+ \nu_e) < 1.5 \times 10^{-5}$ [27]. Preliminary limits of $\mathcal{B}(B^+ \to \mu^+ \nu_\mu) < 2.0 \times 10^{-6}$ [28] and $\mathcal{B}(B^+ \to e^+ \nu_e) < 7.9 \times 10^{-6}$ [29] are also available from the Belle and BaBar collaborations, respectively.

2.2.9 Theory Beyond the Standard Model

We have just mentioned the decays of B mesons in the SM. However if there are any particles which are expected in a theory beyond the SM, the branching fractions would be enhanced. In this section, we briefly introduce one example of theories beyond the SM.

In the Minimal Supersymmetric Standard Model (MSSM), the Higgs doublet($H^{\pm}$) which is predicted to have electric charge and light mass can be created from annihilation of $\bar{b}$ and u quarks and decay into a lepton and a neutrino [30]. Figure 2.4 shows the Feynman diagram of the leptonic decay of a B meson via a charged Higgs boson.

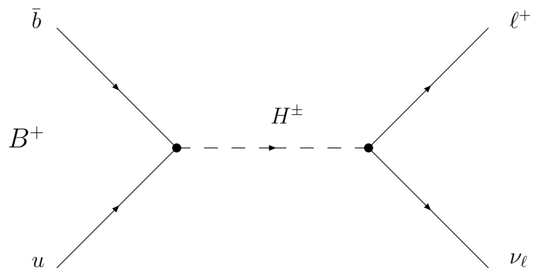


Figure 2.4: Feynman diagram for the leptonic decay of the charged B_d meson via Higgs doublet based on MSSM prediction.

In the MSSM, the Higgs doublet Yukawa coupling constants are controlled by the parameter $\tan\beta = v_2/v_1$, the ratio of the vacuum expectation values of the two doublets, normally expected to be of order m_t/m_b. Then the Equation 2.1 are expanded as

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} V_{ub}[u\gamma_u(1-\gamma^5)\bar{b}][\ell\gamma_u(1-\gamma^5)\nu] - R_\ell[u(1+\gamma^5)\bar{b}][\ell(1-\gamma^5)\nu], \tag{2.34}$$

where

$$R_\ell = \tan^2\beta\left(\frac{m_b m_\ell}{m_{H^+}^2}\right). \tag{2.35}$$

The standard V-A term induces $B^+ \to \ell^+\nu$ decays via axial-vector current, with Equation 2.2, while the pseudoscalar coupling of the $H^{\pm}$ boson is simply related

$$< 0|u\gamma^5\bar{b}|B^+ >= i f_B\left(\frac{m_B^2}{m_b}\right), \tag{2.36}$$

where we have ignored m_u compared to m_b. Equation 2.3 is also expanded and one easily arrives at the amplitude

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} V_{ub} f_B\left[m_\ell - R_\ell\left(\frac{m_B^2}{m_b}\right)\right][\ell\gamma_u(1-\gamma^5)\nu], \tag{2.37}$$

where the helicity suppression is in the SM term, while the charged Higgs term is proportional to R_ℓ. Then the branching fraction is given as

$$\mathcal{B}_{\rm MSSM}(B^+ \to \ell^+\nu_\ell) = \mathcal{B}_{\rm SM} r_H, \tag{2.38}$$

where $\mathcal{B}_{\rm SM}$ shows the Equation 2.4 and r_H is given as

$$r_H = \left[1 - \tan^2\beta\left(\frac{m_{B^+}^2}{m_{H^+}^2}\right)\right]. \tag{2.39}$$

The recent result of the analysis of $B^+ \to \tau^+\nu_\tau$ shows that the measured branching fraction is consistent with the SM prediction within errors. In the SM prediction, the branching fraction of τ decay in a charged B meson is extracted as $\mathcal{B}_{\mathrm{SM}}(B^+ \to \tau^+\nu_\tau) = (1.59 \pm 0.40) \times 10^{-4}$. The measured branching fraction is compared with SM prediction and the r_H is constrained as

$$r_H = \left[1 - \tan^2\beta\left(\frac{m_{B^+}^2}{m_{H^+}^2}\right)\right] = 1.13 \pm 0.51. \tag{2.40}$$

Chapter 3

The Belle Experiments

The Belle experiment is designed to perform precision quantitative studies of B mesons. It is conducted at the High Energy Accelerator Research Organization, known as KEK, which is located in Tsukuba, Japan, as a joint effort of more than 350 physicists from 54 institutes and 10 countries. Its main goal is a precise measurement of CP asymmetries of B meson decays. Studies of CP violation and rare B meson decays require a data sample of many millions of B mesons. They are produced in collisions of electrons and positrons at KEKB, a B factory with asymmetric energies of beams set at the center-of-mass energy best corresponding to the mass of the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ resonance is a vector meson $b\bar{b}$ state, which decays with the strong interaction to a $B\bar{B}$ meson pair. Since the energies of the beams are asymmetric, the particles are boosted int the direction of the more energetic beam. This boost enables the study of time-dependent CP violation by increasing the distances between decay vertices of the two B mesons. The Belle detector is situated at the interaction region of the e^+e^- beams, covering a large portion of the solid angle. Several detector sub-part systems enable reconstruction of tracks and identification of particles produced in the collision.

The KEKB accelerator commissioning began in December 1998, and six months after the Belle detector started its data-taking. Since then it managed to accumulate a data-sample of over 400 million decays of B meson pairs. This chapter briefly describes the experimental apparatus of KEKB and Belle.

3.1 The KEKB

The KEKB is a ring accelerator measuring 3 kilometers in circumference and colliding electrons and positrons at a center-of-mass energy of 10.58 GeV. Electrons with energy of 8.0 GeV and positrons with energy of 3.5 GeV are accelerated in bunches in the High Energy Ring (HER) and Low Energy Ring (LER), respectively. Two rings intersect at the Interaction Point (IP) where bunches of particles continuously collide. To reduce background synchrotron radiation, the beams collide at a finite crossing angle of 22 mrad. The IP is located in Tsukuba Experimental Hall - the site of the Belle detector, see Figure 3.1.

At the IP, electrons and positrons interact in processes like Bhabha scattering, tau and muon

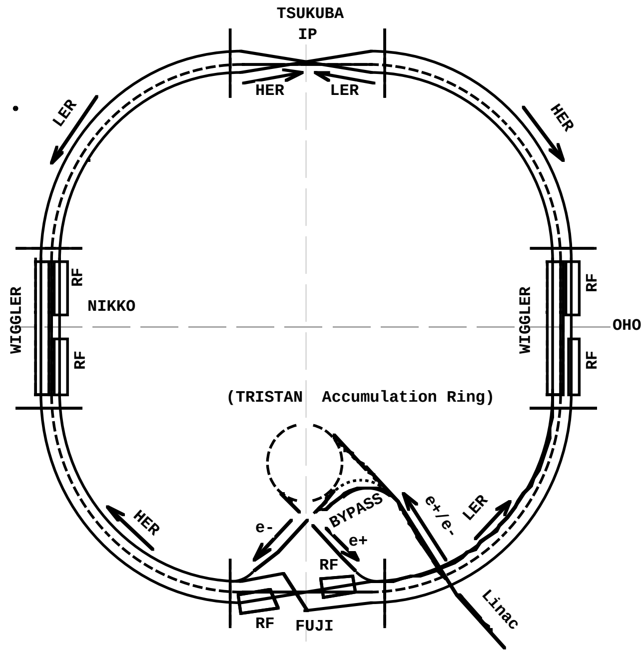


Figure 3.1: The KEKB Storage rings, LER and HER, with the IP located in Tsukuba Experimental Hall.

pair production, quark pair production and two-photon events. Even though the center-of-mass energy is tailored for the $\Upsilon(4S)$ resonance production as illustrated in Figure 3.2, only one $\Upsilon(4S)$ is produced in every e^+e^- interaction. The rate of the production, R is defined as the interaction cross section, σ, multiplied by the luminosity, $\mathcal{L}$, measured in units of cm^2 and $\mathrm{cm}^{-2}s^{-1}$, respectively:

$$R = \sigma\mathcal{L}. \tag{3.1}$$

The interaction cross section for the $\Upsilon(4S)$ production at the center-of-mass energy 10.58 GeV is

$$\sigma(e^+e^- \to \Upsilon(4S)) = 1.1\,\mathrm{nb}, \tag{3.2}$$

where the unit is barn, $b \equiv 10^{-24}\mathrm{cm}^2$. The luminosity is a measure of the beam-colliding performance, and is given by

$$\mathcal{L} = f \cdot n \cdot \frac{N_1 N_2}{A}, \tag{3.3}$$

where n bunches of N_1 and N_2 particles in opposing beams meet f times per second, and the overlapping area of the beams is A.

The estimated maximum luminosity to be achieved in the proposal [31] was $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{sec}^{-1}$, and has already been surpassed. A maximum luminosity of $1.7118 \times 10 - 34\mathrm{cm}^{-2}\,\mathrm{sec}^{-1}$ was achieved on November 15, 2006, and is currently the highest luminosity ever achieved by a collider. The measure of collected data is integrated luminosity ($\mathcal{L}_{int}$):

$$\mathcal{L}_{int} = \int \mathcal{L} dt. \tag{3.4}$$

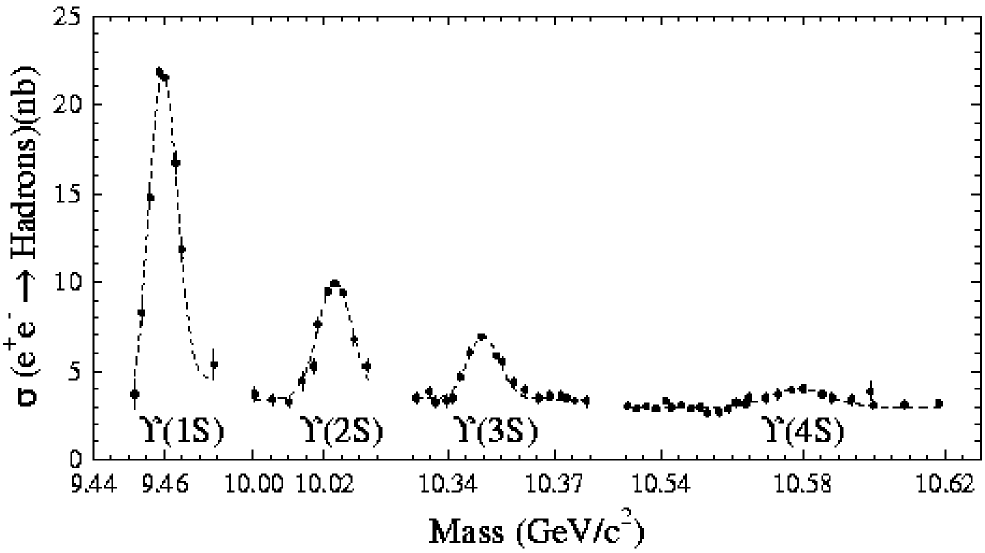


Figure 3.2: Cross section of $e^+e^- \rightarrow$ hadrons in a invariant mass range of $9.44 - 10.62\,\mathrm{GeV}/c^2$.

Taking the detector dead time into account, the Belle detector has accumulated the integrated luminosity of $\mathcal{L}_{int} = 710.254\mathrm{fb}^{-1}$ on December 25, 2006.

3.2 The Belle Detector

The Belle detector is a particle spectrometer configured within a 1.5 T superconducting solenoid and iron structure. It is located at the interaction region of the KEKB beams. Its structure reflects the beam energy asymmetry. It covers 97 % of the total solid angle and consists of seven sub-detectors; the silicon vertex detector (SVD), a central wire drift chamber (CDC), an aerogel Cherenkov counters (ACC), a time of flight counters (TOF), an electro-magnetic calorimeter (ECL), an extreme forward calorimeter (EFC), and a K_{L} and μ detector (KLM). The Belle detector scheme with detector sub-part systems is depicted in Figure 3.3.

The SVD measures B meson decay vertices and aids the CDC in providing charged particle tracking. Specific ionization energy loss measurements made with the CDC are combined with light yield readings from the ACC and time of flight information from the TOF to provide charged kaon and pion identification. Electromagnetic shower measurements and calorimetry, crucial for electron identification and photon detection, are performed by the ECL and the EFC. The KLM is used to identify muons and detect K_{L} mesons. The solenoid magnet provides a magnetic filed needed for measurement of momenta. The following subsections describe the Belle subdetectors. A detailed description of the Belle detector is given in Ref. [32].

3.2.1 Coordinate Systems

The origin of the coordinate system is defined as the position of the nominal IP. The common z axis is defined as the direction of the magnetic field within the solenoid, which coincides with

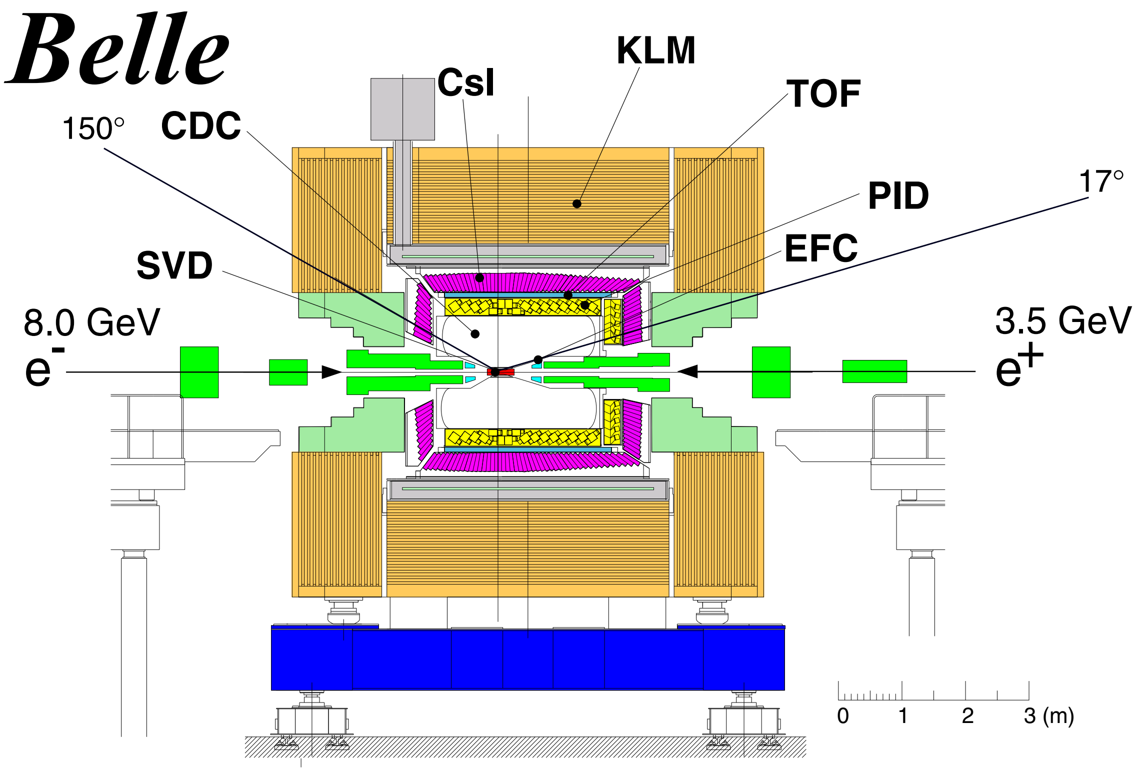


Figure 3.3: Side view of the Belle detector.

the direction of the electron beam. The x and y axes are horizontal and vertical, respectively, and correspond to a right-handed coordinate system. The polar angle is measured relative to the positive z axis. The azimuthal angle ϕ, laying in the $x-y$ plane, is measured relative to the positive x axis. The radius in the cylindrical coordinate system is defined as $r = \sqrt{x^2+y^2}$.

3.2.2 Beam Pipe

The beam pipe encloses the interaction point and maintains the accelerator vacuum. To precisely determine the decay vertices, the SVD should be as close to the IP as possible, but two effects force the SVD to be displaced from the IP: the beam-included heating of the beam pipe and large beam backgrounds due to the multiple Coulomb scattering in the beam pipe wall. These considerations are balanced by providing a double-wall beryllium beam pipe extending from $z = -4.6$ cm to $z = 10.1$ cm with an inner radius of $r = 20$ mm. Helium gas is cycled in the gap between the inner and outer walls to provide cooling and its low Z minimizes Coulomb interactions. The beam pipe is shown in Figure 3.4. In 2003, when a new SVD detector was installed, the existing beam pipe was replaced with a one with smaller dimensions, where the inner radius was reduced to $r = 15$ mm.

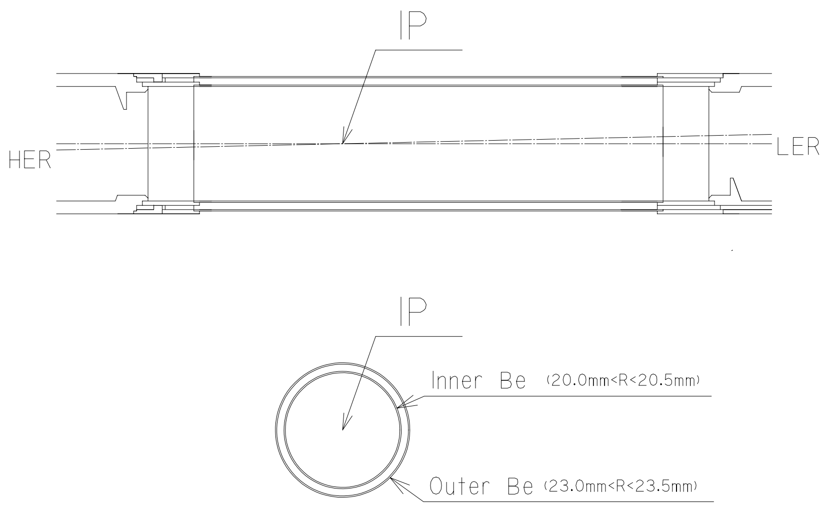


Figure 3.4: The cross section of the beryllium beam pipe at the interaction point.

3.2.3 Silicon Vertex Detector (SVD)

The measurement of the separation of two B meson decay vertices, which can be translated into a life-time difference between neutral B meson decays, is necessary for the measurement of time dependent CP violation in mixing. The $\Upsilon(4S)$ Lorentz boost in the laboratory frame allows measurement of the B meson decay vertices. The average flight distance of B mesons at the Belle detector is $200\,\mu$m, while SVD is able to resolve vertices to within a precision of $100\,\mu$m.

The SVD detects particles passing through a Double Sided Silicon Detector (DSSD), by observing the charge collected by sense-strips on both sides of the DSSD. At the Belle detector this occurrence is known as a SVD hit. The SVD uses S6936 type DSSDs, fabricated by Hamamatsu Photonics. The read-out is based on the VA1 chip, fabricated by Austrian Micro Systems.

The DSSD is essentially a semiconductor with a *pn* junction, operated under reverse bias to reach full depletion. A charged particle passing through the junction liberates electrons from the valence band into the conduction band, creating electron-hole (e^-h^+) pairs. The free e^-h^+ pairs instigate current in p^+ and n^+ strips situated along the surface of the bulk on opposing sides of the DSSD. The DSSD operation is depicted in Figure 3.5.

The p^+ strips are aligned along the beam axis to measure the azimuthal angle, ϕ, while the n^+ strips are aligned perpendicular to the beam axis to measure the z position. The pitch for different configurations can be read off Table 3.1.

The DSSD size is $57.5 \times 33.5 \times 0.3\,\mathrm{mm}^3$ and the DSSD consists of 1280 sense strips and 640 readout pads on each side. Every second sense strip is read out and the current is read using a hybrid card. Either one or two DSSDs connected to a hybrid form a short or long half ladder (HL), respectively. A full ladder consists of two half-ladders, connected together with the hybrids

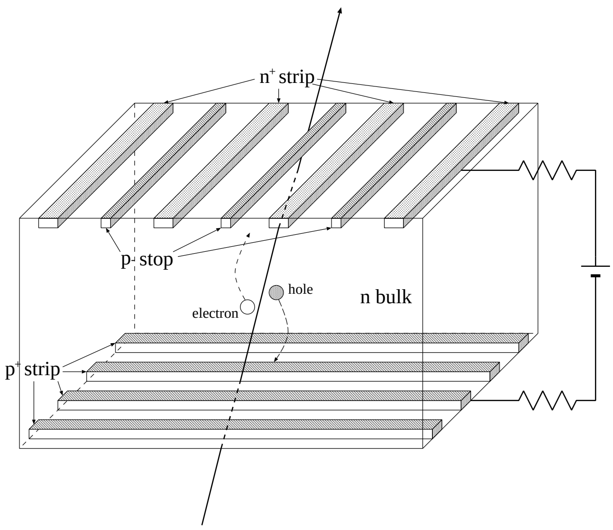


Figure 3.5: Schematic view of a Double Sided Silicon Detector.

	SVD1	SVD2
Beam-pipe radius (mm)	20	15
Number of layers	3	4
Number of DSSD ladders in layers 1/2/3/4	8/10/14/N.A.	6/12/18/18
Number of DSSDs in ladder in layers 1/2/3/4	2/3/4/N.A.	2/3/5/6
Radii of layers 1/2/3/4 (mm)	30.0/45.5/60.5/N.A.	
Angular coverage	$23^\circ \leq \theta \leq 140^\circ$	$23^\circ \leq \theta \leq 140^\circ$
Angular acceptance	0.86	0.86
Total number of channels	81920	110592
Strip pitch(μm) for z	84	75(73 for layer 4)

Table 3.1: Characteristics of the SVD1 and the SVD2

at the ends. Full ladders are arranged in cylindrical layers.

Two SVD configurations were used in the period of the data taking, the SVD1 (1998-2003) and the SVD2 (2003 to-date). Since the SVD detector has to be very close to the beam pipe, it has to endure intensive particle irradiation. The SVD2 has a greatly increased radiation tolerance, and by adding another layer of ladders the spatial resolution was improved as well as the solid angle coverage of the detector. The characteristics of the two configurations are summarized in Table 3.1. Further detail on the SVD can be found in Ref [33].

3.2.4 Central Drift Chamber (CDC)

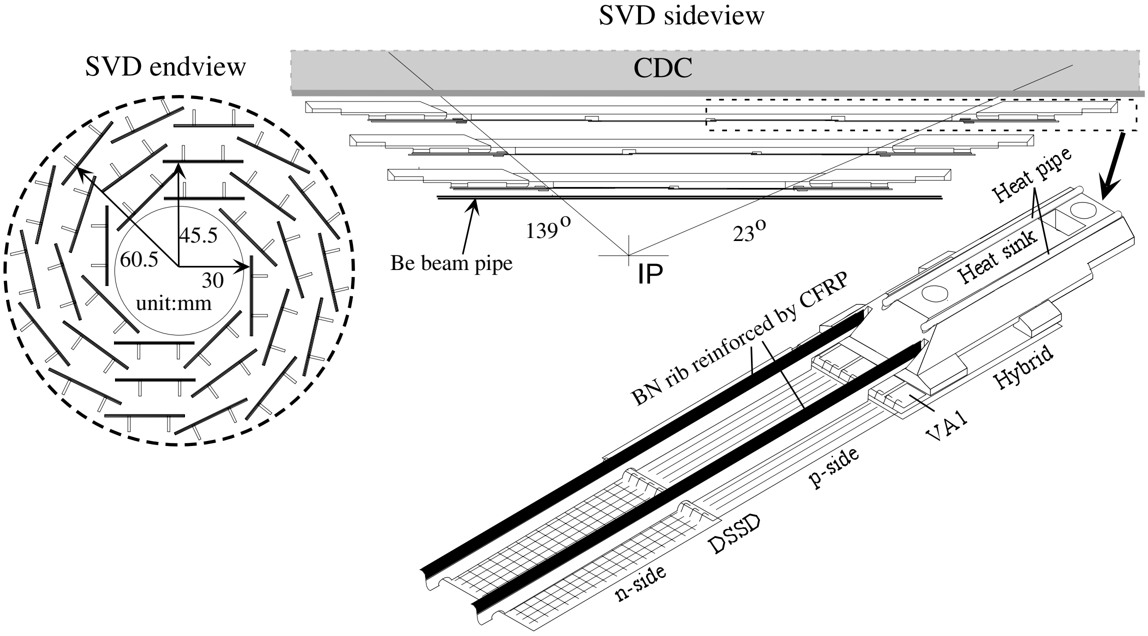
The Central Drift Chamber (CDC) is designed to reconstruct trajectories of charged particles by detecting the ionization of the gas from the passing particles. Particle specific ionization energy loss, dE/dx, is also measured for particle identification purposes. Information on the hits in the CDC is used in triggering.

Figure 3.7 shows the CDC detector views from the side and from the beam axis. The CDC encloses the SVD, extending radially from 77 mm to 880 mm. The CDC consists of 32 axial layers, 18 small angle stereo layers, and 3 cathode strip layers. Axial layers measure the $r - \phi$ position, while stereo layers in conjunction with axial layers, included at a small angle to the beam pipe, measure the z position. The CDC covers a polar angle region of $17^\circ \leq \theta \leq 150^\circ$. The spatial resolution is $10\,\mu$m in $r - \phi$, and is better than 2 mm in the z direction. The CDC contains a total of 8400 drift cells. A drift cell is the functional unit of the CDC, consisting of a positively biased sense wire, surrounded by six negatively biased field wires, strung along the beam direction. When the SVD2 was installed the inner layers of the CDC were removed to accommodate for a larger SVD.

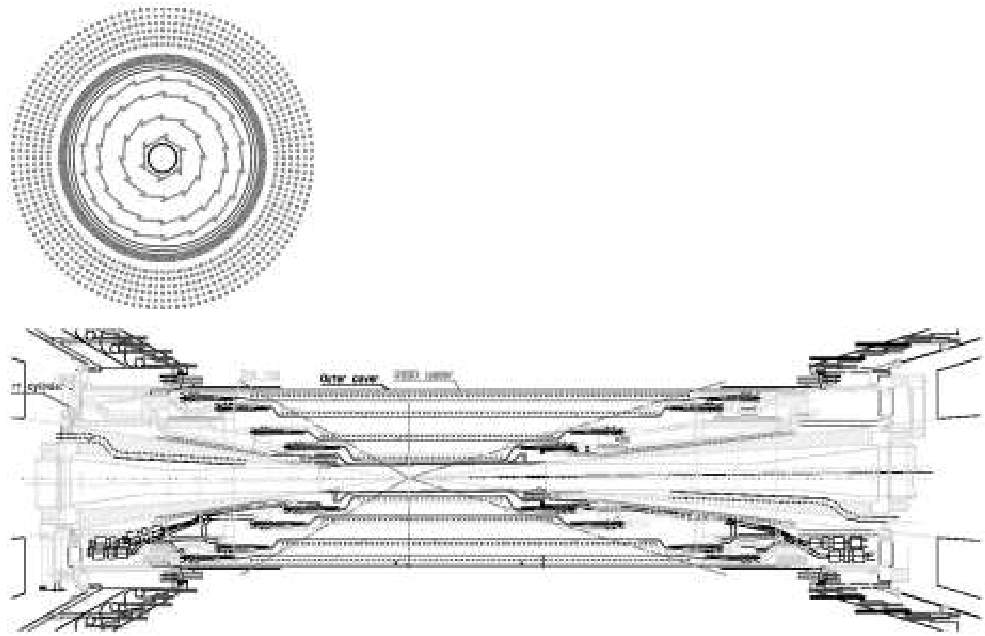
The cells are immersed in a helium-ethane gas mixture of ratio 1 : 1. The helium-ethane gas mixture has relatively long radiation length of 640 m to minimize the effect of multiple Coulomb scatterings on the momentum resolution. The ethane component increases the electron density, which improves the resolution of the ionization-energy-loss measurement.

A charged particle, traversing the cell, ionizes the gas along its path. The ionized electrons and positive ions are attracted to the anode and cathode sense wires, respectively. Their drift in high electromagnetic field near the wire instigates further ionization, resulting in avalanches of electrons and positive ions. When the avalanches reach the sense wire, a current is induced. If the signal by the current is higher than the threshold, a CDC hit is detected. The distance between the ionizing track and the sense wire is estimated from the time taken for the ionization column to form.

Track parameters are determined using a track segment finder, which sorts hits into tracks. A helix, which describes the path of a charged particle in a constant magnetic field, is fitted to the track. The obtained helix parameters are combined with the magnetic field strength to determine the charged particle momentum. The momentum resolution transverse to the beam axis measured



SVD1 (SVD1 end-view and SVD1 side-view)



SVD2 (SVD2 end-view and SVD2 side-view)

Figure 3.6: Silicon vertex detector configurations for the SVD1 and the SVD2.

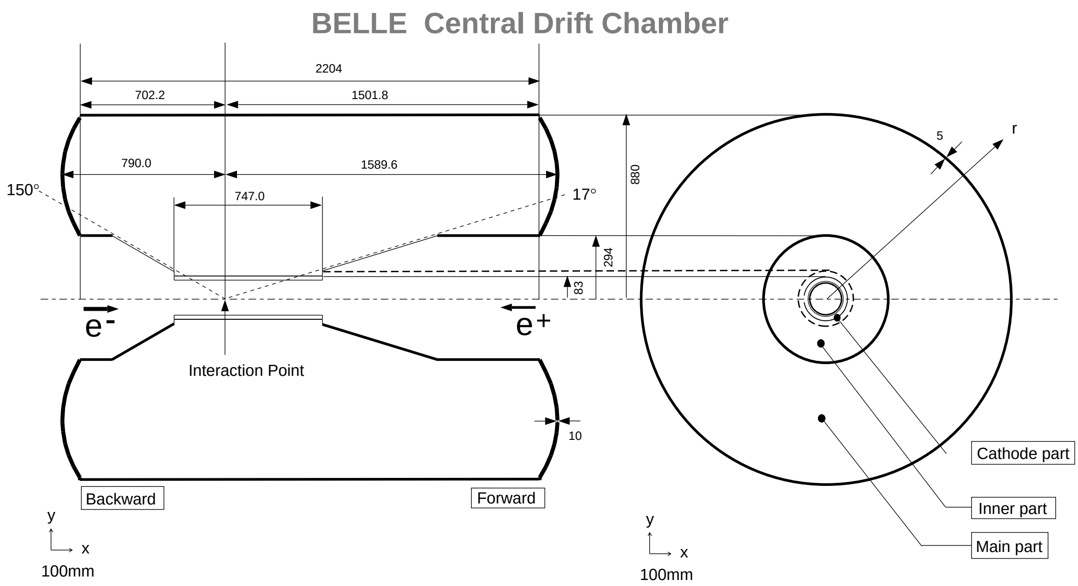


Figure 3.7: The Central Drift Chamber (CDC): side-view (left) and end-view (right).

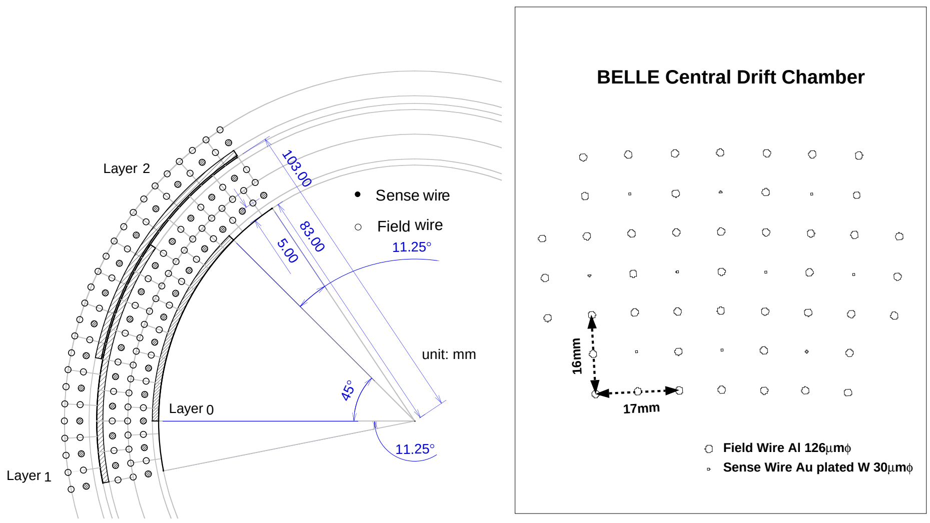


Figure 3.8: Cross-sectional view of cell structure of the CDC. Cathode strips are also drawn by hatches in the left figure.

from cosmic ray data, is

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = \sqrt{(0.20 p_\mathrm{T})^2 + (0.29/\beta)^2}\,\%, \tag{3.5}$$

where p_T is in units of GeV$/c$ and β is the particle velocity divided by the speed of light. Figure 3.9 shows the resolution of the transverse momentum.

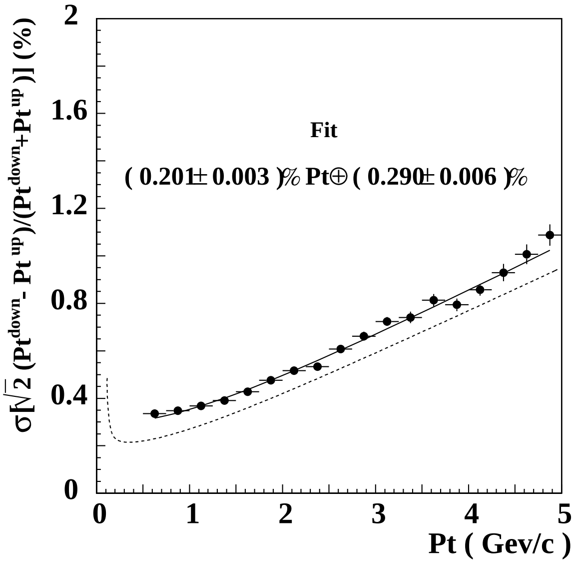


Figure 3.9: Resolution of the transverse momentum of the CDC.

Particle energy loss in a drift cell due to ionization, dE/dx, is determined from a hit amplitude recorded on a sense wire. Since the energy loss depends on a particle velocity at a given momentum, dE/dx distributions differs for different particle masses, as shown in Figure 3.10. The ionization energy loss is measured for each CDC hit and measurements along the trajectory are combined to calculate the truncated mean, $\langle dE/dx \rangle$, of the track.

The $\langle dE/dx \rangle$ resolution, measured on a sample of pions from K_S decays, is 7.8 %. The CDC can be used to distinguish pions from kaons of momenta up to 0.8 GeV$/c$ with a 3σ separation. A detailed description of CDC is presented in Ref [34].

3.2.5 Aerogel Cherenkov Counter (ACC)

The silica Aerogel Cherenkov Counter (ACC) plays a crucial role in discriminating charged pions from kaons. When a particle travels faster than the speed of light in that medium, it will emit Cherenkov light. The emitted light appears in the form of a coherent wavefront at a fixed angle with respect to the trajectory.

To emit Cherenkov light, the particle velocity has to be grater than the threshold value:

$$\beta > \beta_\mathrm{threshold} = \frac{1}{n} \tag{3.6}$$

where n is the refractive index of the medium. Threshold Cherenkov counters exploit the fact that only particles with velocity above $\beta_\mathrm{threshold}$ can emit Cherenkov photons. Since the momentum

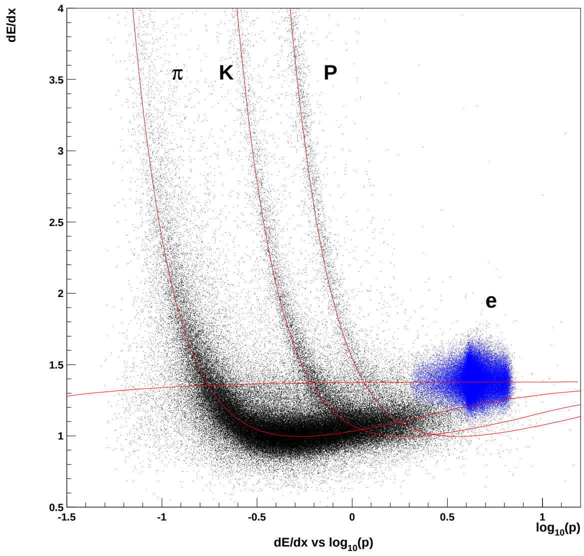


Figure 3.10: Truncated mean of dE/dx versus momentum The points are measurements taken during accelerator operations, and the lines are the expected distributions of each particle type. p is measured in GeV/c.

is measured by the other sub-detectors, the particles can be identified be observing whether Cherenkov photons have been detected or not.

K/π discrimination can be achieved by selecting media with appropriate refractive indices to cover typical momenta. The ACC augments the other detector subsystems by performing excellent K/π separation for momenta between 2.5 and 3.5 GeV/c, and is also able to provide useful information for momenta as low as 1.5 GeV/c and as high as 4.0 GeV/c.

The ACC is divided into barrel and forward end-cap regions and is shown in Figure 3.11. It spans a polar angle region of $17^\circ \leq \theta \leq 127^\circ$. The barrel ACC contains 960 counter modules, segmented into 60 cells in the ϕ direction. The forward end-cap ACC contains 228 counter modules, arranged into 5 concentric layers. Depending on the polar angle, the refractive index of the aerogel tiles ranges from $n = 1.01$ to 1.03. One of concerns when using aerogel is that its transparency is greatly reduced with age due to its hydrophilic property. A special aerogel production procedure has been developed so that it is possible to produce hydrophobic aerogel.

Cherenkov photons are detected by either one or two fine mesh-type photo-multiplier tubes (FM-PMT), which are attached to each counter. The ACC detector is positioned within a strong magnetic field, which drastically reduces the gain and the collection efficiency of photoelectrons. By using a fine FM-PMT with 19 dynodes, a high gain of 10^8 is maintained even in the strong magnetic field. Three different sizes of FM-PMT with radii of 1, 1.25 and 1.5 inches are used. The choice is dependent on the refractive index to keep the constant photon yield, since a lower refractive index results in a lower yield, barrel and end-cap modules are depicted in Figure 3.12.

The pulse height for each FM-PMT has been calibrated using μ-pair events. The average

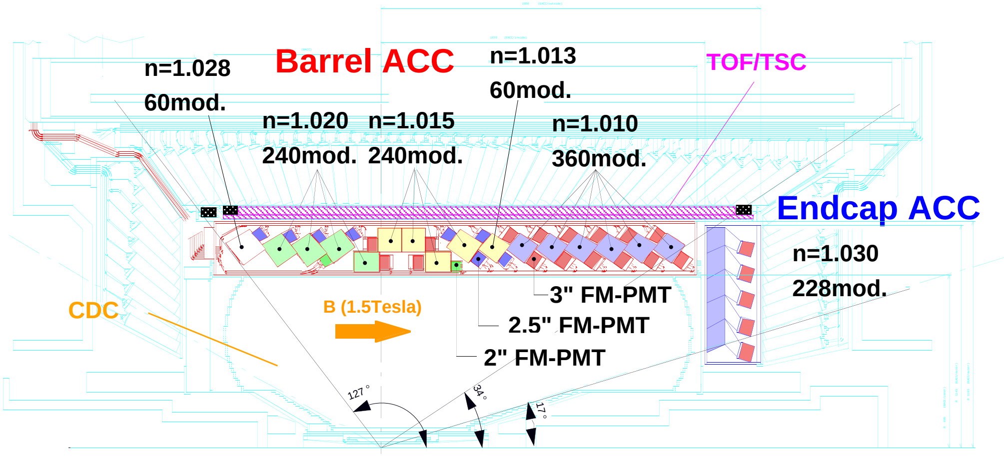


Figure 3.11: The configuration of the Aerogel Cherenkov Counter.

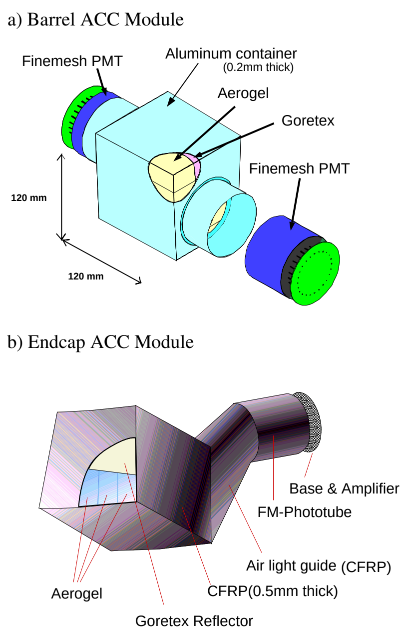


Figure 3.12: Schematic drawing of a typical ACC counter module : (a) barrel and (b) end-cap ACC.

number of detected photoelectrons, $\langle N_{pe} \rangle$, ranges from 10 to 20 for the barrel ACC and from 25 to 30 for the end-cap ACC.

Since pions are the most ubiquitous particles in hadronic events, the ACC performance is measured by its ability to identify kaons amongst pions. The ACC can provide good k/π separation with a kaon efficiency above 80 % and a pion-to-kaon fake rate below 10 %, as demonstrated in Figure 3.13. A more detailed description of ACC is presented in Ref [35, 36].

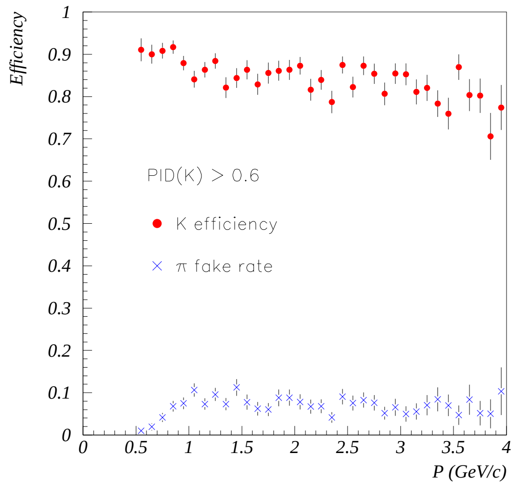


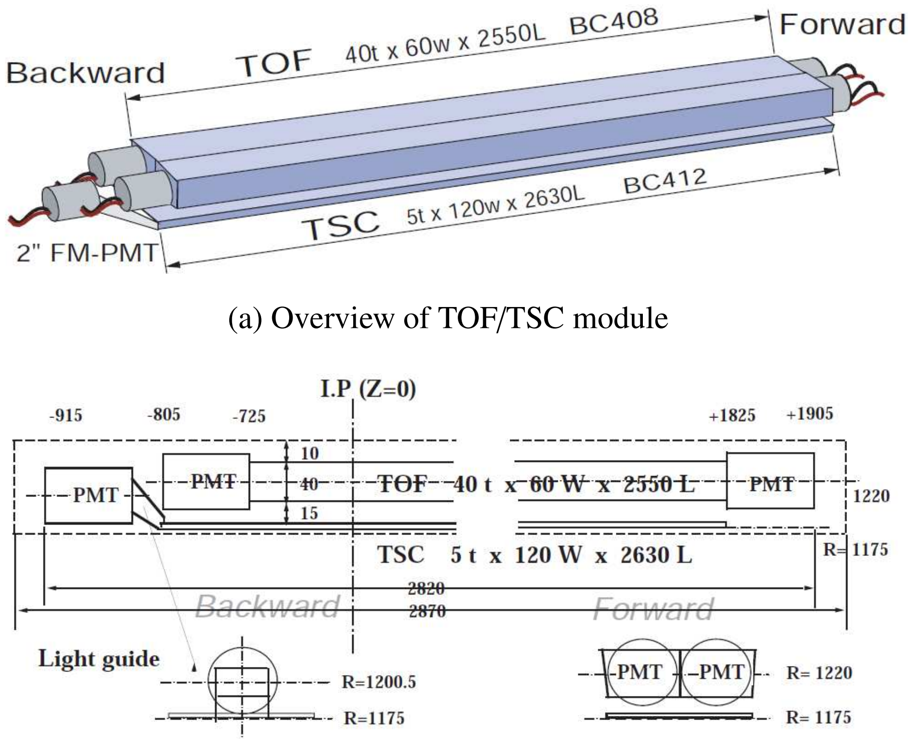
Figure 3.13: Kaon efficiency and pion fake rate, measured with $D^{*+} \to D^0(\to K\pi) + \pi^+$ decays, for the barrel region of the ACC.

3.2.6 Time of Flight Counter (TOF)

The Time of Flight counter (TOF) is used for identification of charged particles in an intermediated momentum range of 0.8 GeV.c to 1.2GeV$/c$. It measures the velocity of particles from the time of flight over the distance, which in turn is determined by the track helix parameters, measured in the CDC. A particle type is determined by combining its momentum (measured by CDC) and velocity obtained by TOF.

The TOF is comprised of long plastic scintillators which are chemical compounds emitting short light pulses after excitation by the passage with charged particles or by photons of high energy. The TOF measures the time of flight of a particle originating at the IP and passing through the scintillator, by detecting the emitted light pulses.

The TOF system consists of 64 modules, concentrically arranged around the z axis at a radius of 1.2 m. A module is made up of two trapezoid-ally shaped time-of-flight counters and one Trigger Scintillation Counter (TSC), separated by a radial gap of 1.5 cm, as shown in Figure 3.14. The thin TSC modules help to reject photon conversion backgrounds by taking a coincidence between TOF and TSC counters for triggering purposes. Scintillation light from a counter is collected by a FM-PMT. The FM-PMT was chosen due to excellent gain in the magnetic field. Two FM-PMTs are used for a TOF counter while one is used for a TSC counter. The FM-PMTs are mounted directly on the scintillator to eliminate the need for light guides.



(a) Overview of TOF/TSC module

(b) Dimensions of TOF/TSC module

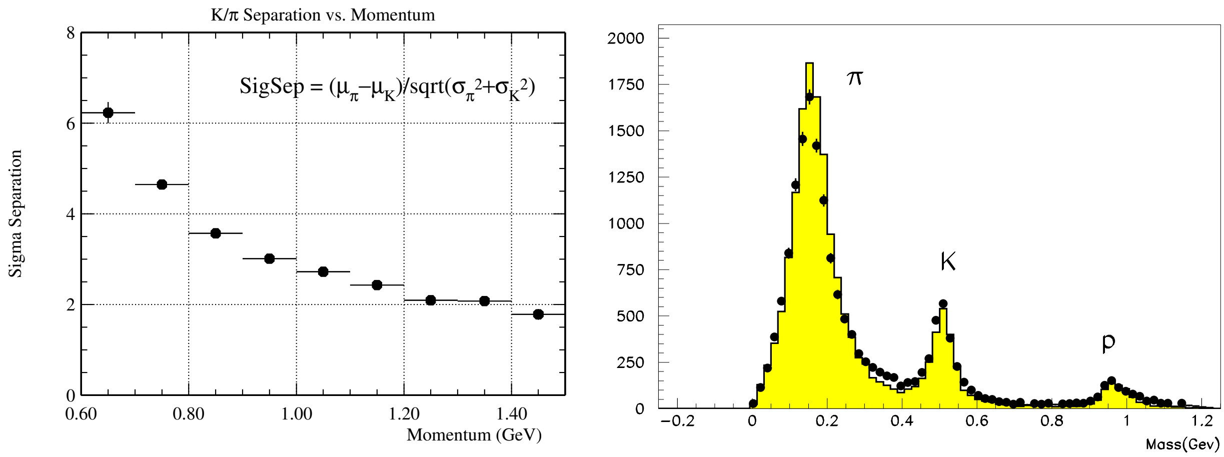
Figure 3.14: Structure of the TOF/TSC module. The numbers in the figure are in units of mm.

Time intervals are measured within a precision of 100 ps. The K/π separation is plotted as a function of momentum in Figure 3.15 (a), it shows that for momenta below 1.0 GeV/c a separation of more than 3σ is achieved. The mass distribution, shown in Figure 3.15 (b), measured from hadronic events, shows a comparison of real data with Monte Carlo simulation for a timing resolution of 100 ps. Clear peaks are evident for pions, kaons and protons. The detailed description of the TOF detector can be found in Ref [37].

3.2.7 Electromagnetic Calorimeter (ECL)

The Electromagnetic Calorimeter (ECL) is designed to measure energies and directions of photons and electrons produced in the Belle detector and it is crucial for electron identification. Fine-grained segmentation of the detector is needed for $\pi^0 \to \gamma\gamma$ reconstruction, since two nearby photons with their opening angle have to be detected.

High energy electrons and photons that enter the calorimeter material produce an electromagnetic shower by interactions with matter, mainly bremsstrahlung and electron-positron pair



(a) (b)

Figure 3.15: Performance of TOF. (a) K/π separation performance of the TOF as a function of momentum (μ_π and μ_K are pion and kaon hypothesis probabilities, respectively). (b) Mass distribution from TOF measurements for particle momenta below 1.2 GeV/c. The CDC momentum measurement and the velocity measurement by TOF are used.

production. A lateral shower shape ensues from Coulomb scattering of particles in the shower. Eventually, all of the incident energy appears as ionization or excitation of the absorbing material.

The ECL consists of a highly segmented array of 8, 736 cesium iodide crystals, doped with thallium (CsI(T1)). The thallium shifts the excitation light into the visible spectrum. The light is detected by a pair of PIN photon diodes placed at the rear of each crystal.

The crystals are arranged into three sections: the backward end-cap, the barrel, and the forward end-cap. The ECL barrel positioned at an inner radius of 1.25 m, is 3.0 m long, and spans the polar angle region of $32\,° \leq \theta \leq 128.7°$. The annular-shaped forward end-cap ECL is situated at $z = +2.0$m, and spans polar angle region of $12.0° \leq \theta \leq 31.4°$. The likewise annular-shaped end-cap ECL is situated at $z = -1.0$ m, and spans polar angle region of $7° \leq \theta \leq 155.7°$. The ECL configuration is shown in Figure 3.16.

A crystal is typically 30 cm long, equivalent to 16.2 radiation lengths (X_0) for electrons and photons, and is chosen to minimize energy resolution deterioration at high energies due to the fluctuation of shower leakages at the back of the crystal. The crystals are designed such that a photon entering a particular crystal at its center will deposit at least 80 % of its energy in that crystal.

A typical crystal in the barrel ECL has a forward and backward faces measuring 55 mm × 55 mm and 65 mm × 65 mm respectively. In the forward and backward end-caps ECL the profiles vary from 44.5 mm to 70.8 mm and from 54 mm to 82 mm respectively. Each crystal has a tower

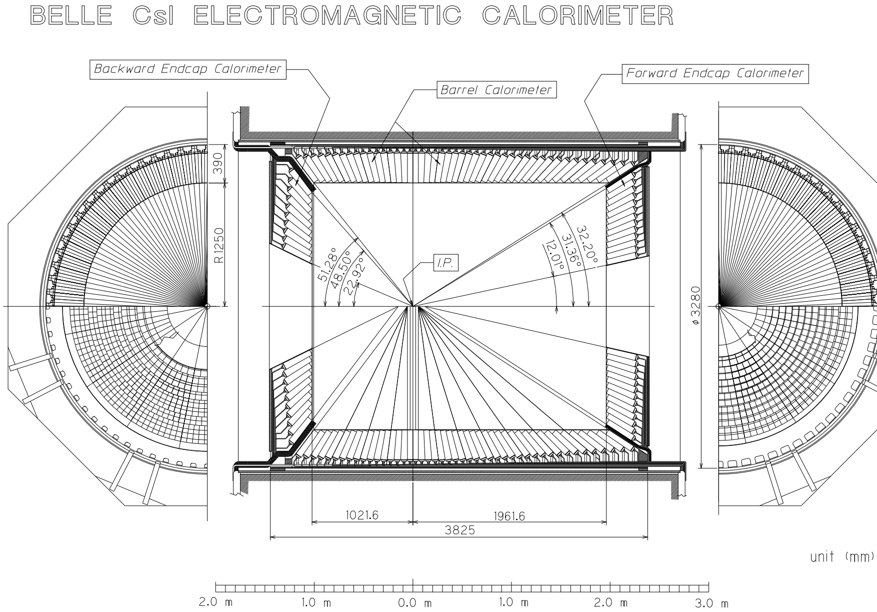


Figure 3.16: The configuration of Electromagnetic Calorimeter, with annular-shaped forward and backward end-caps.

like structure. In the barrel ECL they are tiled at an angle of approximately 1.3° in the θ and ϕ directions to prevent particles escaping through gaps between crystals.

Each crystal is wrapped in a diffuse reflector, a 200 μm thick sheet of Goretex teflon, to enable to light-collection by two photo diodes at the rear side (see Figure 3.17).

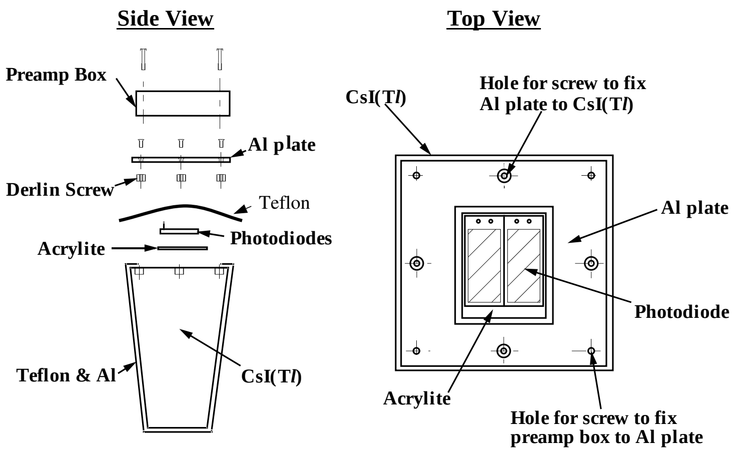


Figure 3.17: Mechanical assembly of the ECL detector.

The ECL is able to measure energies in the range of $0.02 < E_\gamma < 5.40$ GeV. It provides a measured energy resolution of

$$\left(\frac{\sigma_E}{E}\right) = \sqrt{1.34^2 + \left(\frac{0.066}{E}\right)^2 + \left(\frac{0.81}{E^{1/4}}\right)^2}, \tag{3.7}$$

and position resolution of

$$\sigma_{\text{pos}} = \frac{0.5\,\text{cm}}{\sqrt{E}}, \tag{3.8}$$

where E is measured in GeV.

Since pions deposit much less of their energy in the crystal, the difference of the energy deposit in the ECL can be used to distinguish charged pions from electrons, as illustrated in Figure 3.18. The plot also shows the difference between the response of negatively and positively charged pions that is a direct result of their different nuclear cross sections. The peak on the left is from minimum ionizing particles, which does not interact strongly with the material of ECL. Less than 1 % of pions are mis-identified as electrons with momenta above 2 GeV/c.

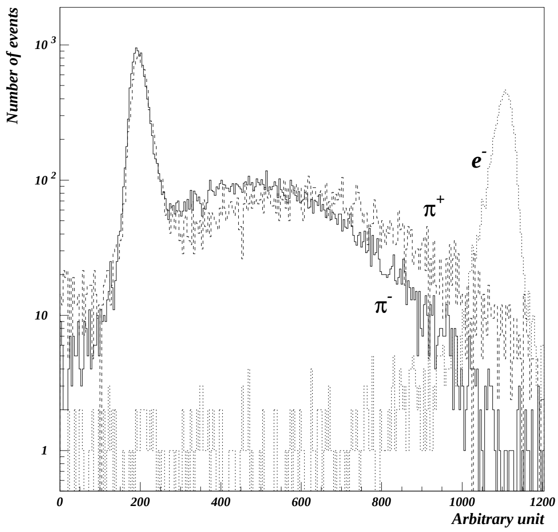


Figure 3.18: Distribution of the energy deposit by electrons (dotted line), by positive pions (dashed line) and by negative pions (solid line) at 1 GeV/c.

3.2.8 K_L/μ Detector (KLM)

The K_L and μ detector (KLM) is the outermost sub-detector system and is designed to identify K_L mesons and muons with high efficiency for momenta greater than 600 MeV/c. A K_L from the IP will typically traverse one interaction length (mean free path before an inelastic interaction) before reaching the KLM, most of which (0.8) is due to the ECL. Another 3.9 interaction lengths are provided by iron plates in the KLM, to produce a shower of ionizing particles when a K_L interacts with matter. The shower location is then measured to provide K_L flight direction, but the fluctuations in the shower size prevent any useful measurement of K_L energy.

Muons of sufficient energy (> 500 MeV) will penetrate the KLM easily, since they do not feel the strong interaction and the Bremsstrahlung radiation loss is much smaller than for electrons. A track penetrating several layers of the KLM is most likely a muon. Since muons suffer smaller

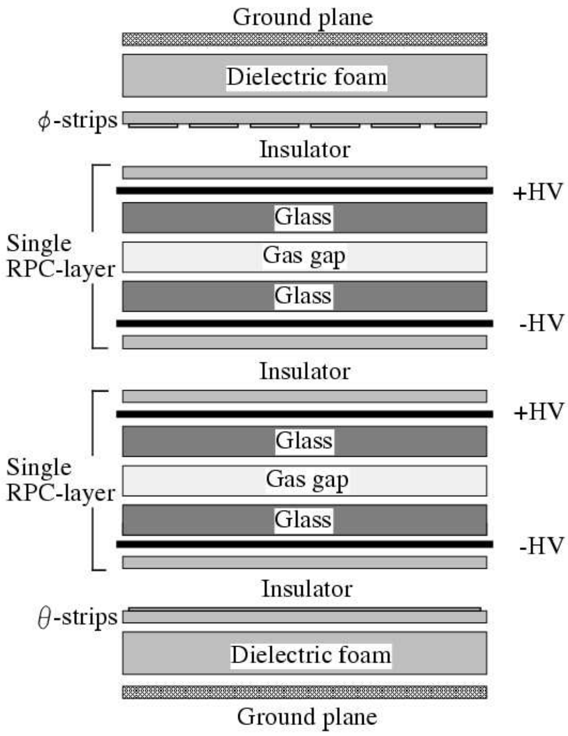


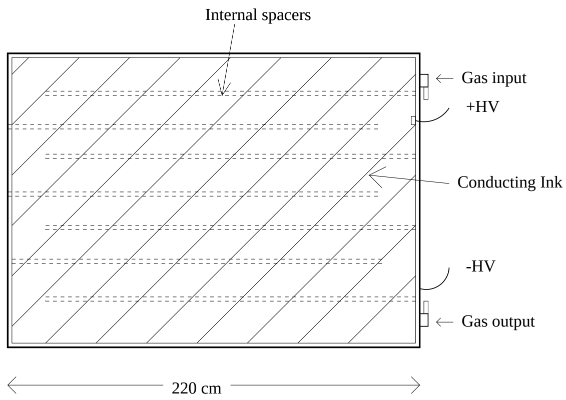
Figure 3.19: Cross section of a KLM super-layer, consisting of two Resistive Plate Counters layers. Ionizing particles instigate a discharge of HV plates, which induces signals in the pickup strips.

deflections in material, they can be distinguished from charged pions and kaons. The separation further improves for higher momenta.

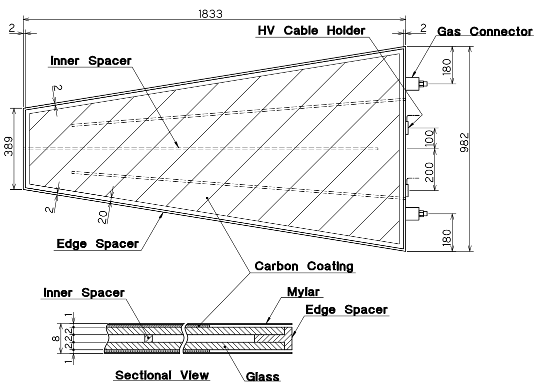
The KLM consists of alternating layers of charged-particle detectors and 4.7 cm thick iron plates. The KLM in the barrel region is octagonally shaped and is made of 15 detector layers and 14 iron layers. The KLM in the forward and backward end-caps contain 14 detector layers each.

A detector layer is a super-layer of two glass-electrode Resistive Plate Counters (RPC), sandwiched between two orthogonal pickup strips. The RPC module consists of two high-voltage plates, insulated by high-bulk-resistivity glass plates from a gas-filled gap, as shown in Figure 3.19. An ionizing particle, traversing the gas-filled gap of the RPC, initiates a streamer in the gas that results in a local discharge of the plates. The discharge induces signals on the external orthogonal pickup strips.

The pickup strips, typically 5 cm wide, provide $\phi - z$ and $\theta - \phi$ information in the barrel and the end-cap regions, respectively. The size of the strips matches the uncertainty due to the multiple scattering of particles as they travel through iron, and limits the spatial resolution to a few centimeters. The barrel and end-caps KLMs contain 240 and 122 RPC modules. The polar angular coverage is $20° < \theta < 155°$. Figure 3.20(a) and 3.20(b) show schematic view of the barrel and end-cap RPCs respectively. The K_L angular resolution measured from the IP is better than 10 mrad. For momenta above 1.5 GeV$/c$ the muon identification efficiency is greater than 90 % with a miss-identification rate of less than 5 %. A more detailed description of KLM detector can



(a) Barrel RPC.



(b) End-Cap RPC.

Figure 3.20: A schematic diagram of Resistive Plate Counters in the KLM detector.

be found in Ref [38].

3.2.9 Solenoid Magnet

A super-conducting solenoid magnet provides a magnetic field of 1.5 T in a cylindrical volume of 3.4m in diameter and 4.4 m in length. The solenoid encases all the sub-detectors except the KLM. The external iron structure of the Belle detector serves as the return path of magnetic flux and as absorber material for the KLM. The solenoid details are shown in Table 3.2. The magnetic field mapping, measured with accelerator final-focus quadrupole magnets located within the solenoid, QCS-R and QCS-L, is shown in Figure 3.21.

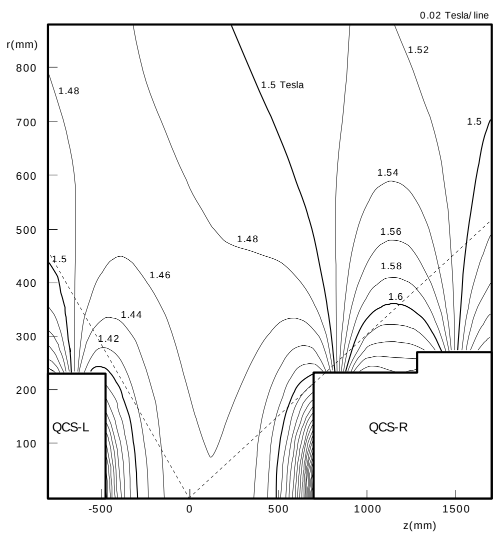


Figure 3.21: Contour plot of the measured magnetic field strength in the Belle detector.

General	Central field	1.5 T
	Length	4.41 m
	Weight	23 t
	Cool-down time	≤ 6 days
	Quench-recovery time	≤ 1 day
Cryostat	Inner/Outer Radius	1.70/2.00 m
Coil	Effective radius	1.8 m
	Length	3.92 m
	Superconductor	NbTi/Cu
	Nominal current	4400 A
	Inductance	3.6 H
	Stored energy	35 MJ
	Typical charging time	0.5 h

Table 3.2: Main parameters of the solenoid magnet.

3.2.10 Extreme Forward Calorimeter (EFC)

The Extreme Forward Calorimeter (EFC) offers electron and photon calorimetry at the extreme forward and backward regions, defined as $6.4^\circ < \theta < 11.5^\circ$ and $163.3 < \theta < 171.2^\circ$, respectively. The EFC is placed on the cryostat front faces of the KEKB accelerator compensation solenoid magnet, which is surrounding the beam pipe.

The EFC is constructed from crystals of Bismuth Germanate (BGO), which was chosen for its ability to withstand radiation doses at mega-rad level, while still providing good energy resolution. The detector is segmented into 32 azimuthal and 5 polar sections for both backward and forward cones. Each crystal is tower-shaped and is aligned to point towards the IP. The arrangement is illustrated in Figure 3.22. Since the BGO crystals are resistive to radiation, the EFC shields the CDC from beam related backgrounds and synchrotron radiation. The EFC is also used as a beam monitor and luminosity meter for KEKB accelerator control. A more detailed description of EFC detector can be found in Ref [32].

3.3 Trigger and Data Acquisition System

3.3.1 Trigger

In the environment of a beam crossing rate of 509 MHz, various processes occur at higher rate than we are able to store in our data acquisition system. Many of the processes are due to the interactions of beams with their residual gas or beam pope material and not of interest for the B physics measurements. Therefore a complex triggering system has to be adopted to focus on the events of interest. Physics of interest includes hadron production, Bhabha scattering, μ-pair and τ-pair production and two-photon events. Two-photon and Bhabha scattering events are

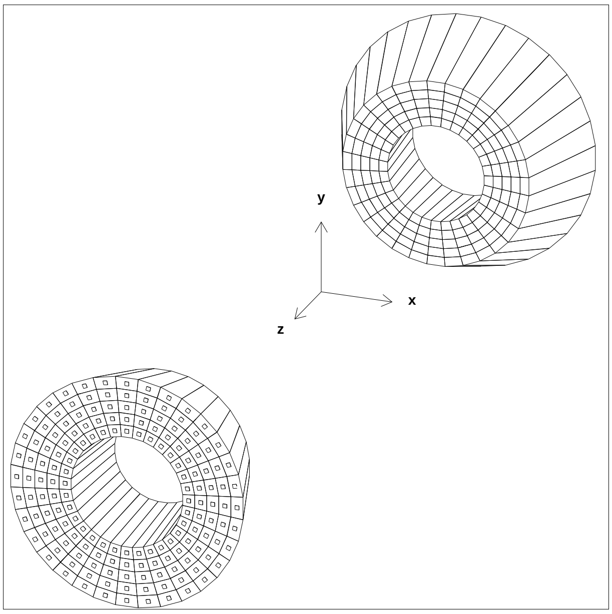


Figure 3.22: An isometric view of the BGO crystals of the forward and backward EFC detectors.

needed for detector calibration and luminosity measurements, but the number of the events to be recorded needs to be reduced about hundred times because the copious amounts of the events are produced.

At an instantaneous luminosity of $10^{34}\,\mathrm{cm^{-2}s^{-1}}$, the rate for physics events of interest is around 100 Hz, and the typical trigger operating output rate is 350 Hz. The Belle data acquisition system can handle rates as high as 500 Hz. Since the beam-related backgrounds depend on accelerator operating conditions, their level cannot be determined accurately and the trigger has to be robust enough to handle large variations in background rates.

Triggering is done using information from each of the sub-detectors, which is processed in parallel and fed to Global Decision Logic (GDL). The trigger is arranged into four levels, denoted as level 0, 1, 3 and 4 respectively.

The level 0 trigger (L0) is a prompt timing signal from the TOF which forces the SVD into the HOLD state.

The level 1 trigger (L1) is implemented in hardware. It is made up of sub-detector triggers which fed to the GDL. The GDL sources information from all sub-detectors without the SVD. All triggers processed in parallel are used by the GDL to characterize the event type. The CDC provides $r-\phi$ and $r-z$ track trigger signals. The TOF trigger system provides an event timing signal and delivers information on a hit multiplicity and topology. The ECL provides two triggers based on total energy deposition and a cluster multiplicity sensitive to different types of hadronic events. The KLM provides a high efficiency trigger for muon tracks. The trigger timing is provided by the TOF, otherwise the ECL is used. The Level 1 trigger configuration is depicted in Figure 3.23. To keep hadronic events, the GDL typically relies on three main trigger classes; multi-tracks from the CDC, total energy deposition and isolated cluster counts from the ECL. Each provides more than 96 % efficiency for hadronic events individually, and 99.5 % efficiency is achieved by the combination.

The level 3 trigger (L3) is implemented in software in an online computer farm. Using an ultra-fast track finder it requires at least one track with an impact parameter in z less than 5.0 cm and the total energy deposit in the ECL to be greater than 3.0 GeV. The trigger reduces the event rate by $50 - 60$ % while retaining 99 % of events of interest.

The level 4 trigger (L4) is implemented in software and performed offline on a computer farm before full event reconstruction. Its purpose is to reduce the amount of data that goes to the full event reconstruction, and its algorithms are optimized for speed. The conditions that activate the trigger are:

- Events are accepted in case that those are tagged by the hardware L1 trigger preselection to be used for luminosity measurement, detector calibration or beam-background studies.
- A total ECL energy deposit is less than $4\,\mathrm{GeV}/c^2$ by the fast cluster-finder routine. To reduce background from cosmic rays, this is vetoed by events with coincident KLM and ECL hits as encoded in L1 trigger information.
- At least one track with r and $|z|$ distances to IP is less than 1.0 cm and 4.0 cm, respectively and p_T is higher than $300\,\mathrm{MeV}/c$.
- For monitoring purposes, 1 % of non-triggered events are kept.

The criteria retain hadronic events with an efficiency of 99.8 % while reducing the total event trigger rate by around 73 %.

3.3.2 Data Acquisition System

The Data Acquisition (DAQ) system manages and stores the data collected at the Belle detector. It is able to process data at the trigger rate of 500 Hz, while having a dead-time of less than 10 %. The system is shown in Figure 3.24. The data from each sub-detector are processed in seven parallel sub-systems, and transformed into full event records by the Event Builder. The full event records are sent to the online farm, where the data are filtered through the L3 software trigger and transformed into the offline event format suitable for offline analysis. A single event size is approximately 30 kBytes of disk space, which translates into a data flow of 15 MBytes/s. The data is sent to the tape library at the computer center 2 km away, where it are written to tapes by a high-speed tape drive. The data monitoring system analyzes events at a rate of approximately 20 Hz of events and can be plugged into the data stream without affecting it.

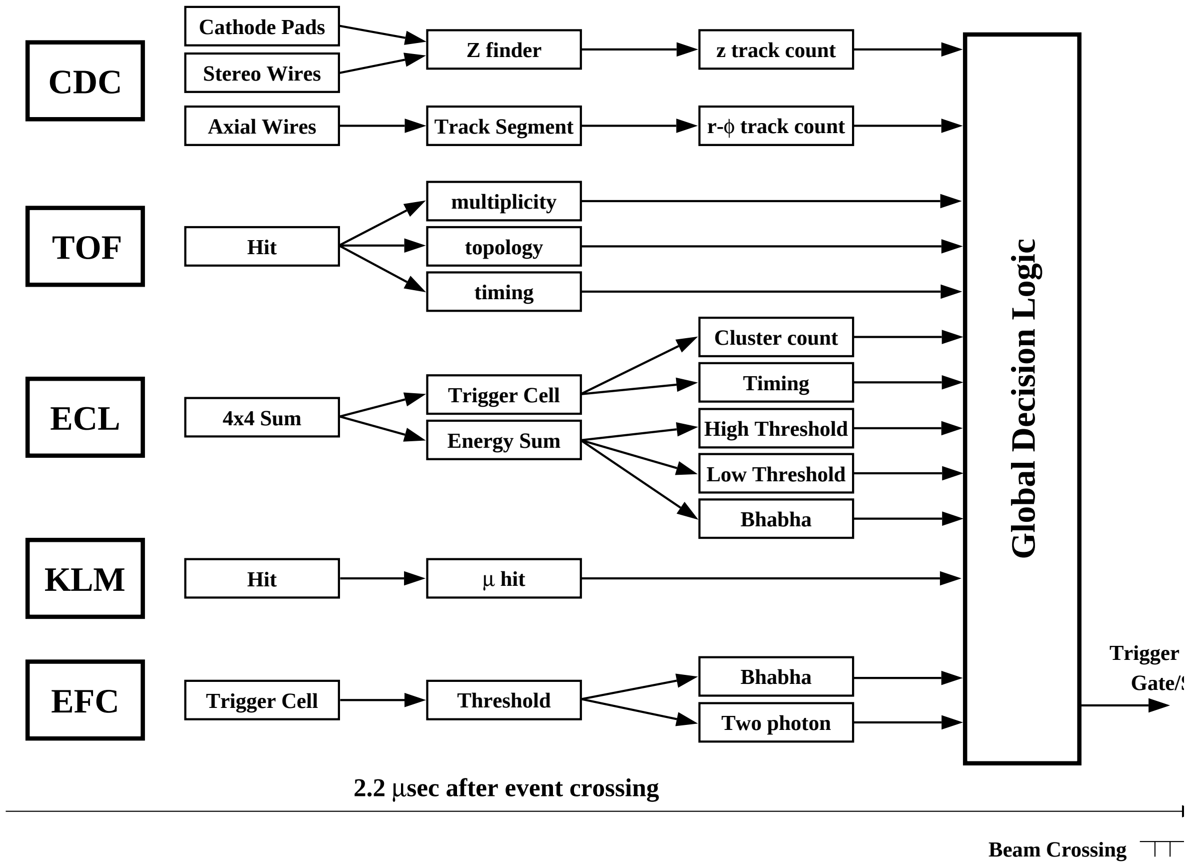


Figure 3.23: The overview of the Level 1 trigger system.

The offline computer farm filters the data written to the tapes through the L4 trigger, where a fast event reconstruction is performed to reject uninteresting data. Afterwards, full reconstruction of the events is performed and the data are translated into a Data Summary Tape (DST) format. A DST is made up of higher level data structures with physical quantities of the decay, for example four-vectors of particle momenta.

Further analysis filters events into hadronic, Bhabha, τ-pair, μ-pair and two-photon event skims. The skims are saved into mini data summary tape (MDST) files. The MDST is a subset of the DST, containing the data needed for physics analyses.

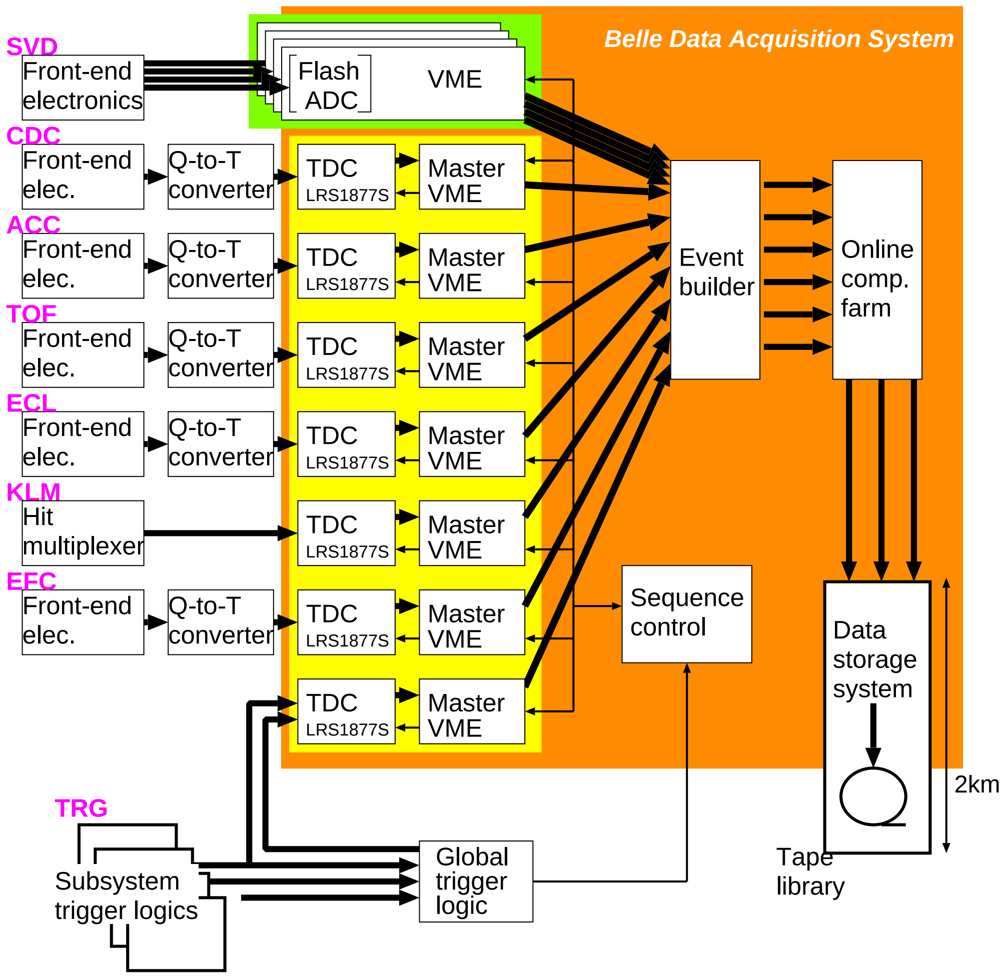


Figure 3.24: The overview of the Belle DAQ system.

3.4 Monte Carlo Simulation

Analysis of data requires a detailed understanding of detector effects, possible background components concerning the analysis and the interpretation of results. A large sample of Monte Carlo (MC) simulation is used, in which the theoretical understanding of physical processes in observed decays and our knowledge of detector effects are incorporated. The amount of the sample is usually several times larger than that of data collected. Two levels of simulation are present, one focused on the physics of decays and the other on simulation of the interaction of particles with the detector.

3.4.1 Event Generators

Event generators focus on the description of physical processes occurring at the decay of particles produced in the e^+e^- collision. The description includes decay chains of all the particles and the kinematical properties of their decays, such as position four-vectors of all decay vertices and momentum four-vectors of all decaying particles.

Two event generators have been used in the simulation of the current analysis, `QQ98` [39] and `EVTGEN` [40]. Both generators are dedicated for modeling B meson decays. Hadronic continuum events, namely $e^+e^- \to q\bar{q}$ interactions where $q = (u, d, s, c)$ is a quark flavor, are generated using `JETSET` [41, 42] which is based on the LUND string fragmentation model [43]. `QQ98` uses a decay table in which decay modes, their decay models, branching fractions, lifetimes and decay parameters are given by the user. The decay table information is usually composed from world averages. The `EVTGEN` event generator is also used in Belle analyses, and has the advantage in that it uses decay amplitudes instead of probabilities, and can simulate the entire decay tree from the amplitudes of each branch. Both are phenomenological in nature and rely on a detailed description of decays of interest.

3.4.2 Simulation of Detector Response

After the decay chains are generated, they are passed to modules that propagate each particle through the detector. The detector geometry is described using `GEANT` [44] which simulates the passage of elementary particles through the matter. A set of detector simulation modules based on `GEANT` is grouped in `GSIM` which is framework to use the `GEANT` library. Detector parameters are continually updated with current experimental conditions and information from real data studies.

3.5 Particle Reconstruction

3.5.1 Reconstruction of Charged Particle Tracks

Charged particles crossing the detector leave tracks in the tracking detectors, the CDC and the SVD (for the description see 3.2.4 and 3.2.3). The event timing by TOF and Level4 Trigger is used to discriminate between hits produced by tracks from beam background and tracks of interest. The tracks are first searched by using hit information obtained from the CDC, where axial wire hits provide $r - \phi$ coordinates, while stereo wire hits measure z coordinates. Since the degree of non-uniformity of the magnetic field is small, the hits of these reconstructed tracks are fitted with a helix with the following parameters: κ, (the reciprocal to the transverse momentum), the slope of the track, and three pivot point coordinates (the point of closest approach to the detector origin). The helix also neglects energy loss due to ionization and multiple scattering.

Then, the hits in the SVD are matched to the fitted tracks and the final tracks are fitted through a non-homogeneous magnetic field using Kalman filter algorithm, where energy loss due to ionization and multiple scattering is accounted for. Finally, to enable muon identification, the tracks are extrapolated all the way to the KLM by solving equations of motion with a Runge-Kutta method. Corrections are applied to the momentum obtained from helix parameters to compensate for the stronger non-uniform magnetic field effects in the extreme forward and background regions. The corrections are calculated from observed shifts of invariant mass peak positions of known particles [45].

The tracking provides both track parameters and their error matrices that are needed for reliable fitting of kinematical constraints.

3.5.2 Reconstruction of Photon Clusters

The ECL is constructed in such a way that a large part of the energy of an electro-magnetic shower produced by a photon is deposited in the ECL. The crystal with the largest energy deposit is taken as a seed of a cluster of hits and the energies of 3×3 and 5×5 counters around the seed are summed up to form the cluster energy. The position of the cluster is obtained by calculating the “center of gravity” of energy, and the momentum vector of each photon is calculated from the position and the energy of the cluster. If clusters from different photons overlap, the overlapping regions is unfolded by comparing the ratio of non-overlapping energy depositions of the two clusters. The ECL is calibrated to obtain the global correction factors and the correction factors of each crystal.

3.5.3 Charged Particle Identification

Present analysis depends strongly on an efficient particle identification. Prompt charged leptons, electrons and muons, are used to recognize semileptonic decays, while a presence of a kaon in the decay signals the background $b \to c$ transition. For the former the leptons have to be successfully separated from hadrons, while for the latter kaon/pion separation is crucial. The particle identification is done based on the information from several detector sub-systems: CDC, ACC, TOF, EFC and KLM.

3.5.4 Muon Identification

Muons are heavy charged leptons that loose their energy mainly by ionization in the detector material. Muons with momenta grater than $500\,\mathrm{MeV}/c$ can easily penetrate to outermost part of the detector, the KLM. To identify a track produced by a muon, the reconstructed track is extrapolated to the KLM and associated hits are searched for within 25 cm of the extrapolated track. Two quantities are used to test the hypothesis that a track is a muon: the difference between the measured and expected ranges of the track (ΔR), and the normalized transverse deviations of all hits associated with the track (χ_r^2). The probability for a hypothesis is constructed by multiplying the separate probabilities (assuming a weak correlation between the two quantities): $p(\Delta R, \chi_r^2) = p_1(\Delta R) \cdot p_2(\chi_r^2)$. Muon candidates are selected based on the value of the normalized ratio

$$\texttt{Prob}(\mu) = \frac{p_\mu}{p_\mu + p_\pi + p_K}. \tag{3.9}$$

The efficiency for muon selection and the pion fake rate for two different $\texttt{Prob}(\mu)$ selections are shown in Figure 3.25.

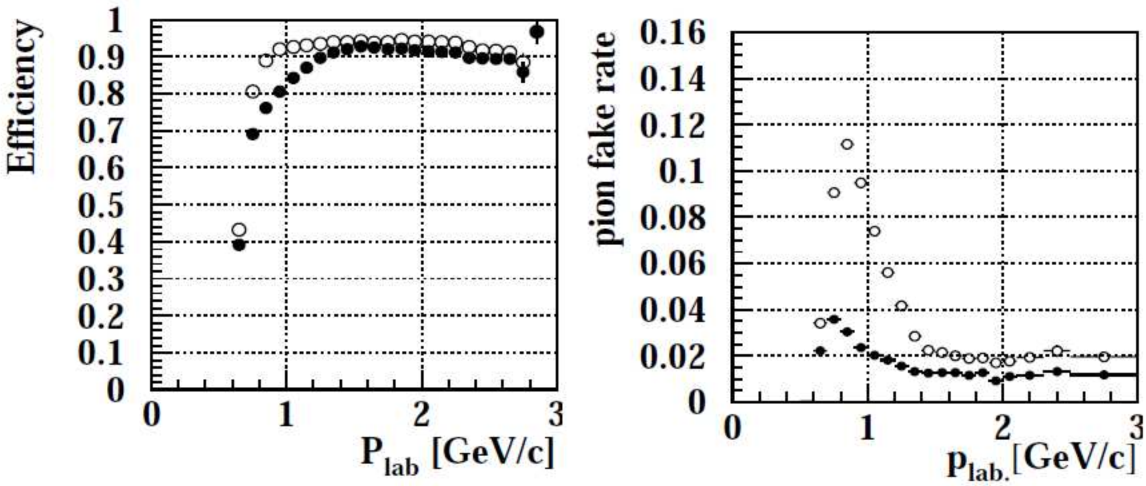


Figure 3.25: The efficiency for muon selection (left) and the pion fake rate (right) in the barrel as a function of the lab momentum, measured in $e^+e^- \to e^+e^-, \mu^+\mu^-$. Open circles for Prob($\mu$) > 0.1, closed circles for Prob(μ) > 0.9. From Ref [46].

3.5.5 Electron Identification

An electron produce a narrow in ECL in which the electrons deposits nearly all of its energy. The energy deposition and the difference in the velocity at a given measured momentum, obtained from CDC and ACC, are used in the electron identification. Information from TOF is not included, since the timing resolution does not permit separation of electrons from pions.

Five discriminating variables are used in the electron identification:

Track to Cluster Matching: a electron track is required to match the position of an ECL cluster. The matching is assessed by a χ^2 like variable based on the separation of the extrapolated track and the center of the ECL cluster.

E/p: the ration of deposited energy in the ECL to the momentum measured by the CDC. Since a electron leaves a large part of its energy in the calorimeter, and its mass is negligible compared to the energy, $E \approx p$ and $E/p \approx 1$ (see Figure 3.26(a)). Hadron leaves only a fraction of its energy in the ECL and has the ratio below 1 - as well as a small part of electrons that have lost some energy in the material before reaching the ECL.

E9/E25: since the shape of the distribution of the electron energy deposit is narrow, the transverse shower shape in compared by observing the ratio of deposited energy in 3×3 (E_9) and $5{\times}5$ (E_{25}) crystals. The ratio is close to one (≈ 0.95) for electrons, while it differs from one for hadrons, since their passage instigates more than one shower (see Figure 3.26(b)).

dE/dx**:** the energy loss due to ionization along a charged track's trajectory is measured by the CDC. The energy loss is dependent on particle's velocity β, which provides excellent separation between electrons and pions for momenta greater than $0.5\,\mathrm{GeV}/c$, as shown in Figure 3.26(c).

ACC light yield $\langle N_{pe} \rangle$: the presence or absence of photoelectrons from Cherenkov effect in the ACC can reveal the type of the passing particle, since the threshold for emitting photons is different for different particles. the threshold value is a few MeV for electrons and in the momentum range $0.5 - 1\,\mathrm{GeV}/c$ for pions. The separation of electrons and pions is thus possible only in the momentum range below $1\,\mathrm{GeV}/c$.

Likelihood for electron and pion hypothesis is constructed by combining the probability density functions from the five variables. The overall likelihood used for identification of an electron is defined as the sum of products of likelihoods from a single variable:

$$\mathtt{Prob}(e) = \frac{\Pi_{i=1}^{5} \mathcal{L}_i^e}{\Pi_{i=1}^{5} \mathcal{L}_i^e + \Pi_{i=1}^{5} \mathcal{L}_i^\pi}. \tag{3.10}$$

The distribution of the overall normalized likelihood can be seen in Figure 3.26(d). The efficiency for electron selection and the pion fake rate as a function of the lab-frame momentum, measured in radiative Bhabha events, is shown in Figure 3.27.

3.5.6 Identification of Charged Hadrons: K/π Separation

The identification of charged hadrons, mainly kaons and pions, is performed using the combined information on the specific ionization dE/dx measurement (CDC), the time-of-flight measurement (TOF) and the measurement of the number of photoelectrons in the ACC, to cover the typical momenta of hadrons (see Figure 3.28). The refractive indices of aerogel Cherenkov radiators in the ACC are optimized for successful separation for high momentum hadrons ($1.2 < p < 3.5\,\mathrm{GeV}/c$), and the likelihood for different particle hypotheses is calculated from the obtained light yield.

The TOF is used to measure particle velocities from the time used for a particle to fly over a certain distance, and is useful for separation of kaons and pions with low momentum below $1.2\,\mathrm{GeV}/c$. The likelihood is calculated from the difference between the expected time of the

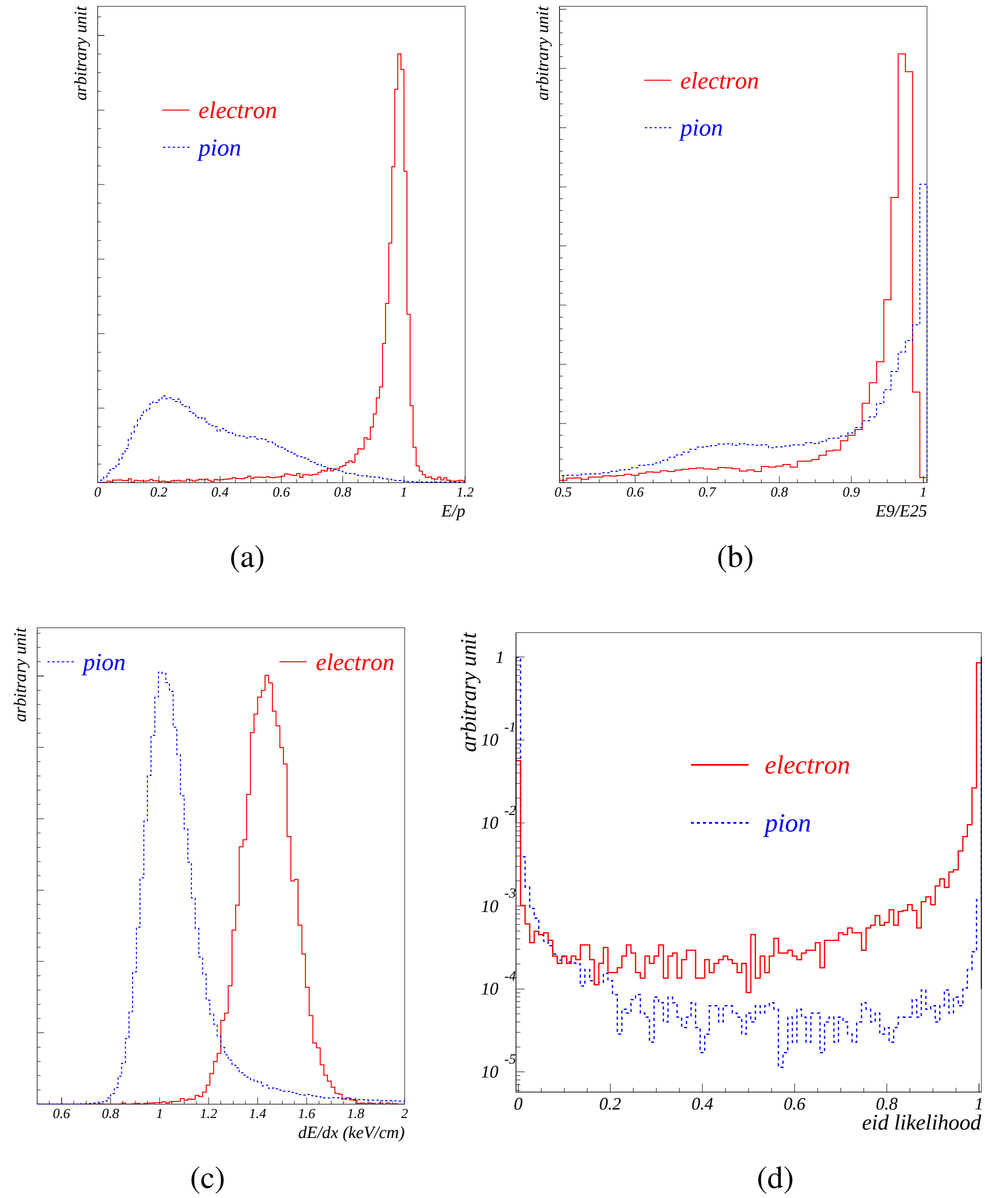


Figure 3.26: (a) Ratio of energy deposit to track momentum, E/p, (b) Transverse energy shape, $E9/E25$, (c) Rate of ionization energy loss dE/dx, for electrons (solid line) and pions (broken line). (d) The electron likelihood, Prob(e), for electrons (solid line) and pions (broken line). From Ref. [47].

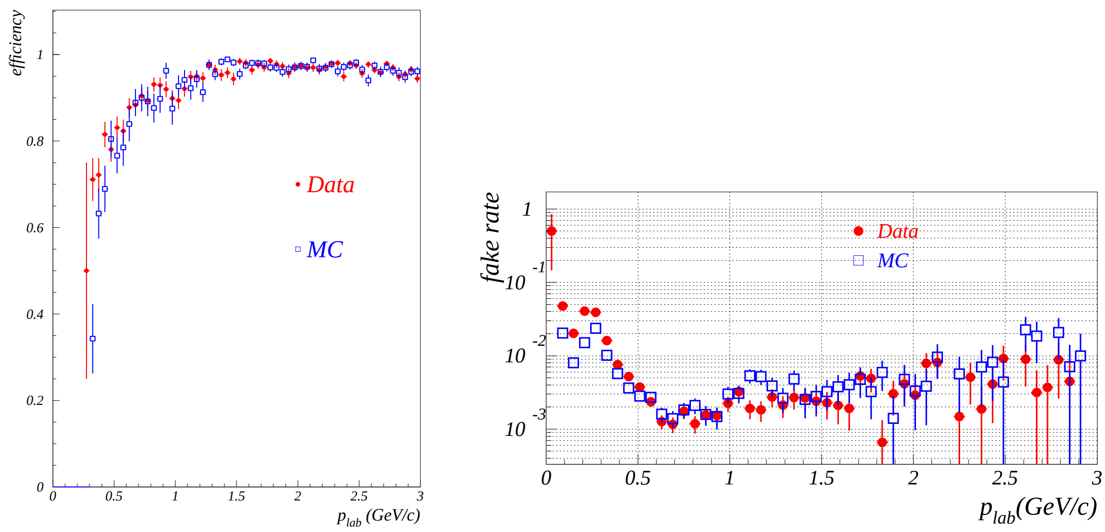


Figure 3.27: The efficiency for electron selection (left) and the pion fake rate (right) as a function of the momentum in the laboratory frame, measured in radiative Bhabha events. From Ref. [47].

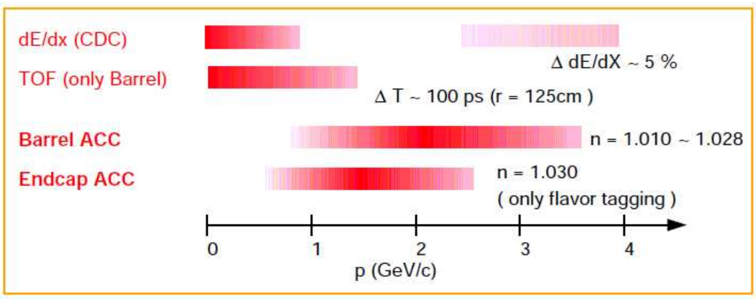


Figure 3.28: Momentum coverage of kaon/pion separation at the Belle detector.

flight for a hypothesis and the measured time:

$$\mathcal{L}_{\text{TOF}} = \frac{e^{-\frac{1}{2}\chi^2_{\text{TOF}}}}{\sqrt{2\pi}\sigma_{\text{TOF}}}, \quad \chi^2_{\text{TOF}} = \left[\frac{t_{meas} - t_{hyp}}{\sigma_{\text{TOF}}}\right]^2. \tag{3.11}$$

Similarly, the likelihood obtained from the measurement of ionization loss is obtained as:

$$\mathcal{L}_{dE/dx} = \frac{e^{-\frac{1}{2}\chi^2_{dE/dx}}}{\sqrt{2\pi}\sigma_{dE/dx}}, \quad \chi^2_{dE/dx} = \left[\frac{(dE/dx)_{meas} - (dE/dx)_{hyp}}{\sigma_{dE/dx}}\right]^2. \tag{3.12}$$

The total likelihood of a hypothesis is obtained as the product of three kinds of the likelihood:

$$\mathcal{L} = \mathcal{L}_{\text{ACC}}(hyp) \times \mathcal{L}_{\text{TOF}}(hyp) \times \mathcal{L}_{dE/dx}(hyp). \tag{3.13}$$

The separation of charged hadrons is achieved by calculating the probability .

`Prob(signal/background)` of the signal particle hypothesis, when separating it from the background particle. The probability is calculated from the likelihoods of signal and background hypotheses:

$$\texttt{Prob(signal/background)} = \frac{\mathcal{L}(\text{signal})}{\mathcal{L}(\text{signal}) + \mathcal{L}(\text{background})}. \tag{3.14}$$

The signal and background particles can be any of the following: e, p, K, π. The kaon selection efficiency and the pion fake rate for $\texttt{Prob}(K/\pi) > 0.6$ is shown in Figure 3.13.

Chapter 4

Data Sets

This analysis is based on an integrated luminosity of 253 fb^{-1} data sample accumulated at the $\Upsilon(4S)$ resonance (on-resonance) with the Belle detector. It corresponds approximately to $(276.6 \pm 3.1) \times 10^6$ pairs of $B\bar{B}$. The off-resonance data sample is accumulated at a center-of-mass (CM) energy 60 MeV below the $\Upsilon(4S)$ resonance, and is corresponding to an integrated luminosity of 28.1 fb^{-1}. It is used to estimate the continuum background events as light quark decay ($e^+e^- \to q\bar{q}$, where $q = u, d, s, c$). The other B meson decays except for the signal decays, $B^\pm \to \mu^\pm\nu_\mu$ and $B^\pm \to e^\pm\nu_e$, are also background events. The background events from other B meson decays are estimated by MC sample generated by the EvtGen [40] event generator. The generic $B\bar{B}$ decays MC sample, in which the semi-leptonic B meson decays via $b \to c$ transition ($B \to X_c\ell\nu_\ell$) are dominant, are generated to be an integrated luminosity of about 430 fb^{-1}. The $B \to X_u\ell\nu$ decay basically has one higher momentum lepton track than one from other B meson decays ($B \to X_c\ell\nu_\ell$). The $B \to X_u\ell\nu$ MC samples are generated to be an integrated luminosity of 970 fb^{-1}. The signal decay MC samples for $B^\pm \to \mu^\pm\nu_\mu$ and for $B^\pm \to e^\pm\nu_e$ have been respectively generated to be $60K$ events, which are corresponding to integrated luminosities of 1.2×10^5 fb^{-1} and 5.5×10^{10} fb^{-1}. All MC samples also depend on the same conditions of the Belle experiment, for example beam energies and background conditions.

The amount of each data sets are summarized in Table 4.1.

Data Type	Integrated luminosity[fb^{-1}]
On resonance	253.12
Off resonance	28.11
Generic $B\bar{B}$ MC	430
$X_u\ell\nu$	970
Signal MC($B^\pm \to \mu^\pm\nu$)	1.2×10^5
Signal MC($B^\pm \to e\nu$)	5.5×10^{10}

Table 4.1: Table of data sets used for the analysis. The luminosities of the signal MC assume that cross sections are 4.7×10^{-7} for the muon decay and 1.0×10^{-12} for the electron decay.

Chapter 5

Analysis

5.1 Particle Identification

In the signal decays $B^+ \rightarrow \mu^+\nu_\mu$, $B^+ \rightarrow e^+\nu_e$, one lepton should be detected in the signal side, therefore all detected particles except for one signal candidate lepton are inclusively reconstructed to be a companion B meson.

Before the inclusive reconstruction for a companion B meson and identification of a signal candidate lepton, particle types of all tracks and ECL clusters need to be identified. Though the general probabilities (likelihood value) of particle identification have been computed (see section 3.5 and so on), we should determine all particle types depending on the probabilities, the momentum and other informations. The Figure 5.1 shows the flow of the particle identifications. The following subsections describe each of the particle identifications in details.

5.1.1 K_S reconstruction

Firstly, K_S mesons are reconstructed from vector meson particle candidates (`MDST_Vee2`), which have been already accumulated to be candidates of K_S mesons and gamma conversion. A K_S meson can be identified by using secondary vertex reconstruction. We follow the standard criteria used for the Belle experiment, which are summarized in table 5.1. An invariant mass of two pions is required $0.46 < M_{\pi\pi} < 0.53\,GeV/c^2$.

items	$p_{K_S} < 0.5\,\mathrm{GeV}/c$	$0.5 < p_{K_S} < 1.5\,\mathrm{GeV}/c$	$p_{K_S} > 1.5\,\mathrm{GeV}/c$
dz at second vertex	$< 0.8\,cm$	$< 1.8\,cm$	$< 2.4\,cm$
dr of impact parameter	$> 0.05\,cm$	$> 0.03\,cm$	$> 0.02\,cm$
Deflection angle	$< 0.3\,rad.$	$< 0.1\,rad.$	$< 0.03\,rad.$
K_S flight length	N.A.	$> 0.08\,cm$	$> 0.22\,cm$

Table 5.1: Reconstruction criteria of K_S mesons for various K_S momentum ranges

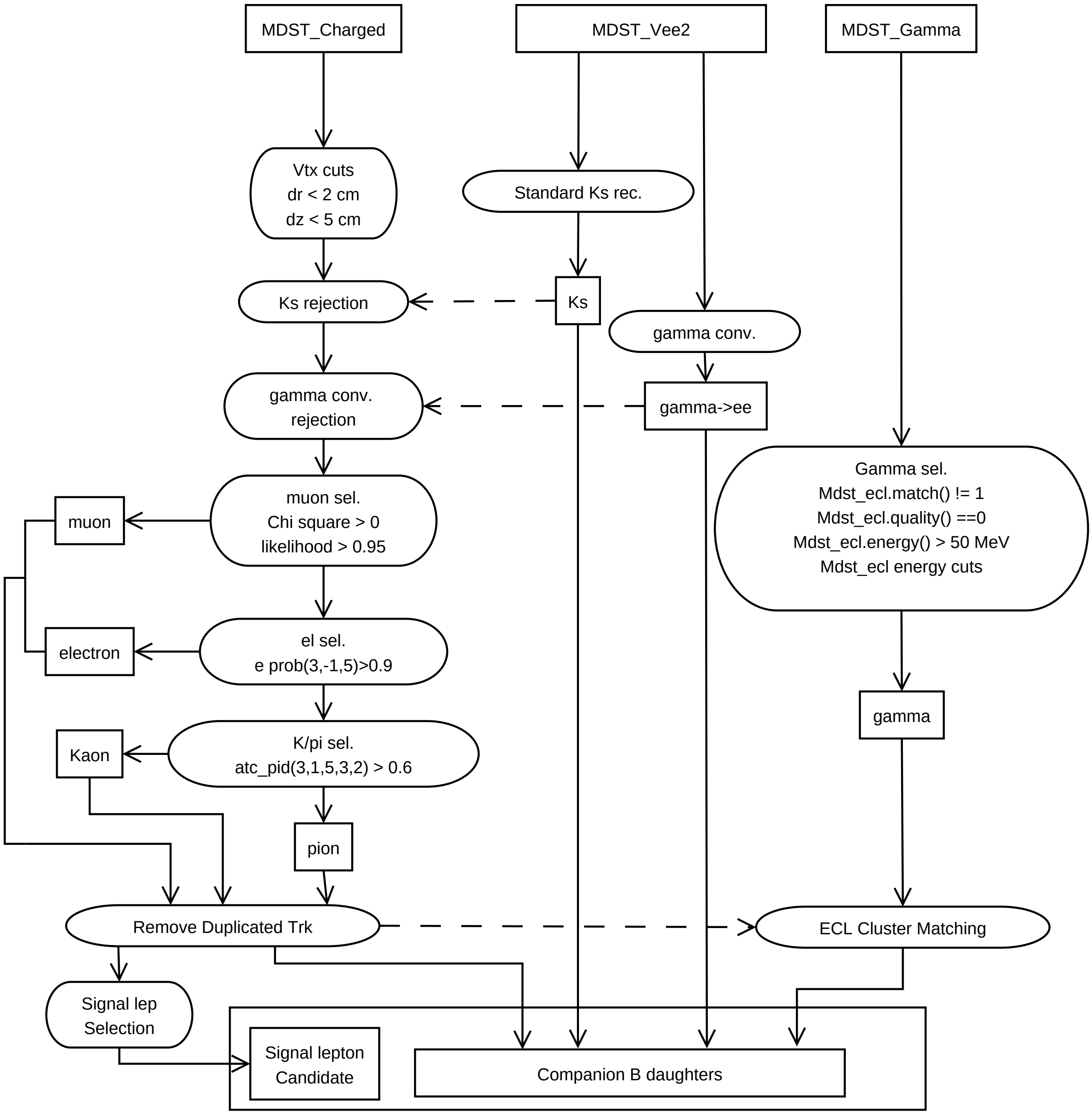


Figure 5.1: Particle identification (PID) flow chart: The boxes included MDST show the MDST data classes. The round boxes represent PID criteria. A particle name in a box shows an identified particle. The normal arrows show flows of particle identification. Dash arrows show that two connected (round) boxes are related.

5.1.2 Reconstruction of γ Conversion

Some of γs passing through the Belle detector are converted into electron-positron pairs through the electromagnetic interaction. The events are called γ(photon) conversion. They are reconstructed from `MDST_Vee2` An invariant mass of an electron-positron pair are required $M_{ee} < 0.1\,GeV/c^2$.

5.1.3 Identification of Charged Particle

All charged tracks are accumulated in the `MDST_Charged` data. Charged tracks are classified four kinds of particles, which are muon, electron, kaon and pion. If a charged track has been reconstructed to be a K_S meson or γ conversion, it is explicitly vetoed. We require tight identification criteria for leptons in order to reject punch-through pions and kaons polluting lepton purity in higher momentum region. The following criteria are required for charged tracks.

- length of radial direction cut : $dr < 2.0\,cm$
- length of beam axis cut : $dz < 5.0\,cm$
- Particles reconstructed to be K_S and photon conversion are vetoed.
- Remove duplicated tracks (See next section)

The following criteria are required for the determination of particle type.

- μ : muon probability > 0.95 (see section 3.5.5)
- e : electron probability > 0.9 (see section 3.5.4)
- K : kaon probability > 0.6 (see section 3.5.6)
- π : otherwise failing the criteria for μ, e and K

Duplicated tracks

For the identified charged particles, we check whether they are the fake track or not. When the transverse momentum(p_T) of the track is as low as it curls back inside of the CDC, and the track finder sometimes makes duplicated tracks (see Figure 5.2). The duplicated tracks may have the same or opposite charge of the original track. A duplicated pair with the same charge have the almost same transverse momentum and the opening angle ($\phi_{\rm 2tracks}$) between the original track and the same charge duplicated track should be small. A duplicated pair with opposite charge have almost same momentum but the opening angle is close to 180°. We use information based on the impact parameters. A pair of duplicated tracks are defined as follows:

- same charge : $\Delta p_T < 0.1\,GeV/c$ and $\cos\phi_{\rm 2tracks} < 0.9$

- opposite charge : $\Delta p_T < 0.1\,GeV/c$ and $\cos\phi_{\text{2tracks}} < -0.9$
- A duplicated track $p_T < 0.3\,GeV/c$

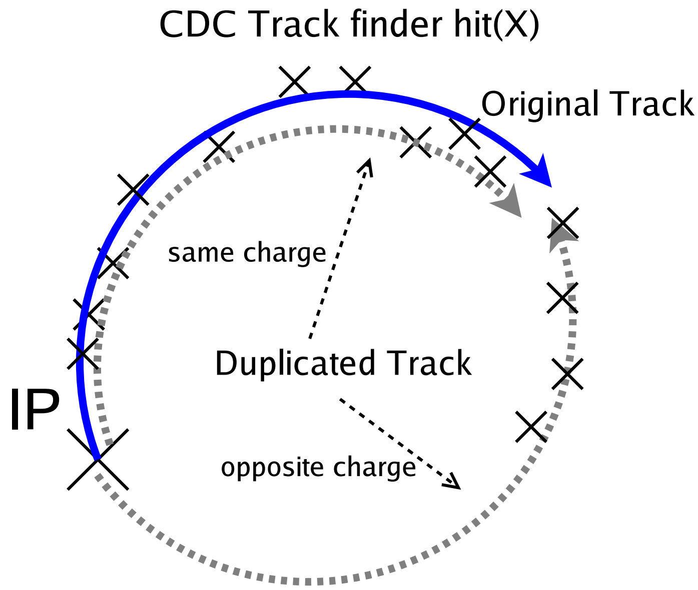


Figure 5.2: Schema of duplicated tracks with a same charge track and an opposite charge track. Thick curve shows an original track and dashed curves show duplicated tracks. Small x marks show hits for the CDC track finder. A large x mark shows the IP position.

In order to determine which track is the duplicated track(fake track), we compare $|\Delta r[cm]/20|^2 + |\Delta z[cm]/100|^2$ between two tracks and select the one which has the smaller value. If there are more candidates, we select the smallest one [48].

A charged particle which is regarded as a duplicated track is removed.

5.1.4 Identification of Gamma (γ)

ECL clusters made by the beam background can be removed by introducing energy cuts. We use the different energy cuts for the barrel part and the end-cap parts, since the effect of the beam background is severe in the end-caps. γ-candidate clusters satisfy the following criteria:

- $E_\gamma > 0.05\,GeV$: barrel region
- $E_\gamma > 0.10\,GeV$: forward end-cap region
- $E_\gamma > 0.15\,GeV$: backward end-cap region

The ECL clusters associated to tracks are removed from gamma candidates [48].

5.1.5 Signal Lepton Selection

The signal candidate lepton is selected from particles identified as lepton, and the momentum in the CM frame is required to be $2.2 < p^* < 3.0\,GeV/c$. Because the signal decay modes are two body decay, the lepton has about a half of momentum of B meson mass. Especially, the momentum must be monochromatic momentum spectrum in the B rest frame as shown in Figure 5.3. However the signal side B meson can not be reconstructed because the neutrino can not be detected. The signal lepton momentum is approximated by information of the companion B meson according to the following formula. The method of the reconstruction for the companion B meson is described in next section.

$$p_\ell^B \simeq p_\ell^* \left(1 - \frac{|\vec{p}^*_{B^{\rm sig}}|}{m_B} \cos\theta_{\ell - B^{\rm sig}}\right) \tag{5.1}$$
$$\simeq p_\ell^*(1 + 0.06\cos\theta_{\ell - B^{\rm comp}}) \tag{5.2}$$

where p_ℓ^B and p_ℓ^* are respectively lepton momenta in the B rest frame and the CM frame, $\cos\theta_{\ell - X}$ represents the cosine of the opening angle between the signal candidate lepton and the signal B meson or between the signal candidate lepton and the companion B meson. Moreover the approximation $-\frac{|\vec{p}^*_{B^{\rm sig}}|}{m_B}\cos\theta\ell - B^{\rm sig} \simeq +0.06\cos\theta_{\ell - B^{\rm comp}}$ tells that the momentum of the signal side B meson is extracted by the companion B meson and the magnitude of momentum $|\vec{p}^*_{B^{\rm sig}}|$ is approximately $\sqrt{E^2_{\rm beam} - m_B^2} \simeq 0.32\,GeV/c$ where $E_{\rm beam}$ is a half of the beam collision energy and is approximately 5.29 GeV. Figure 5.3 shows the simulated distributions of the lepton momentum in the B rest frame. The peak value of the distributions is shown about a half of B mass.

5.1.6 Companion B Meson Reconstruction

We call a companion B meson for the opposite side of a signal B meson. The companion B meson is inclusively reconstructed from all identified particles except for the signal lepton. Four-momentum of the companion B meson is defined by:

$$\vec{p}^*_{B^{\rm comp}} = \sum_i \vec{p}^*_i \tag{5.3}$$

where $\vec{p}^*_i$ represents a momentum of an identified particle and i is running among all the particles except for the signal lepton candidate.

Then we define the beam constraint mass M_{bc} and ΔE for the companion B meson. These variables are described by:

$$M_{\rm bc} = \sqrt{E^2_{\rm beam} - |\vec{p}^*_{B^{\rm comp}}|^2} \tag{5.4}$$
$$\Delta E = E_{B^{\rm comp}} - E_{\rm beam} \tag{5.5}$$

where $E_{\rm beam}$ is a half of beam collision energy in the CM frame, and $\vec{p}^*_{B^{\rm comp}}$ is the CM frame momentum and $E_{B^{\rm comp}}$ is the CM frame energy for the reconstructed companion B meson. Figure 5.4

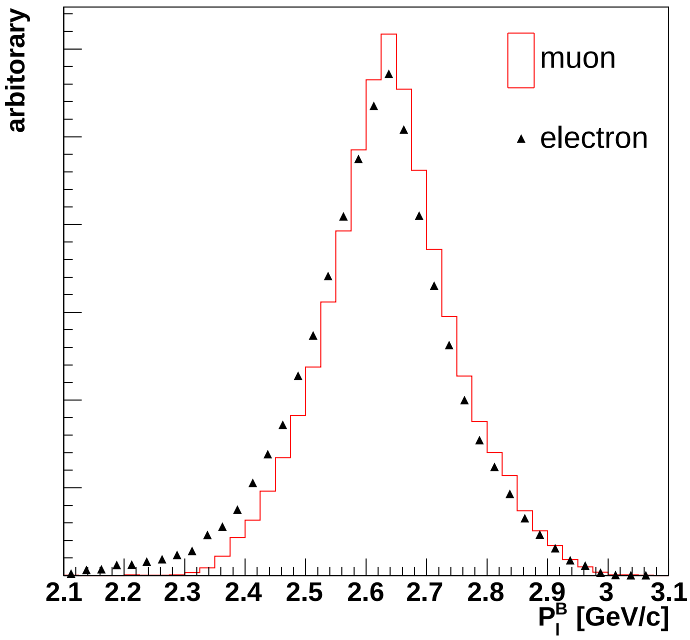


Figure 5.3: p_ℓ^B distribution for the signal MC samples, where p_ℓ^B represents the lepton momentum in the B rest frame.

and Figure 5.5 show the $M_{\rm bc}$ and ΔE distributions for each decay mode after pre-selections (see next section) have been applied, respectively. If collection of the companion B mesons are correct, the $M_{\rm bc}$ distribution have a peak at 5.28 GeV$/c^2$ like a signal MC distribution.

5.2 Event Selection

In this section, we explain the event selection and its optimization. Firstly, the pre-selection which have been already applied to all data samples at the event selection stage is described. The definitions of the signal region, the fit region and the sideband region are also described. After the above explanation have done, the event selection is described. The selection is classified into three sections by the features of the selection. The event selection is constraints on the signal candidate lepton, continuum suppression and neutrino reconstruction. They are explained according to this order. Finally all selection criteria are determined by the results of the selection optimization.

5.2.1 Pre-selection

Most of the events collected by the Belle detector are obviously irrelevant events for this analysis and are removed in advance of the event selection. The number of tracks identified lepton is limited to one, the momentum of the signal candidate lepton in the CM frame is also limited and $M_{\rm bc}$ is required as follows:

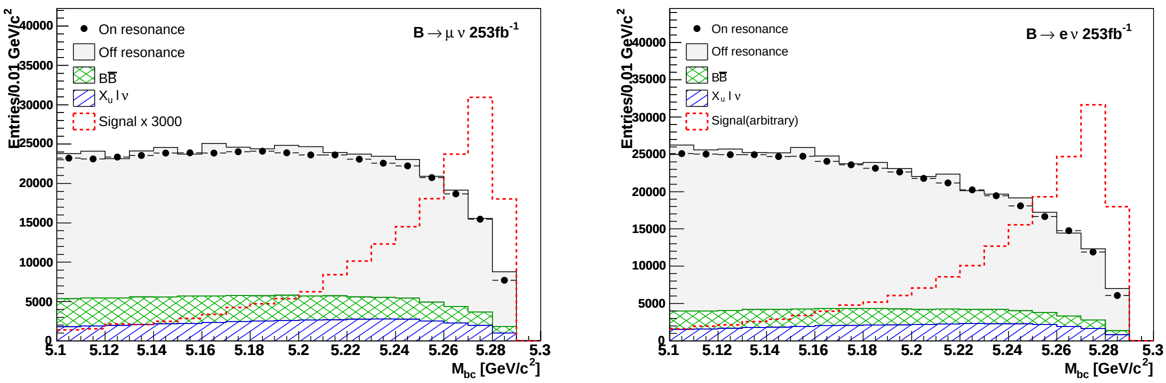


Figure 5.4: $M_{\rm bc}$ distributions of the companion B meson for the muon mode (right) and the electron mode (left) just after the pre-selection have been applied. Points show the on-resonance data, and solid histograms show the expected background due to rare $B \to X_u \ell \nu$ decays (hatched, from MC); other $B\bar{B}$ events, mainly $B \to X_c \ell \nu$ decays (cross-hatched, also from MC); and continuum events (light shaded, taken from scaled off-resonance data). The dashed histograms represent the signal as predicted by the MC with arbitrary normalization. The following similar histograms have the same style.

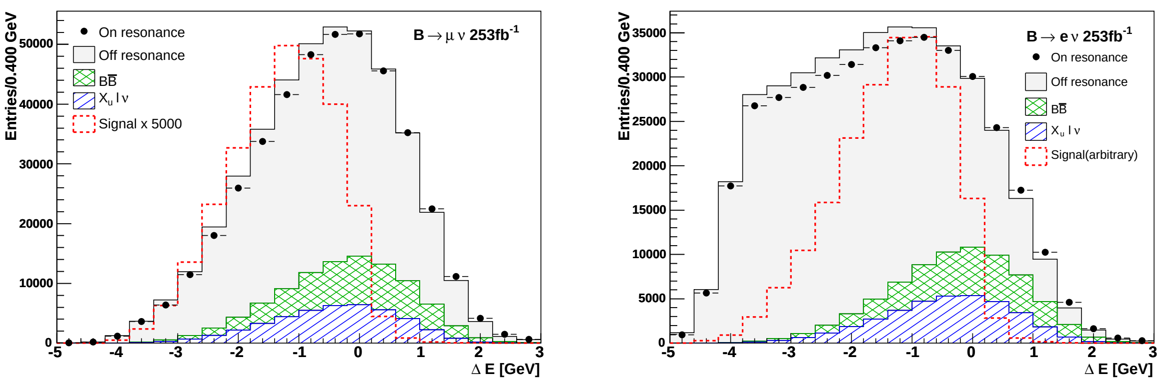


Figure 5.5: ΔE distributions for the muon mode (right) and the electron mode (left) just after pre-selection have been applied. The style of these histograms are the same as Figure 5.4.

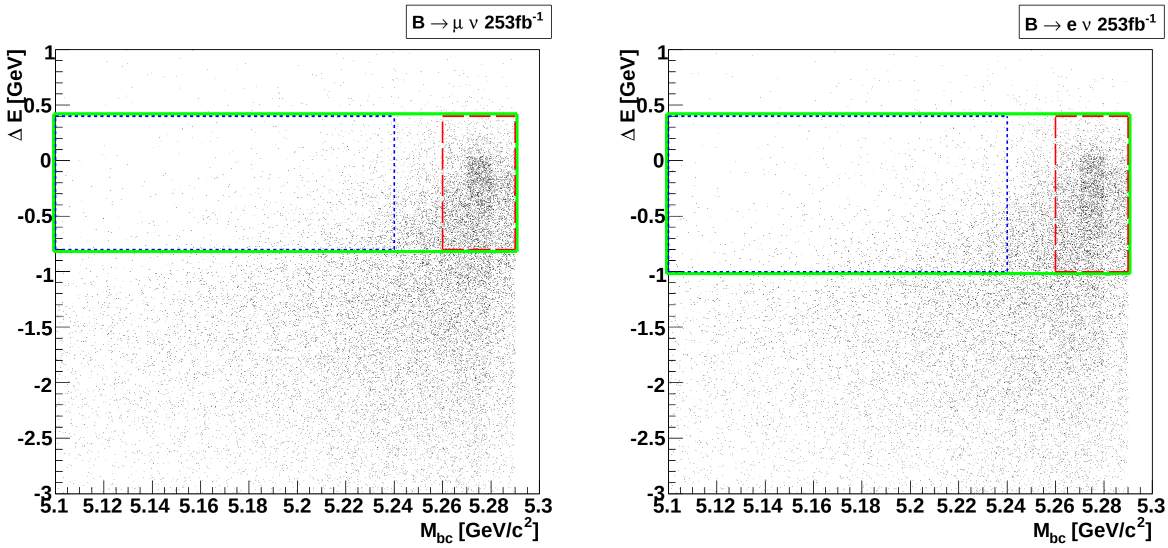


Figure 5.6: Two-dimensional distribution of ΔE vs M_{bc} for the signal MC events. The solid box shows the fit region, the long-dashed box is the signal region and the short-dashed box is the M_{bc} sideband region.

- $2.2 < p^*_\ell < 3.0\,\mathrm{GeV}/c$
- $M_{\mathrm{bc}} > 4.9\,\mathrm{GeV}/c^2$
- N_ℓ equal to 1.

All data samples have been applied the above pre-selection criteria.

5.2.2 Definition of Regions

We have already defined M_{bc} and ΔE for a companion B meson. Figure 5.6 shows the scatter plots of M_{bc} vs ΔE for the signal MC.

The *signal region* in the $M_{\mathrm{bc}} - \Delta E$ space is defined by $5.26 < M_{\mathrm{bc}} < 5.29\,\mathrm{GeV}/c^2$ and $-0.8(-1.0) < \Delta E < 0.4\,\mathrm{GeV}$ for the muon (electron) mode. This region is used to optimize the selection criteria and to extract the signal yield. The width of the *fit region* in M_{bc} increases to $5.10 < M_{\mathrm{bc}} < 5.29\,\mathrm{GeV}/c^2$ compared to the signal region. The fit region is used to fit to extract the signal yield and to set upper limits of the branching fraction. The M_{bc} *sideband region* is also defined in the fit region as $5.10\,\mathrm{GeV}/c^2 < M_{\mathrm{bc}} < 5.24\,\mathrm{GeV}/c^2$.

Table 5.2 summarizes the definition of each region.

5.2.3 Signal Candidate Lepton

We have already described that the signal lepton momentum has about a half of the B meson mass because of the two-body decay. The signal lepton momentum is basically the highest among the ones of particles from the other decays. Figure 5.7 shows the distribution of the

		M_{bc} GeV$/c^2$	ΔE GeV
muon mode	signal region	[5.26, 5.29]	[−0.8, 0.4]
	fit region	[5.10.5.29]	[−0.8, 0.4]
electron mode	signal region	> 5.26	[−1.0, 0.4]
	fit region	[5.10.5.24]	[−1.0, 0.4]

Table 5.2: Summary table of the signal region and the fit region

lepton momentum (p_ℓ^B) in the B rest frame just after the pre-selection have been applied. The p_ℓ^B is required to be $2.6 < p_\mu^B < 2.84\,\mathrm{GeV}/c$ for the muon mode, and $2.6 < p_e^B < 2.8\,\mathrm{GeV}/c$ for the electron mode. These cuts are asymmetric with respect to the peak value of the signal p_ℓ^B distribution because the peaking background events ($B\bar{B}$ and $X_u\ell\nu$) are explicitly removed. This selection can reject greater than 99 % of the peaking background events.

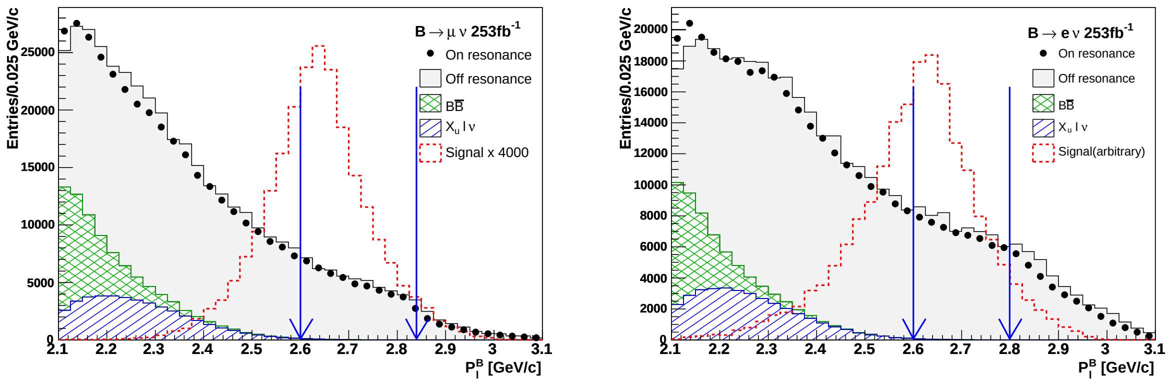


Figure 5.7: The signal candidate lepton momentum in the B rest frame after the pre-selection have been applied. The arrows show the selection criteria. The histograms for the signal MC are normalized to make thm visible.

The signal lepton is required to be the most probable lepton. Figure 5.8 shows the distribution for the cosine of the polar angle of the signal lepton ($\cos\theta_\ell$), where the polar angle is defined by the angle between the opposite direction of the electron beam and the flight direction of the particle. There are higher momentum tracks in the end-cap region from the continuum background because the continuum background event (decays from light quark $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$)) is a jet-like event. Moreover the end-cap region have worse efficiency to detect a lepton. The signal lepton track is required as $-0.5 < \cos\theta_\mu < 0.85$ for the signal muon and $-0.5 < \cos\theta_e < 0.8$ for the signal electron. These selection criteria assure good purity for the lepton samples.

The charged kaons with high momentum may punch through the Belle detector like the signal muon with high momentum. To remove the punch-through kaons, we apply kaon probability to be less than 0.25 for the signal tracks. Figure 5.9 shows the distritbution of kaon probability. This restriction is applied for the signal candidate muon only.

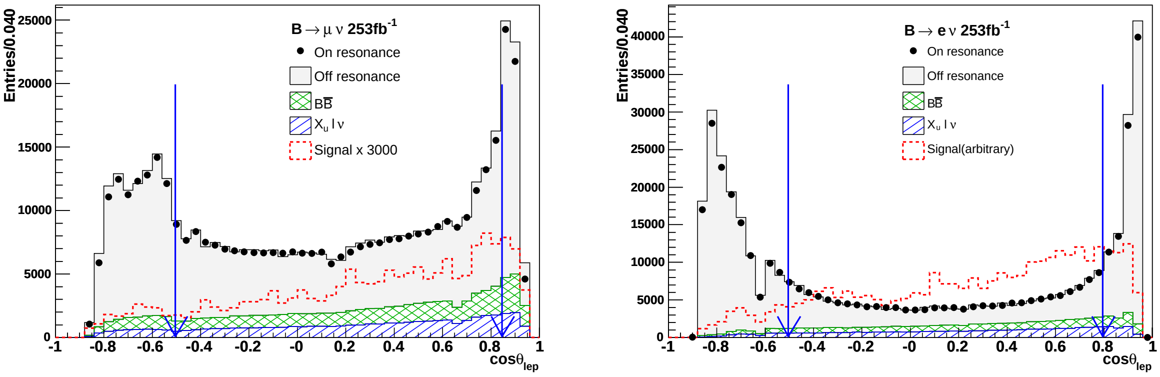


Figure 5.8: Distributions of the cosine of the polar angle of the signal lepton direction. The arrows show the selection criteria. We require the signal candidate lepton goes to the barrel region.

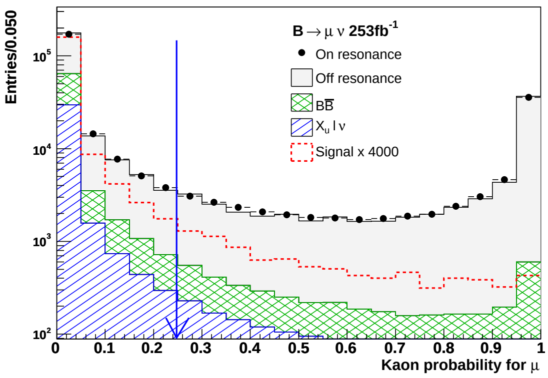
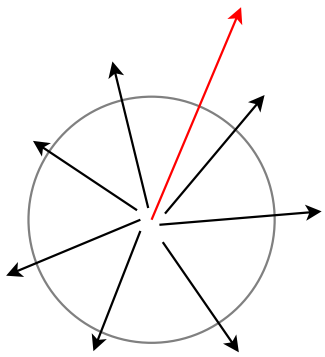


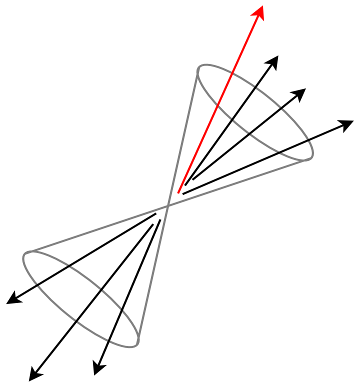
Figure 5.9: The probability distribution of a charged kaon for a muon candidate after $\cos\theta_\mu$ selection criterion have been applied. This selection is only for the muon mode.

5.2.4 Continuum Suppression

The signal lepton momentum is the highest momentum among the momenta of particles from B meson decays. Figure 5.7 shows the distribution of the lepton momentum in the B rest frame. The dominant background is the continuum background, which estimated by is the off-resonance data and is almost $q\bar{q}$ ($q = u, d, s, c$) background. In order to suppress $q\bar{q}$ background, we use the difference of the event topology between the signal events and the continuum events. Because B mesons are produced almost at rest in the CM frame, decay particles from the signal events distribute spherically. On the other hands, since each quark of $q\bar{q}$ events is produced with high momentum, $q\bar{q}$ events tend to have two jets (see Figure 5.10).



(a) spherical event



(b) two jets event

Figure 5.10: Event topology.

We define the following variables in the CM frame.

$$R_l = \frac{\Sigma_i |p_i||p_{\ell ep}| P_l \cos\theta_{i\ell ep}}{\Sigma_i |p_i||p_\ell|} \quad , \quad r_l = \frac{\Sigma_{ij} |p_i||p_j| P_l \cos\theta_{ij}}{\Sigma_{ij} |p_i||p_j|} \tag{5.6}$$

where P_l is a Legendre polynomial of order of l, and i, j run over all the particles from the reconstructed companion B. Then we combine five of them into a Fisher discriminant [49]

$$F = \sum_{l=2,4} \alpha_l R_l + \sum_{l=2,3,4} \beta_l r_l, \tag{5.7}$$

where the coefficients α_l, β_l are optimized to maximize the discrimination between the signal and the continuum background where $q\bar{q}$ MC samples are used as continuum background. Figure 5.11 shows the optimized distributions to discriminate the $q\bar{q}$ MC background events from the signal MC events.

The terms R_1, R_3 and r_1 in Equation 5.7 are excluded from Fisher discriminant, because either of these terms is found to have a correlation with $M_{\rm bc}$. We call F the *Super Fox-Wolfram* (SFW) [50] variable, since the terms are combined in such a way as to enhance the discriminating power of the original Fox-Wolfram moments [49]. We select events with $F > 0.3$ for the

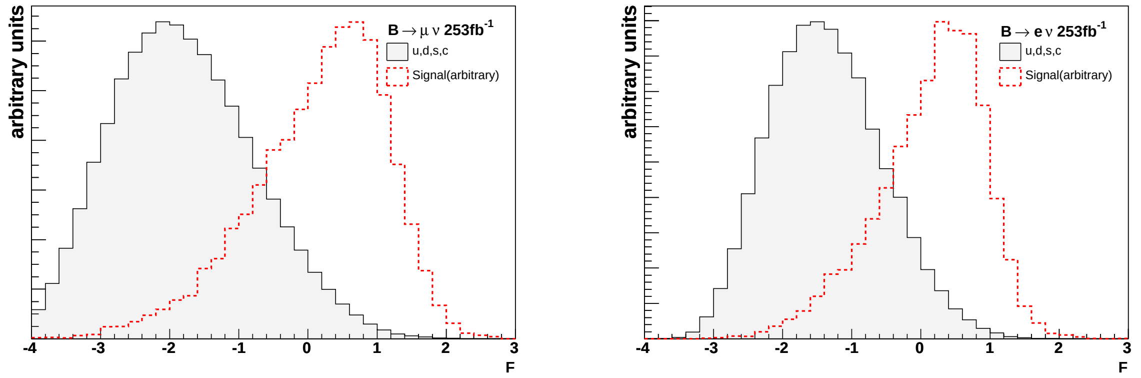


Figure 5.11: Optimized Fisher discriminate momenta distribution. Dashed histogram shows the signal MC. Light shaded histogram shows the $q\bar{q}$ ($q = u, d, s, c$) MC samples.

muon mode and $F > 0$ for the electron mode, which are a compromise between the statistical significance of the signal and the size of the systematic error due to the fitting to extract the signal yields. Figure 5.12 shows the SFW distributions for the on-resonance data, for the background MC events, for the off-resonance data and for the signal MC events. Then greater than 99 % and 95 % of the continuum background events are suppressed by this selection for the muon mode and for the electron mode, respectively. While this selection keeps approximately 54 % and 62 % of the signal events for the muon mode and for the electron mode respectively.

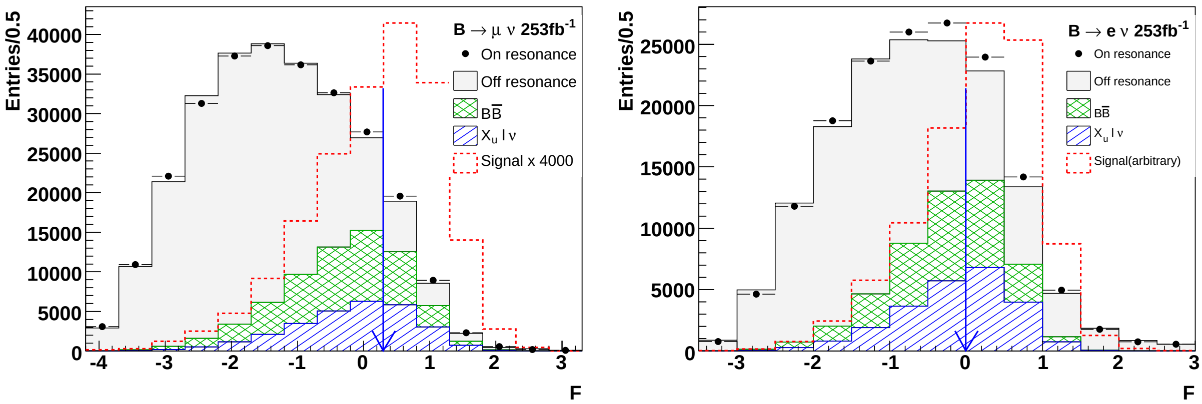


Figure 5.12: SFW distributions for the muon mode (left) and for the electron mode (right).

5.2.5 Neutrino Reconstruction

The neutrino from the signal decay flies away with momentum as large as the signal lepton momentum. Therefore the signal decay has large missing momentum. Then we regard the signal neutrino is the source of the missing momentum and the missing momentum in the laboratory frame is defined by

$$\vec{p}_\nu = \vec{p}^{\text{missing}} = -\sum_i \vec{p}_i \tag{5.8}$$

where i runs over all detected particles in an event. The continuum event has also large missing momentum because a jet-like event likely flies away near the beam axis where particles can not be detected. However the difference of the feature for the missing momentum between the continuum background and the signal decay is useful to discriminate the signal events from the continuum background events. Figure 5.13 shows the distribution for the cosine of the polar angle of the missing momentum. We reject the events with the missing momentum to be close to the beam axis. At last the variables $\cos\theta_{\text{missing}}$ are required to be $\cos\theta_{\text{missing}} < 0.84$ for the muon mode and $\cos\theta_{\text{missing}} < 0.82$ for the electron mode.

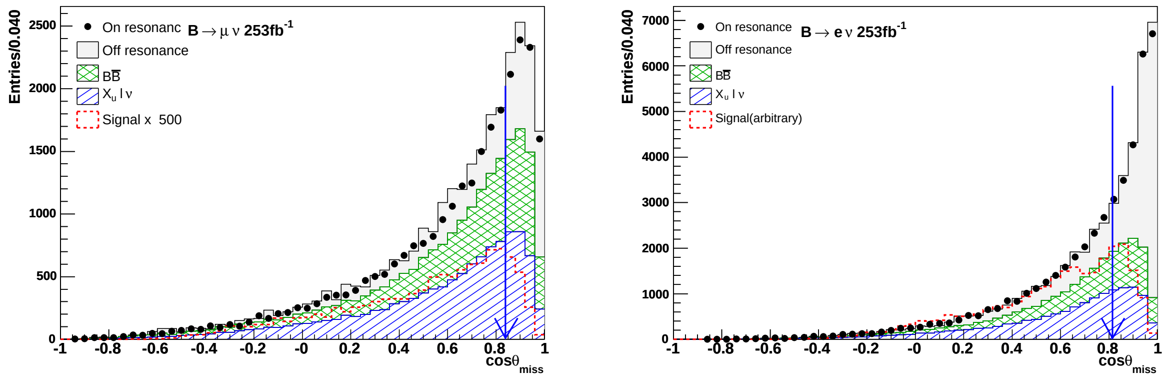


Figure 5.13: Distributions of the cosine of poler angle of the missing momentum in the laboratory frame for the muon mode(left) and the electron mode(right).

The direction of the missing momentum and the signal lepton momentum close to the beam axis have been already restricted. The neutrino from the signal decay must have large transverse momentum. Therefore we require the transverse missing momentum to be greater than $1.75\,\text{GeV}/c$. Figure 5.14 shows the distribution of the transverse missing momentum in the laboratory frame.

5.2.6 Selection Optimization

We optimize the selection criteria. The optimized criteria are determined by the figure-merit of $N_S/\sqrt{N_S + N_B}$ vs ϵ_{signal}, where N_S and N_B represent the number of events for signal MC and background MC in the signal region, and ϵ_{signal} is the signal efficiency.

The off-resonance data sample is too small, and few events remain in case of tight selection criteria applied. Therefore instead of the off-resonance data samples, $e^+e^- \to q\bar{q}$ $(q = u, d, s, c)$ MC samples, τ-pair MC samples and two-photon process MC $(\gamma\gamma \to \mu^+\mu^-/e^+e^-)$ are used as the continuum background.

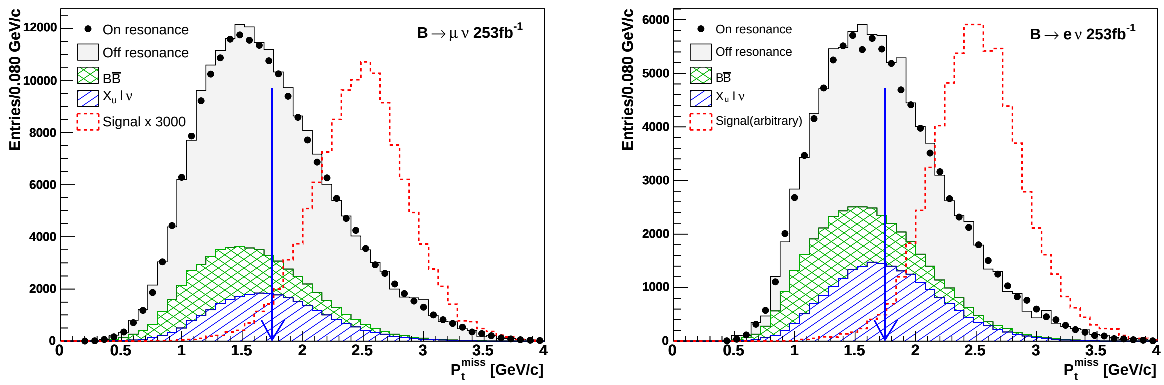


Figure 5.14: Transverse missing momentum distribution for the muon mode (left) and for the electron mode (right). The selection criteria for $\cos\theta_\ell$ and $\cos\theta_{\text{missing}}$ have been already applied to these histograms.

Each criterion of the selection and signal region changes step by step in the reasonable range. Each significance ($N_S\,/\,\sqrt{N_S + N_B}$) and signal efficiency (ϵ_{signal}) by the changed selection criterion are extracted every steps. Then we get a two-dimensional scatter plot of $N_S\,/\,\sqrt{N_S + N_B}$ vs ϵ_{signal} as Figure 5.15.

Then we set the criteria so as to get the best significance and signal efficiency.

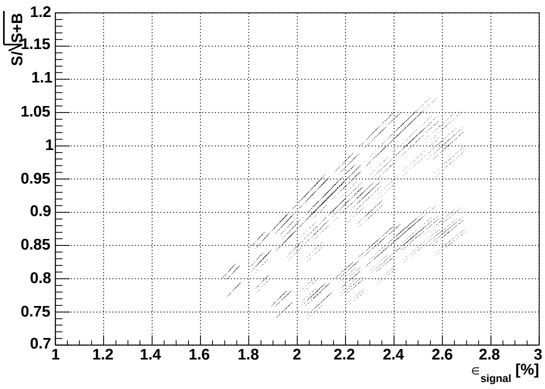


Figure 5.15: Scatter plot of the significance ($N_S\,/\,\sqrt{N_S + N_B}$) vs the signal efficiency

We can get accumulated efficiencies for every event selection step as Table 5.3 and 5.4.

After all selection criteria have been applied, the signal efficiencies in the signal region are respectively given by:

$$\epsilon_\mu^{\text{sig}} = 2.18 \pm 0.06\,\% \text{ (muon mode)} \tag{5.9}$$

$$\epsilon_e^{\text{sig}} = 2.39 \pm 0.06\,\% \text{ (electron mode).} \tag{5.10}$$

Event selection criteria	signal	$B\bar{B}$	$X_u\ell\nu$	Off resonance
pre-selection	43.9	1.99×10^{-2}	1.72	3.93×10^{-2}
$-0.5 < cos\theta_\mu < 0.85$	35.4	1.47×10^{-2}	1.35	2.41×10^{-2}
$F > 0.3$	17.7	3.57×10^{-3}	4.60×10^{-1}	1.31×10^{-3}
K prob. < 0.25	16.9	3.35×10^{-3}	4.41×10^{-1}	9.69×10^{-4}
$\cos\theta_{\mathrm{miss}} < 0.84$	14.6	2.34×10^{-3}	3.37×10^{-1}	5.61×10^{-4}
$p_T^{\mathrm{miss}} > 1.75\,\mathrm{GeV}/c$	14.1	7.56×10^{-4}	1.79×10^{-1}	2.88×10^{-4}
$2.6 < p_\mu^B < 2.84\,\mathrm{GeV}/c$	9.02	1.05×10^{-6}	1.42×10^{-3}	7.06×10^{-5}
fit region	3.15	4.16×10^{-7}	2.78×10^{-4}	5.43×10^{-6}
signal region	2.18	$< 2.11 \times 10^{-7}$	1.13×10^{-4}	$< 1.08 \times 10^{-6}$

Table 5.3: Accumulated efficiencies for the muon decay mode [%]. We applied from the top to the bottom of the event selection criteira in turn. “fit region” and “signal region”, which are defined in section 5.2.2, show the efficiencies in the regions after all selection criteria have been applied.

Event selection criteria	signal	$B\bar{B}$	$X_u\ell\nu$	Off resonance
pre-selection	43.1	1.35×10^{-2}	1.45	3.96×10^{-2}
$-0.5 < \cos\theta_e < 0.8$	33.6	9.90×10^{-3}	1.14	1.25×10^{-2}
$F > 0$	20.9	3.81×10^{-3}	5.50×10^{-1}	2.68×10^{-3}
$\cos\theta_{\mathrm{miss}} < 0.84$	16.5	2.34×10^{-3}	3.85×10^{-1}	9.52×10^{-4}
$p_T^{\mathrm{miss}} > 1.75\,\mathrm{GeV}/c$	16.1	7.45×10^{-4}	2.09×10^{-1}	5.14×10^{-4}
$2.6 < p_e^B < 2.8\,\mathrm{GeV}/c$	8.51	4.22×10^{-7}	1.32×10^{-3}	1.36×10^{-4}
fit region	3.86	$< 2.11 \times 10^{-7}$	3.81×10^{-4}	1.30×10^{-5}
signal region	2.39	$< 2.11 \times 10^{-7}$	1.44×10^{-4}	$< 1.09 \times 10^{-6}$

Table 5.4: Accumulated efficiencies for the electron decay mode [%]. We applied from the top to the bottom of the event selection criteria in turn. “fit region” and “signal region”, which are defined in section 5.2.2, show the efficiencies in the regions after all selection criteria have been applied.

The signal efficiencies in the fit region are given by:

$$\epsilon_{\mu}^{\text{fit}} = 3.15 \pm 0.07\,\% \quad \text{(muon mode)} \tag{5.11}$$
$$\epsilon_{e}^{\text{fit}} = 3.86 \pm 0.08\,\% \quad \text{(electron mode).} \tag{5.12}$$

5.3 Signal Extraction

All the selection criteria have been applied to the on-resonance data sample. Then we can find 12 events for the muon mode and 15 events for the electron mode in the signal region as Figure 5.16.

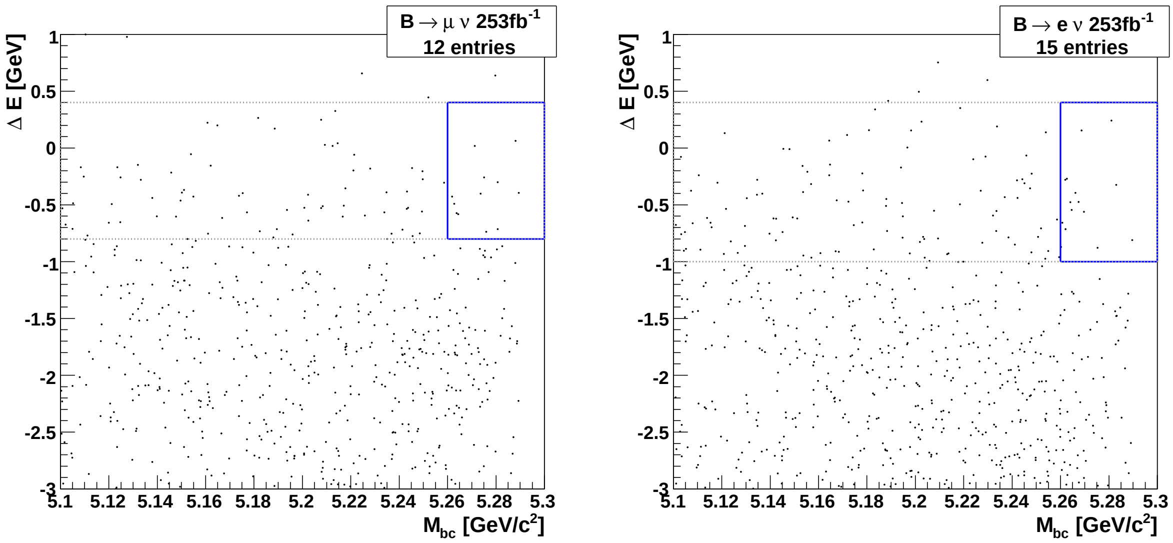


Figure 5.16: $M_{\text{bc}} - \Delta E$ scatter plots for the on-resonance data for the muon mode (left) and for the electron mode (right). Boxes are the signal regions and dotted boxes are the fit regions.

Figure 5.17 shows the M_{bc} distributions which are projections of events in the fit region to the M_{bc} axis.

We try to extract the signal yields in the signal regions by using an unbinned maximum likelihood fit on the fit region. In order to fit the on-resonance data plots, two kinds of functions must be determined as two probability density functions (PDFs) for the signal shape and for the background shape. They are explained in the next section.

5.3.1 Probability Density Functions

The probability distribution functions (PDFs) are defined as the following procedure:

1. M_{bc} distribution of the signal MC event sample with all the event selection cuts is fitted by Crystal Ball function(Figure 5.18)

2. The MC and the off-resonance data samples with looser selection criteria are fitted by ARGUS function(Figure 5.19)

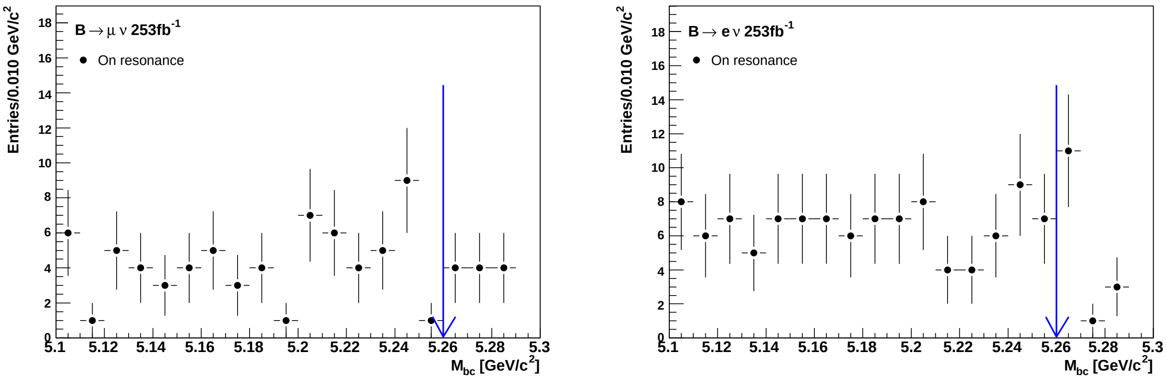


Figure 5.17: M_{bc} distributions in the fit region for the muon mode (left) and for the electron (right) after all selection criteria have been applied. The arrows show the signal region.

Signal Shape

As we described above, we assume that the M_{bc} distribution for the signal events is described by the Crystal Ball function. It consists of a Gaussian core portion and an exponential low-end tail, below a certain threshold. The Crystal Ball function is given by:

$$f(M_{\mathrm{bc}};\alpha,n,\bar{M}_{\mathrm{bc}},\sigma) = N\cdot\begin{cases}\exp\left(-\frac{(M_{\mathrm{bc}}-\bar{M}_{\mathrm{bc}})^2}{2\sigma^2}\right), & \text{for } \frac{M_{\mathrm{bc}}-\bar{M}_{\mathrm{bc}}}{\sigma} > \alpha\\ A\cdot\left(B-\frac{M_{\mathrm{bc}}-\bar{M}_{\mathrm{bc}}}{\sigma}\right), & \text{for } \frac{M_{\mathrm{bc}}-\bar{M}_{\mathrm{bc}}}{\sigma} \le \alpha,\end{cases} \tag{5.13}$$

where

$$A = \left(\frac{n}{|\alpha|}\right)^n\cdot\exp\left(-\frac{|\alpha|^2}{2}\right), \tag{5.14}$$

$$B = \frac{n}{|\alpha|} - |\alpha|. \tag{5.15}$$

N is a normalization parameter and $\bar{M}_{\mathrm{bc}}, \sigma, \alpha$ and n are shape parameters which are fitted with the data. Then the M_{bc} distributions for the signal MC events, which have been made by applying all the selection criteria except for the criterion the M_{bc}, are fitted by the Equation 5.13. Figure 5.18 shows the fitted M_{bc} distributions for the signal MC. All parameters are given by Table 5.5. Figure 5.18 shows the signal shape for each mode. In order to define the signal PDFs, the integration values of these shapes are normalized to be 1. The fitting parameters can be found in the Table 5.5.

Background Shape

We assume that the M_{bc} distribution of background conforms to ARGUS function. The ARGUS function is given by:

$$\frac{dN}{dM_{\mathrm{bc}}} = C\times M_{\mathrm{bc}}\sqrt{1-\frac{M_{\mathrm{bc}}^2}{E_{\mathrm{beam}}^2}}\times\exp\left[-\zeta\left(1-\frac{M_{\mathrm{bc}}^2}{E_{\mathrm{beam}}^2}\right)\right] \tag{5.16}$$

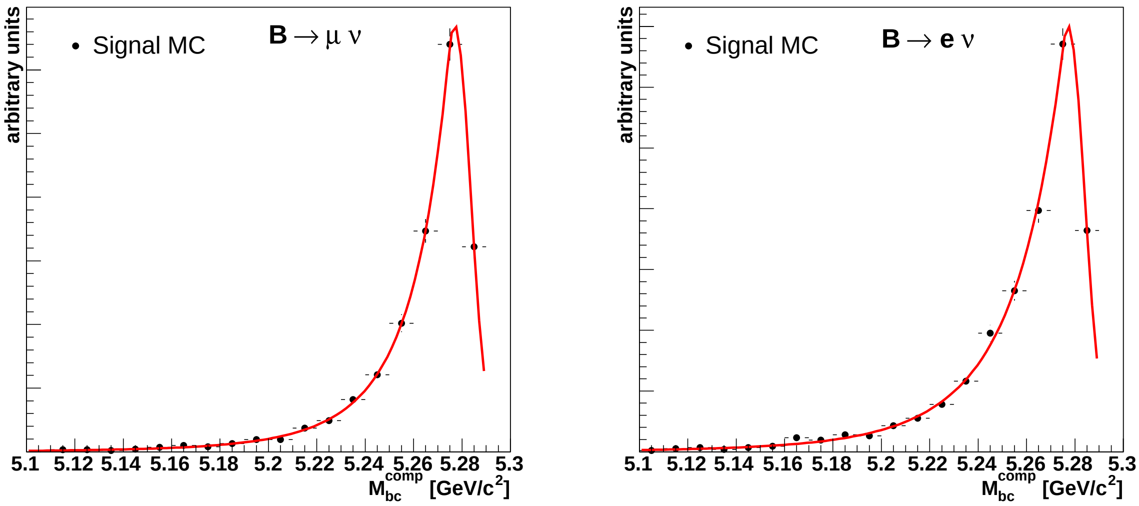


Figure 5.18: Signal PDF distributions for the muon mode (left) and for the electron mode (right). The points show the signal MC in the fit region after all the selection criteria have been applied. The solid lines show the signal PDF by fitting the signal MC by the Crystal Ball function.

parameters	muon mode	electron mode
$\bar{M}_{bc}$	5.2770 ± 0.0005	5.2773 ± 0.0002
σ	$(6.59 \pm 0.58) \times 10^{-3}$	$(6.77 \pm 0.24) \times 10^{-2}$
n	5.23 ± 0.52	6.50 ± 0.72
α	$(4.27 \pm 0.30) \times 10^{-1}$	$(3.33 \pm 0.12) \times 10^{-1}$

Table 5.5: Parameters determined by the fit in the Crystal Ball function for the muon mode and the electron mode.

where C is a normalization parameter and ζ is a shape parameter. Besides $E_{\rm beam}$ is a half of beam collision energy and is fixed at 5.29 GeV.

Because the MC samples and the off-resonance data sample do not have enough statistics after all selection criteria have been applied, we apply looser selection criteria to the MC samples and the off-resonance data sample to be fitted by the ARGUS function. The looser selection criteria are defined in Table 5.6.

$-0.6 < cos\theta_\ell < 0.86$
$SFW > -0.8$
$cos\theta_{\rm miss} < 0.95$
$p_T^{\rm miss} > 1.2\,GeV/c$
$2.5 < P_\ell^B < 3.1\,GeV/c$

Table 5.6: Looser selection criteria.

The looser selection may not exactly represent the distribution of $M_{\rm bc}$ in case that all the event selection criteria are applied, but the looser selection is defined in order that each background component has the almost same ratio as the component after all the event selection criteria have been applied. A comparison of the background ratios between the looser selection criteria and the standard event selection criteria is shown in Table 5.7. Using a lot of the B decays MC events, we estimate the contamination of the peaking background is a few percent of all background events.

Mode	BG components	Looser selection [%]	Standard selection [%]
Muon	$B\bar{B}$	0.8	2.3
	$X_u\ell\nu$	3.6	7.3
	off-resonance	95.6	90.4
Electron	$B\bar{B}$	0.3	0.0
	$X_u\ell\nu$	6.4	4.4
	off-resonance	93.2	95.6

Table 5.7: The background components' ratio.

Figure 5.19 shows background shapes by fitting the off-resonance data and the background MC samples by the ARGUS function. In order to define the background PDFs, the integration values of these shapes are normalized to be 1.

Then we get the background PDF for the $M_{\rm bc}$ distribution in the fit region with the shape parameter(ζ) 13.9 ± 2.4 (18.2 ± 3.4) for the muon (electron) mode.

Background Estimation

We have already mentioned that the background $M_{\rm bc}$ shape in the fit region are modeled on the ARGUS function in the previous section. In this section we complement the explanation of the background components and estimate the number of background events in the signal region.

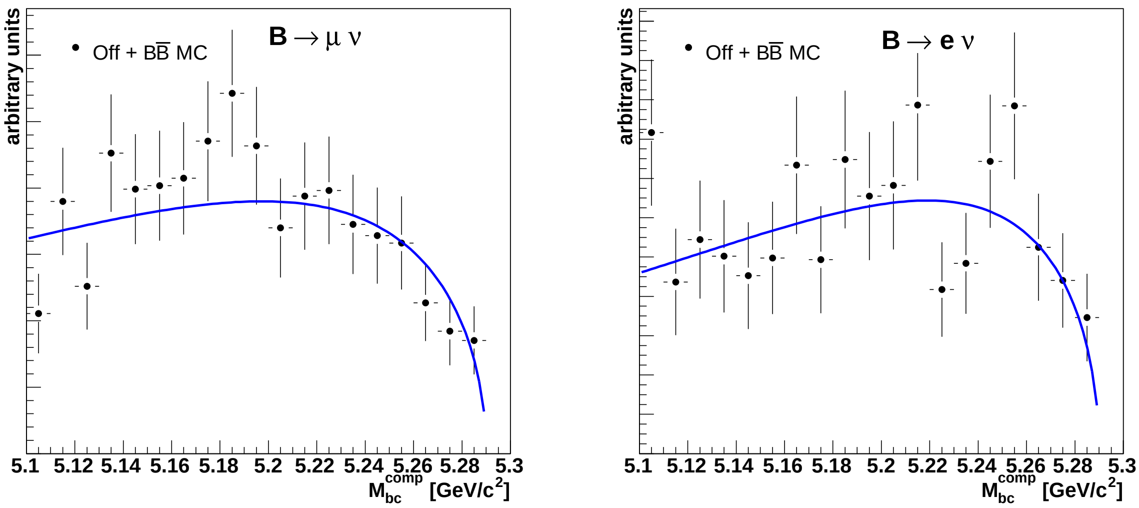


Figure 5.19: The background PDF for the muon mode (left) and for the electron mode (right).The plots show the combined histograms for the background MC and off-resonance data.

The $B \to X_u \ell \nu_\ell$ background tends to make a peak in the $M_{\rm bc}$ signal region because of the decay of a B meson. The peak has a larger width than the signal events. Figure 5.20 shows the $M_{\rm bc}$ distribution for the $B \to X_u \ell \nu_\ell$ background in the fit region after all the event selection criteria have been applied. Then the events in the signal region are polluted with 2.9 ± 0.9 (3.6 ± 1.0) events of $B \to X_u \ell \nu_\ell$ background for the muon (electron) mode, respectively.

In order to study the pollution of signal region by the peaking background, we investigate the $M_{\rm bc}$ distribution for the on-resonance data in the ΔE sideband region which are defined by $\Delta E < -1.0\,\mathrm{GeV}$ for the muon mode and $\Delta E < -1.2\,\mathrm{GeV}$ for the electron mode. Figure 5.21 shows the $M_{\rm bc}$ distributions for the on-resonance data in the ΔE sideband region and the curves fitted by the ARGUS function to the data points. These distributions indicate that the peaking background events are negligible to model the background shape on the ARGUS function.

After all, we can estimate the number of the background events in the signal region by assuming the $M_{\rm bc}$ background shape is described by the background PDF and using the on-resonance data in the sideband region. On the other hand, the background PDF is extrapolated through the signal region as follows:

$$N_{\rm BG} = N_{\rm sideband} \times \frac{\int_{5.26}^{5.29} A(M_{\rm bc}) dM_{\rm bc}}{\int_{5.10}^{5.24} A(M_{\rm bc}) dM_{\rm bc}}, \tag{5.17}$$

where $N_{\rm BG}$ is the number of expected background events, $N_{\rm sideband}$ is the number of events in the on-resonance data in the $M_{\rm bc}$ sideband and $A(M_{\rm bc})$ shows the background PDF normalized by the $N_{\rm sideband}$. Figure 5.22 shows the results of the extrapolation of the background shape in the signal region.

Then we can estimate the number of the background events are 7.4 ± 1.0 events for the muon mode and 13.4 ± 1.4 events for the electron mode. These expected background events consist of

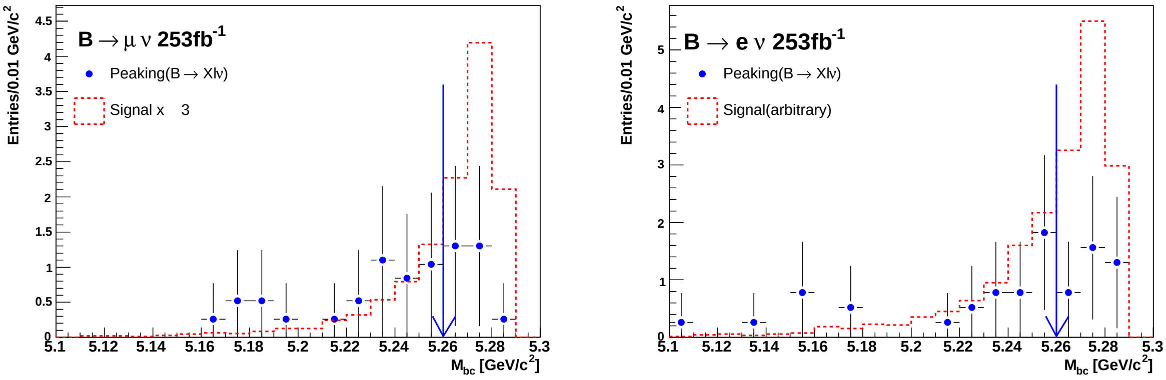


Figure 5.20: $M_{\rm bc}$ distribution after all selection criteria have been applied. Dots show the MC samples of the B meson decay in $B \to X_u \ell \nu_\ell$ background. Dashed lines are the signal MC. Arrows indicate the edge of the signal region in $M_{\rm bc}$.

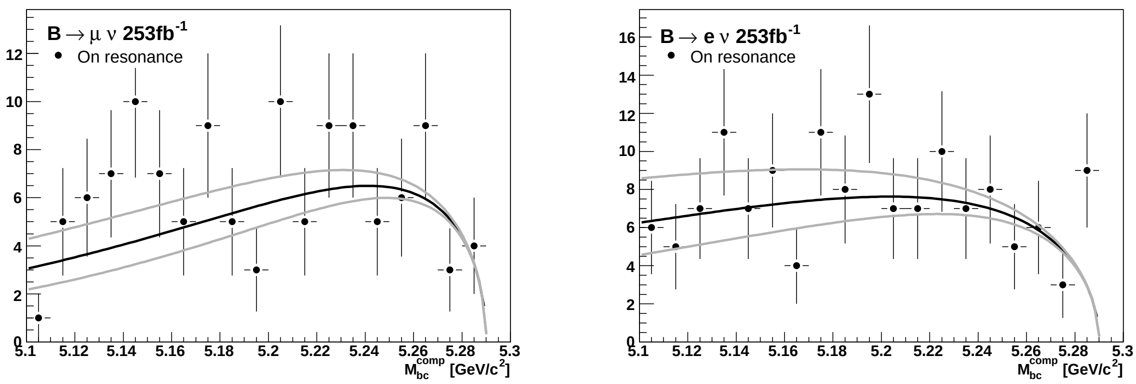


Figure 5.21: $M_{\rm bc}$ distribution for the on-resonance data in the ΔE sideband region after all selection criteria have been applied. The solid curves show the results fitted by the ARGUS function and the light curves show the curves of the ARGUS function with the 1-σ errors of the shape parameters.

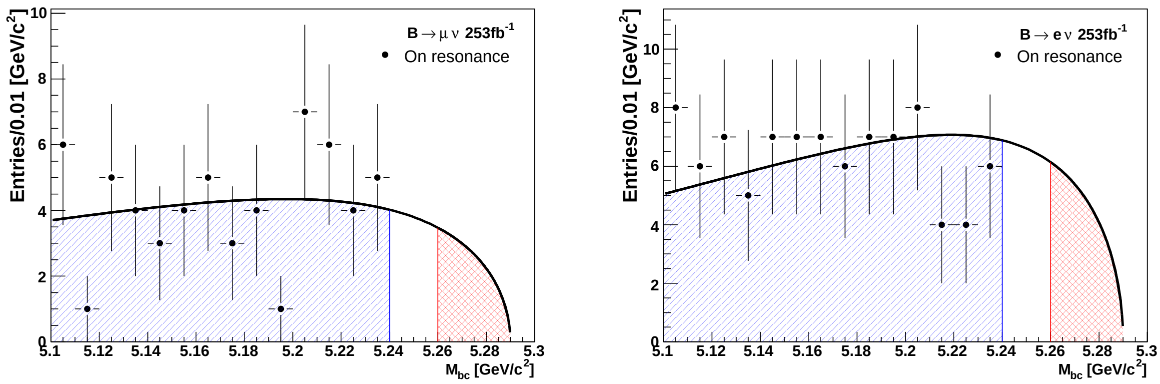


Figure 5.22: Fitted $M_{\rm bc}$ distribution in the $M_{\rm bc}$ sideband region by the ARGUS function as the background PDF. Dots show the on-resonance data. Hatched region is the $M_{\rm bc}$ sideband region and cross-hatched region is the signal region.

approximately 24 % of the peaking background in which $B \to X_u \ell \nu$ is dominant and 76 % of the continuum background.

5.3.2 Likelihood Function

In order to fit the on-resonance data sample by an unbinned maximum likelihood, the likelihood function $\mathcal{L}(n_s)$ for the number of the signal is defined by

$$\mathcal{L}(n_s) = \frac{e^{-(n_s+n_b)}}{N!} \prod_{i=1}^{N} (n_s f_s(i) + n_b f_b(i)) \tag{5.18}$$

where n_b and n_s represent the number of background and signal events, N is the number of observed events, N_b is the number of expected background, The extracted PDFs are used for this unbinned maximum likelihood fit. The negative log likelihood function is minimized using MINUIT with respect to n_b for each n_s which is equal to $\epsilon \times N_{B\bar{B}} \times \mathcal{B}(B \to \ell\nu)$.

5.3.3 Unbinned Maximum Likelihood Fit

The $M_{\rm bc}$ distributions for the on-resonance data with all selection criteria shown in Figure 5.17 are fitted by using the signal and background PDFs and likelihood function. Figure 5.23 shows likelihood distributions of the signal yield.

The signal yield extracted from the fit is 4.1 ± 3.1 events for the muon mode and -1.8 ± 3.3 events for the electron mode in the signal region. For the SM branching fractions, we expect 2.8 ± 0.2 and $(7.3 \pm 1.4) \times 10^{-5}$ events for the muon mode and the electron mode, respectively. Table 5.8 shows the summary table of the signal yield. Figure 5.24 shows the $M_{\rm bc}$ distributions of the events in the ΔE signal region together with the fit results. The significance of the signal in the muon mode is 1.3σ, which is defined as $\sqrt{2\ln(\mathcal{L}_{\rm max}/\mathcal{L}_0)}$ where $\mathcal{L}_{\rm max}$ is the likelihood value

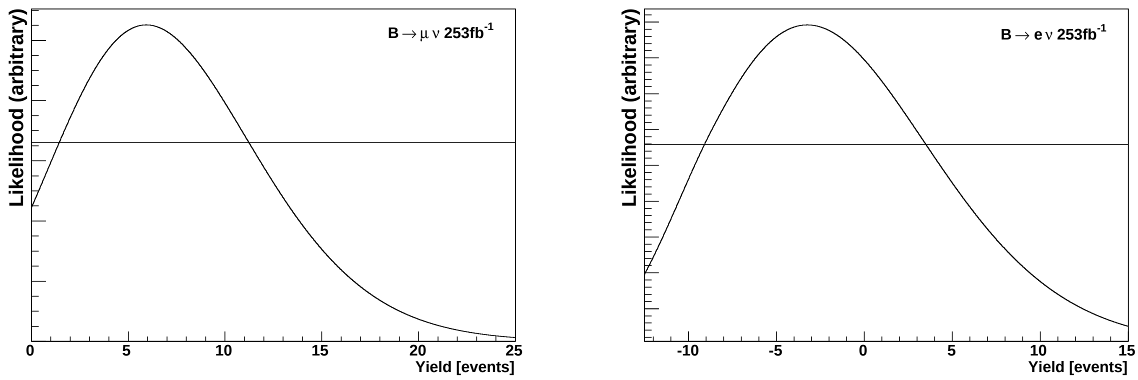


Figure 5.23: Likelihood distributions of the signal yield. The horizontal lines show the one sigma error of the signal yield.

for the best-fit signal yield and $\mathcal{L}_0$ is the likelihood value for no signal event. No excess of events is observed in the electron mode. Therefore we try to set upper limits of the branching fractions for both modes.

	observed	yield (signal region)	SM expected
muon mode	12	4.1 ± 3.1	2.8 ± 0.2
electron mode	15	-1.8 ± 3.3	$(7.3 \pm 1.4) \times 10^{-5}$

Table 5.8: Signal yield summary table.

5.4 Systematic Uncertainties

To set upper limits on the $B^+ \to \ell^+ \nu_\ell$ branching fractions, we have to evaluate systematic uncertainties. Various uncertainties should be considered depending on the methods to extract upper limits. We should consider the following uncertainties.

1. Uncertainty of the number of the signal yield
2. Uncertainty of the $M_{\rm bc}$ distribution shape

5.4.1 Systematic Uncertainty of Signal Yield

The systematic uncertainty of the signal yield depends on that of the signal efficiency($\epsilon_{\rm sig}$) and the number of $B\bar{B}$ events($N_{B\bar{B}}$) because the number of the signal events($n_{\rm s}$) are calculated by using $N_{B\bar{B}}$ and $\epsilon_{\rm sig}$ via $n_{\rm s} = N_{B\bar{B}} \times \mathcal{B}(B \to \ell\nu) \times \epsilon_{\rm sig}$. Then we assume the systematic uncertainty of the signal yield comes from the uncertainties of:

- Number of $B\bar{B}$,

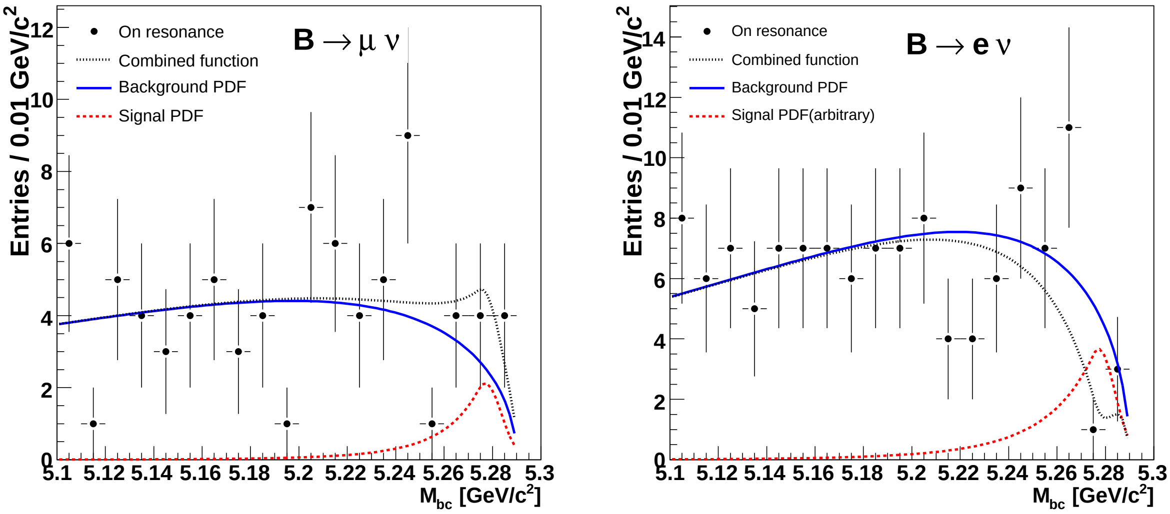


Figure 5.24: $M_{\rm bc}$ distributions for the events in the ΔE signal region, together with the fit results (dotted lines). The solid curves are the background contributions. The dashed curves are the signal contributions. The signal contribution in the electron mode is multiplied by a factor of −4 to make it visible on the plot.

- Tracking for signal candidate lepton,
- Lepton ID,
- Statistics of MC samples and
- Event selections.

The uncertainty of the number of the $B\bar{B}$ obtained are 1.1%, because we have gotten the number of the $B\bar{B}$ events of $(276.6 \pm 3.1) \times 10^6$.

The uncertainty of the tracking efficiency for the signal yield is set to 2.0 %/track by the Belle detector study.

The systematic uncertainties of the identification of the signal muon and signal electron are obtained by the study for the 2-photon processes $\gamma\gamma \to \mu^+\mu^-$ or $\gamma\gamma \to e^+e^-$. The identification uncertainties are determined according to the magnitude and the direction of the lepton momentum in the laboratory frame. Table 5.9 shows the systematic uncertainties dependent on the lepton momentum(p_ℓ) and direction(θ) for the beam axis.

Figure 5.25 shows the $\cos\theta_\ell$ distributions for the signal MC after the other selection criteria have been applied. The systematic uncertainty of the signal lepton identification are calculated after all selection criteria have been applied.

Then the systematic uncertainties of the muon identification for the signal yield and the electron identification are calculated to be 4.4% and 1.1%, respectively.

The uncertainty of the MC statistics for the signal yield is calculated 2.7% for the both modes.

muon mode region	Uncertainty [$\times 100\,\%$]	electron mode region	Uncertainty [$\times 100\,\%$]
$25^\circ \leq \theta < 37^\circ$	$\lvert 0.091 - 0.001 \times p_\ell \rvert$		
$37^\circ \leq \theta < 51^\circ$	$\lvert -0.221 + 0.092 \times p_\ell \rvert$	$35^\circ \leq \theta < 40^\circ$	0.022
$51^\circ \leq \theta < 117^\circ$	0.022	$40^\circ \leq \theta < 60^\circ$	0.015
$117^\circ \leq \theta < 130^\circ$	0.051	$60^\circ \leq \theta < 125^\circ$	0.007

Table 5.9: Systematic uncertainties of the lepton identification dependent on the lepton momentum (p_ℓ [GeV$/c$]) and the direction (θ) with respect to the beam axis.

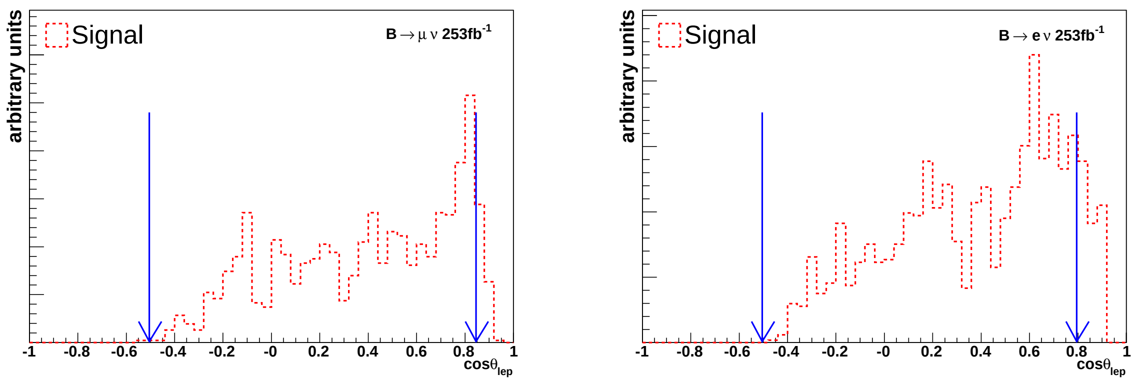


Figure 5.25: The $\cos\theta_\ell$ distribution for the signal MC after the other selection criteria have been applied.

The signal efficiencies are determined by using the signal MC events after all selection criteria have been applied. In order to evaluate the systematic uncertainties for the event shape difference between the on-resonance data and the combined samples of the $B\bar{B}$ MC events and the off-resonance data, we study the mode of $B^{\pm} \to D^{(*)0}\pi^{\pm}$ with fully reconstructed B mesons. The decay mode is also two-body decay so that it is topologically very similar to our signal decays. We regard the pion as the signal lepton and the $D^{(*)0}$ as the signal neutrino. The pion momentum spectrum is, however different from the lepton one for the leptonic decay mode because a $D^{(*)0}$ meson is very heavier than a neutrino. The distribution of the pion momentum in the B rest frame is shown in Figure 5.26. Hence the selection for the lepton momentum has not been applied in the comparison.

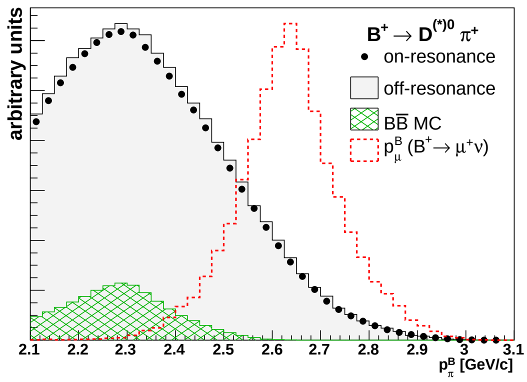


Figure 5.26: Pion momentum spectrum : These histogram show the momentum in the B rest frame. Dashed histogram shows the muon momentum for $B \to \mu\nu$ MC.

We compare the efficiency between the combined sample (MC + off-resonance) and the on-resonance data after the selection criteria for the SFW, the cosine of the polar angle of the signal lepton(pion) momentum, the cosine of the polar angle of the missing momentum and the transverse missing momentum have been applied.

In this signal decay mode $B^{+} \to D^{(*)0}\pi^{+}$, we assume the $M_{\rm bc}$ distribution is parameterized by a Crystal Ball function. The continuum and combinatorial backgrounds are also described by an ARGUS function. Two yields of the signal decay for the on-resonance data sample and the combined sample are extracted by integrating the Crystal Ball functions. The efficiency is defined by

$$\epsilon = \frac{S_{\rm acc}}{S_{\rm acc} + S_{\rm rej}} \tag{5.19}$$

where $S_{\rm acc}$ and $S_{\rm rej}$ are integration values for the accepted and rejected events by the selection criteria, respectively. Figure 5.27 shows the fitted results on the $M_{\rm bc}$ distribution of the rejected and accepted events by the selection criteria for the muon mode in the combined sample. Figure 5.28 shows similar as Figure 5.27 in the on-resonance data sample. Figure 5.29 and 5.30 show also the similar fitted results for the combined sample and for the on-resonance data sample for the electron mode.

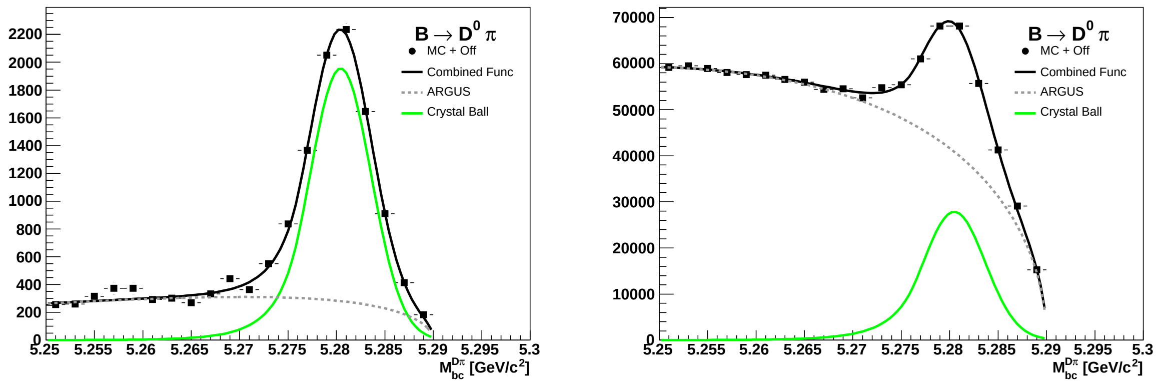


Figure 5.27: $M_{\rm bc}$ distributions of companion B mesons for the combined sample of the MC events and the off-resonance data. Entries of left figure are accepted by the muon mode selection criteria. Entries of right figure are rejected by the muon mode selection criteria. Dotted is the MC + off-resonance combined sample. Solid curves show the fitted results by the combined function of the ARGUS function and the Crystal Ball function. Dashed curves show the background component (ARGUS part). Light curves are the signal component by the Crystal Ball function.

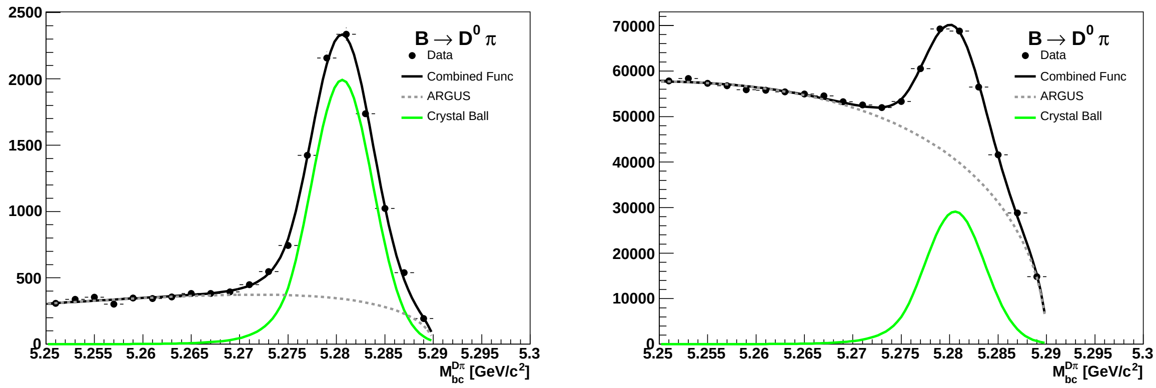


Figure 5.28: $M_{\rm bc}$ distributions of companion B mesons for the on-resonance sample. Dotted shows the on-resonance data sample. The other configurations and styles are the same as Figure 5.27.

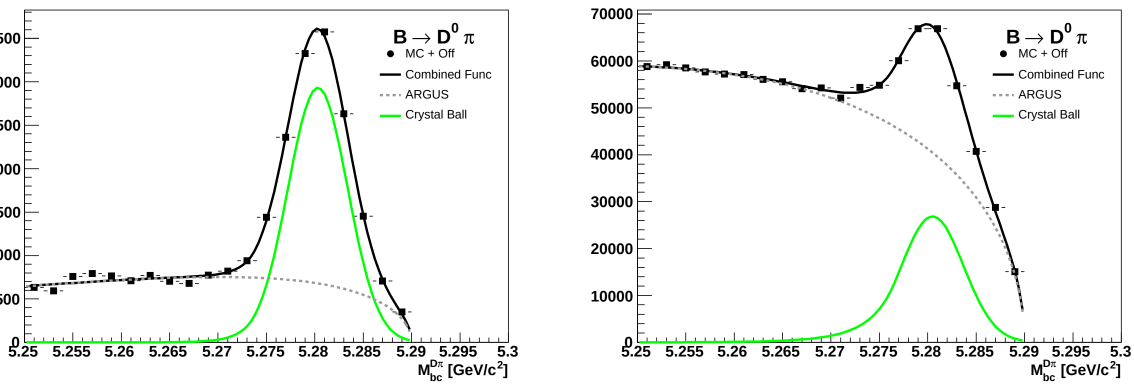


Figure 5.29: M_{bc} distributions of companion B mesons for the combined sample of the MC events and the off-resonance data. sample. The selection criteria for the electron mode are applied. The other configurations and styles are the same as Figure 5.27.

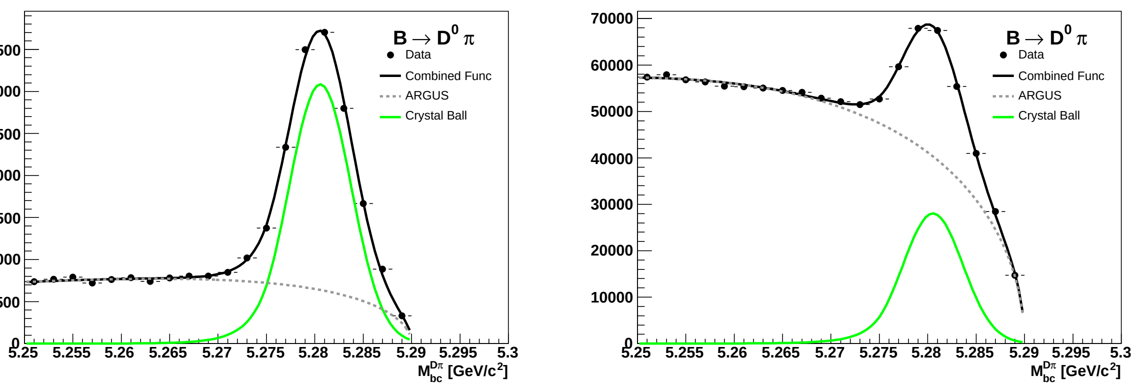


Figure 5.30: M_{bc} distributions of companion B mesons for the on-resonance sample. Dotted shows the on-resonance data sample. The selection criteria for the electron mode are applied. The other configurations and styles are the same as Figure 5.27.

Then the efficiencies are calculated and summarized in Table 5.10.

Mode	$\epsilon_{\rm on}[\%]$	$\epsilon_{\rm MC+off}[\%]$	$\epsilon_{\rm on}/\epsilon_{\rm MC+off}$
Muon	6.53 ± 0.11	6.48 ± 0.21	0.989 ± 0.036
Electron	10.19 ± 0.15	9.04 ± 0.30	1.127 ± 0.041

Table 5.10: Comparison of efficiencies for fully reconstructed $B \to D\pi$ samples

The center value of the efficiency ratio for the muon mode in the Table 5.10 is almost 1. The ratio for the electron mode shows approximately 13 % of the difference of effciency between the on-resonance data sample and the combined samples. The correction for the signal efficiency for the muon mode is negligible because of little difference. The factor 1.127 is assigned as a correction factor for the efficiency of the electron mode. Then the systematic uncertainty are extracted by the ratio between the center value of ratio between efficiencies ($\epsilon_{on}/\epsilon_{MC+off}$) and its error. We get 3.6 % for the muon mode selection and 3.6 % for the electron mode selection. The summary table of the systematic uncertainties for the signal efficiency is show in Table 5.12.

5.4.2 Systematic Uncertainty of $M_{\rm bc}$ Distribution Shape

We have already gotten the PDFs for the signal shape and background shape in Section 5.3 but have not considered the uncertainties for the shapes. In this section, we explain the systematic uncertainties for the shapes. The systematic uncertainty is obtained by recalculating the signal yields by shifting the PDF parameter except for normalization parameter to their own one sigma errors one by one. All parameters' errors except for the normalization parameter and their contributions for the signal yields are shown in Table 5.11. Finally, we get the systematic uncertainties of the PDFs for the muon (electron) mode of 6.5 % (3.2 %) for the signal PDF and 8.1 % (15.7 %) for the background PDF by calculating the square-root of the quadratic sum of each contributions from all the parameters.

5.4.3 Summary of the Systematic Uncertainties

Table 5.12 summarizes the contributions to the systematic uncertainties. The total systematic uncertainty is calculated from the square-root of the quadratic sum of all uncertainties.

5.5 Limits on Branching Fraction

In Section 5.3, we have already mentioned the signal excesses are little or nothing and we will set the upper limits of the branching fractions for the both modes. We have also described the PDF on the $M_{\rm bc}$ distribution in the fit region for the signal and the background events and the likelihood function of the number of the signal yield in Section 5.3. A likelihood function of a

mode	function	parameter	parameter ± error	contribution[%]
muon	Crystal Ball	$\bar{x}$	$(5.28 \pm 2.5) \times 10^{-4}$	0.2
		σ	$(6.59 \pm 0.11) \times 10^{-3}$	5.1
		n	5.23 ± 0.72	1.5
		a	$(4.26 \pm 0.22) \times 10^{-1}$	3.7
		Total		6.5
	ARGUS	shape parameter	13.9 ± 2.4	8.1
electron	Crystal Ball	$\bar{x}$	$(5.28 \pm 2.1) \times 10^{-4}$	0.4
		σ	$(6.50 \pm 0.11) \times 10^{-3}$	2.1
		n	5.23 ± 0.72	1.3
		a	$(3.33 \pm 0.12) \times 10^{-1}$	2.1
		Total		3.2
	ARGUS	shape parameter	18.2 ± 3.4	15.7

Table 5.11: Parameter errors for the signal PDF (Crystal Ball function) and background PDF (ARGUS) and their contributions for the signal yield. Total systematic uncertainty is calculated from the square-root of the quadratic sum of the contributions.

Sources		Muon Mode	Electron Mode
$N_{B\bar{B}}$		1.1%	1.1%
Signal Efficiency	Lepton ID	4.4%	1.1%
	Tracking	1.0%	1.0%
	MC statistics	2.3%	2.1%
	$B^+ \to D^0\pi^+$	3.6%	3.6%
$M_{\rm bc}$ Shape	Signal	6.5%	3.2%
	Background	8.1%	15.7%
Total		12.2%	16.7%

Table 5.12: Summary of systematic uncertainties

branching fraction ($\mathcal{B}$) is proportional to the number of the signal events (n_s) as

$$\mathcal{L}(\mathcal{B}) \propto \mathcal{L}(n_s), \tag{5.20}$$

where $\mathcal{L}(n_s)$ have been defined by Equation 5.18. The 90 % C.L. upper limit on the branching fraction is calculated by

$$0.9 = \frac{\int_0^{\mathcal{B}'_{90}} \mathcal{L}(\mathcal{B}')d\mathcal{B}'}{\int_0^{\infty} \mathcal{L}(\mathcal{B}')d\mathcal{B}'} \tag{5.21}$$

where $\mathcal{B}'$ is defined with random Gaussian number(g) by

$$\mathcal{B}' = \mathcal{B} + g\Delta\mathcal{B} \tag{5.22}$$

where $\Delta\mathcal{B}$ is smeared by systematic uncertainties. Then Figure 5.31 shows likelihood distributions for the branching fraction without and with the inclusion of the systematic uncertainties, respectively.

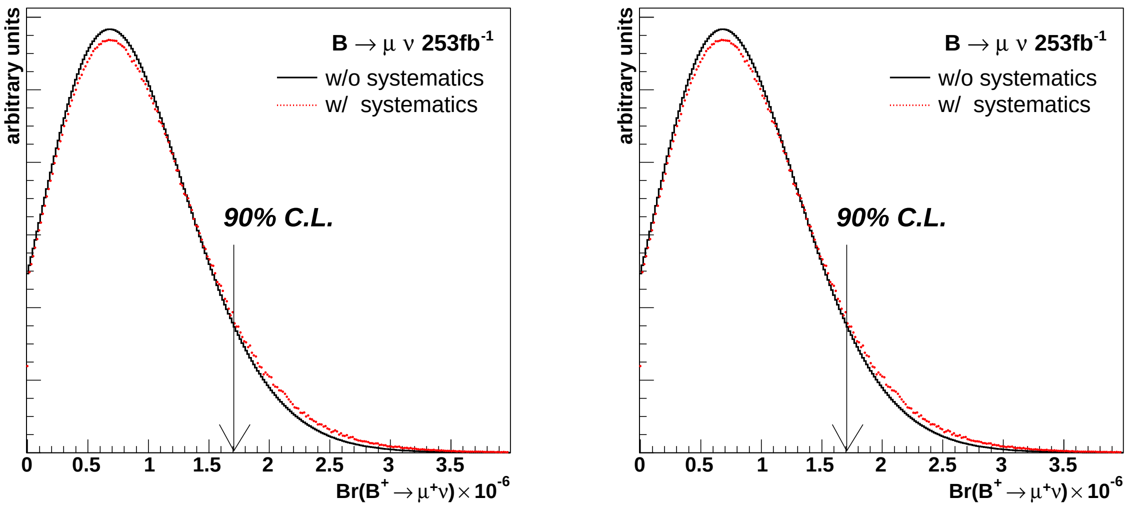
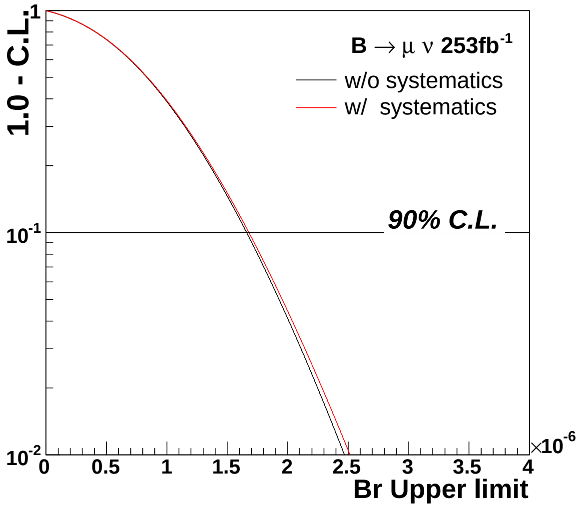


Figure 5.31: Likelihood function dependence on the branching fractions. The solid and dotted curves represent the likelihood functions without and with the inclusion of systematic uncertainties, respectively. The arrows indicate the upper limits on the branching fractions at 90 % confidence level.

Figure 5.32 shows the plots of the confidence level versus branching fraction. At last, the upper limits of the branching fractions are extracted with the inclusion of the systematic uncertainties as follows:

$$\mathcal{B}(B \to \mu\nu) < 1.7 \times 10^{-6} \text{ at } 90\,\% \text{ C.L.} \tag{5.23}$$

$$\mathcal{B}(B \to e\nu) < 9.8 \times 10^{-7} \text{ at } 90\,\% \text{ C.L.} \tag{5.24}$$



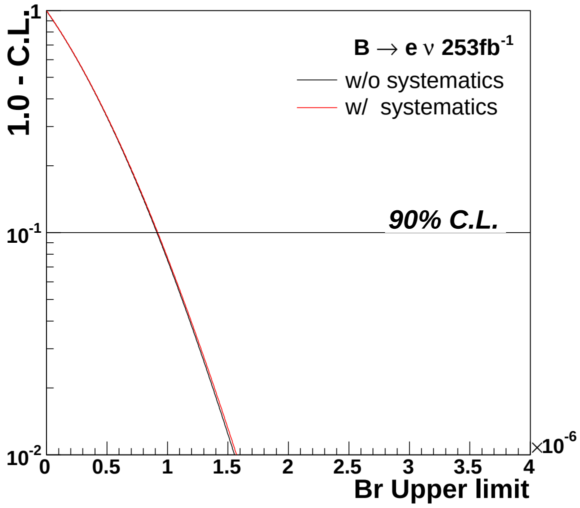


Figure 5.32: Confidence level versus branching fraction distribution for the muon mode (left) and for the electron mode (right).

5.5.1 Expected Sensitivity

The expected sensitivity for the unbinned maximum likelihood fit method to extract the upper limit of the branching fraction is computed by using toy MC studies with a null signal hypothesis. The null signal hypothesis means that the $M_{\rm bc}$ distribution, of the toy samples, conforms closely to the background PDF (parameterized ARGUS function) in the fit region, and the number of events in the toy samples in the fit region is expected to be the number of the background events by the normalized background PDF. The toy MC procedure are listed as follows:

1. Estimation of the number of the background events in the fit region.

   - The number of the background events in the fit region are determined by the same method to estimate the background events in the signal region (see Section 5.3.1).

     Muon Mode 73.0 ± 7.9 events

     Electron Mode 115.6 ± 9.8 events

2. The number of toy MC events in a sample are generated according to the Gaussian distribution with mean value of the background estimation (73.0 for the muon mode and 115.6 for the electron mode) and the σ of the 1σ error of the background estimation (±7.9 for the muon mode and ±9.8 for the electron mode). 3000 sets of the toy MC samples are generated. Figure 5.33 show the distributions of the number of events in the generated toy MC samples.

3. The generated toy MC samples are distributed according to the background PDF.

4. The upper limit of the branching fraction at 90 % confidence level is extracted for every toy MC sample. Figure 5.34 shows the distribution of the extracted upper limit of the branching fraction.

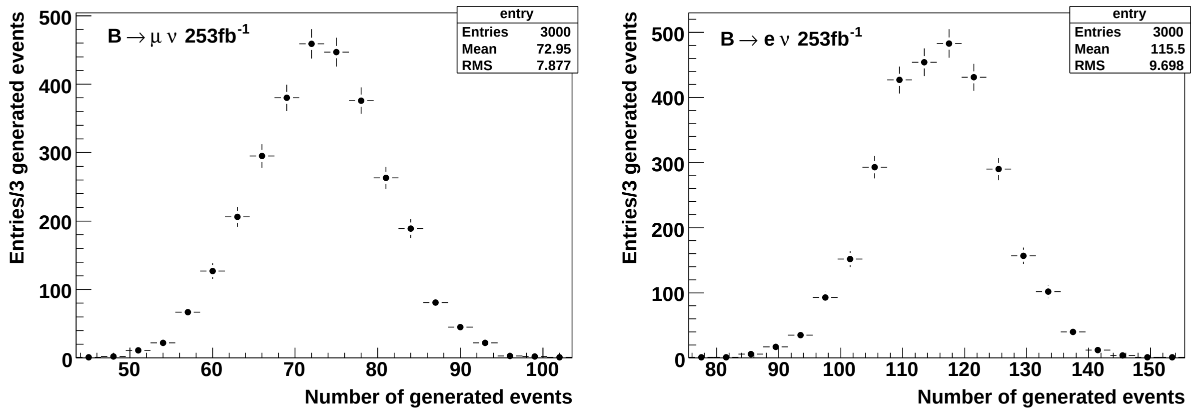


Figure 5.33: Generated number of toy MC samples distributions as a function of events contained in a sample. The means indicate the expected number of the background in the fit region in a sample and the RMS indicate its statistical errors.

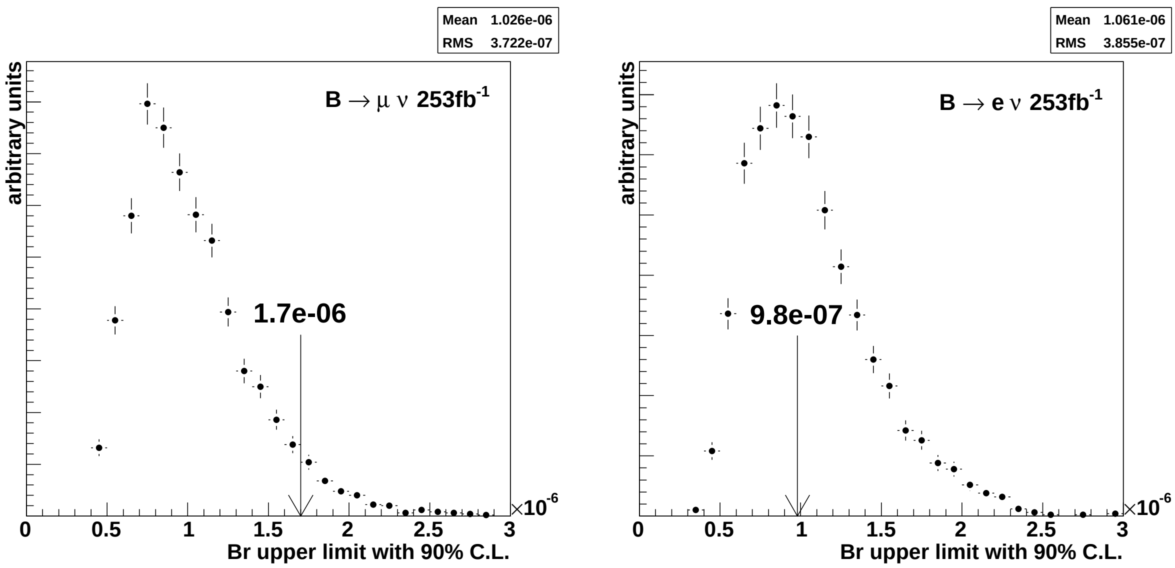


Figure 5.34: Computed upper limits of the branching fraction for the muon mode (left) and the electron mode (right).

Then we get the expected sensitivities for this upper limit extraction method to be 1.0×10^{-6} for the muon mode and 1.1×10^{-6} for the electron mode. The sensitivities are determined to be the average of the distribution for the upper limits of the toy samples. In our analysis case, the probabilities of better limits than our results of the upper limits 1.7×10^{-6} for the muon mode and 9.8×10^{-7} for the electron mode, are 94.3 % for the muon mode and 50.3 % for the electron mode, respectively. Our toy study indicate our result for the muon mode rare case.

Chapter 6

Conclusion and Discussion

6.1 Conclusion

We have searched for the purely leptonic decays $B^+ \to \mu^+\nu_\mu$ and $B^+ \to e^+\nu_e$ using the 253 fb^{-1} data collected by the Belle detector at the KEKB e^+e^- asymmetric-energy collider. We have found no evidence of the signal in either decay mode. We set upper limits on the branching fractions:

$$\mathcal{B}(B^+ \to \mu^+\nu_\mu) < 1.7 \times 10^{-6}, \quad (6.1)$$

$$\mathcal{B}(B^+ \to e^+\nu_e) < 9.8 \times 10^{-7} \quad (6.2)$$

at 90 % confidence level.

In the following section, we discuss the result of the upper limits of the branching fraction with other upper limits.

6.2 Discussion

Obtained upper limits are the most stringent to date and improve the previous published limits [26, 27] by a factor of 4 for $B^+ \to \mu^+\nu_\mu$ and 15 for $B^+ \to e^+\nu_e$. Figure 6.1 shows the upper limits' changes and the branching fraction expected by the SM. Our results are consistent with the SM predictions.

The B meson decay constant, f_B can be only measured by the leptonic B meson decays. However we can not set the branching fractions but the upper limits. Therefore we just introduce the recent result of the f_B by first direct measurement $B^+ \to \tau^+\nu_\tau$ [25]. The f_B has been also set in the paper as follows:

$$f_B = 0.229^{+0.036}_{-0.031}(\mathrm{stat})^{+0.034}_{-0.037}(\mathrm{syst})\ \mathrm{GeV}. \quad (6.3)$$

In our study, the signal yields are extracted as 4.1 ± 3.1 for the muon mode and -1.8 ± 3.3 for the electron mode and the significance of the signal yield for the muon mode is $1.3\,\sigma$. (see Section 5.3). For the muon mode, especially the ratio of the error for the signal yield is approximately 75 %. Because this error of the signal yield dominantly come from data statistics,

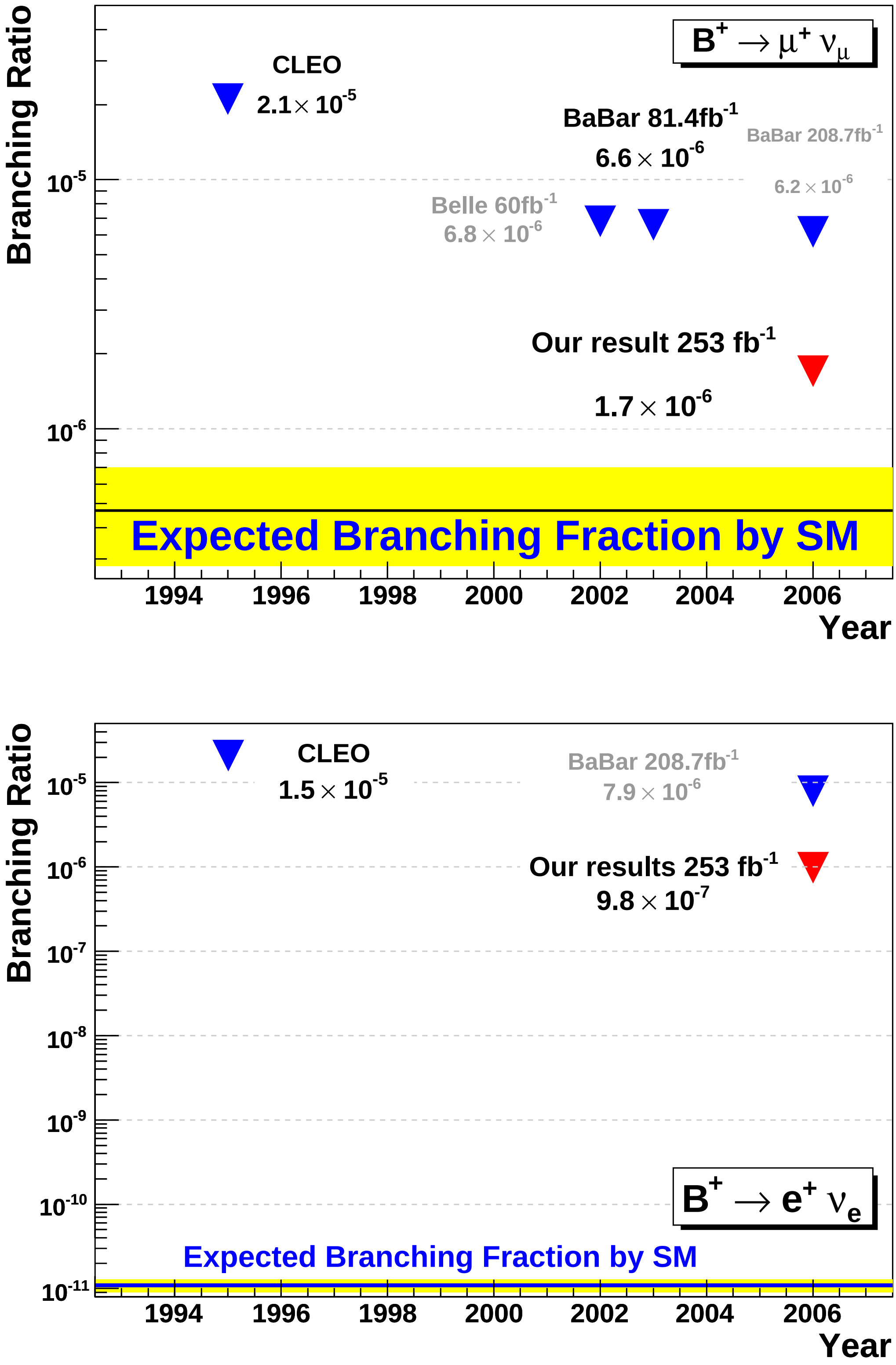


Figure 6.1: Changes of the upper limits of the branching fractions. The thick texts are the upper limits of the branching fractions with publication and the light texts are preliminary results.

if we could analyze $4-5$ times the statistics of the data samples, we might be able to find the evidence of the $B^+ \to \mu^+\nu_\mu$ decay by the same analysis method. However we might be able not to ignore the peaking background ($B \to X\ell\nu$) because of high statistics in the signal region. Therefore we could not discriminate whether the peak in the $M_{\rm bc}$ distribution is the signal and the peaking background. Then the signal regions should be defined in the distribution of the lepton momentum in the B rest frame (p_ℓ^B) where the signal has peak distribution, or we should choice other analysis methods.

Appendix A

Limits on Branching Fraction by Other Methods

We have been set upper limits on the branching fraction by the likelihood function Equation 5.18. To compare the results in Section 5.5 with extractions by other methods, we explain about them in the following sections.

A.1 Additional Systematic Uncertainties

In order to extract the upper limits on the branching fraction by the other methods, we consider additional systematic uncertainties which appear in the other methods. There are systematic uncertainties for the background estimation.

A.1.1 Systematic Uncertainty for Background Estimation

Background estimation is defined by

$$N_{\mathrm{BG}} = N_{\mathrm{sideband}} \times \frac{S_{\text{signal region}}}{S_{\text{sideband}}} \tag{A.1}$$

where N_{sideband} is the number of events in the M_{bc} sideband region and the ΔE signal region for the on-resonance data to which all the selection criteria have been applied, and $S_{\text{signal region}}$ and S_{sideband} indicate the integration values of the M_{bc} distribution in the M_{bc} signal region and the M_{bc} sideband region, respectively.

We consider three sources of the systematic uncertainties from the background estimation. They are indicated as follows:

- The MC + off-resonance data statistics with the looser selection criteria to fit the ARGUS function.
- The on-resonance data statistics in the sideband region with the standard selection criteria.
- Difference for the efficiency between the looser selection criteria and the standard selection criteria.

The first source is the statistics in the $M_{\rm bc}$ distribution for the MC sample and the off-resonance data sample after the looser selection criteria have been applied because $S_{\text{signal region}}/S_{\text{sideband}}$ is determined by the results of fitting to this $M_{\rm bc}$ distribution by the ARGUS function. The shape parameter error and systematic uncertainty are shown in Table A.1.

Mode	shape parameter	$S_{\text{signal region}}/S_{\text{sideband}}$	uncertainty[%]
Muon	13.9 ± 2.4	$0.128^{+0.012}_{-0.011}$	9.4
Electron	18.2 ± 3.4	$0.151^{+0.020}_{-0.018}$	13.2

Table A.1: Uncertainties for the background estimation by statistics of the MC + off-resonance sample with the looser selection criteria.

The second source the uncertainty source comes from a statistical error of N_{sideband}. It is shown in Table A.2.

Mode	Data in $M_{\rm bc}$ side band	uncertainty [%]
Muon	58.0 ± 7.6	13.1
Electron	89.0 ± 9.4	10.6

Table A.2: Uncertainties for the background estimation from the on-resonance data statistics uncertainties.

The third source of the systematic uncertainty comes from the difference of the $M_{\rm bc}$ distributions between with the looser selection criteria and with the standard selection criteria. Because the ΔE signal region is possible to contain the signal (see Figure A.1), the ΔE lower sideband region(see Figure A.1) in $M_{\rm bc} - \Delta E$ plane is used to check the difference.

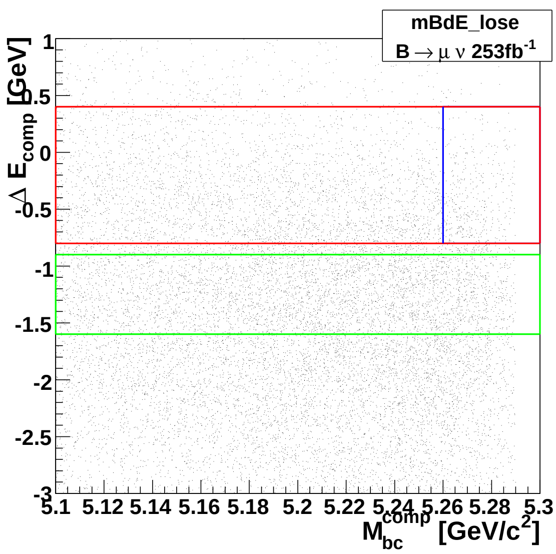


Figure A.1: $M_{\rm bc} - \Delta E$ plane plot for the on-resonance data with the looser selection criteria.

We want to understand the difference between the looser selection and the standard selection for the MC + off-resonance data sample. But the statistics is very poor. Therefore we refer the case of the on-resonance data sample. Firstly we should compare the M_{bc} distribution for the on-resonance data sample with the MC + off-resonance sample when the looser selection criteria have been applied to the all samples. Then it needs to be confirmed that they are almost equivalent. The second check is to compare the M_{bc} distributions between the looser selection and the standard selection. It is done by the on-resonance data sample. The two step are summarized as follows:

1. Comparison of the M_{bc} distributions between the MC + off-resonance samples and the on-resonance data sample with the looser selection (1st check).

2. Comparison of the M_{bc} distributions between the looser selection and the standard selection for the on-resonance data sample (2nd check).

In order to compare the M_{bc} distribution, we compare $S_{\text{signal region}}/S_{\text{sideband}}$ to check the M_{bc} distribution where $S_{\text{signal region}}$ and S_{sideband} represent integration values of the fitted ARGUS function in the M_{bc} signal region and sideband region, respectively. The ratio of the integration values between the signal region and the sideband region is determined by the shape of the fitted ARGUS function. Then the comparison of the 1st check shows in Figure A.2 and Figure A.3.

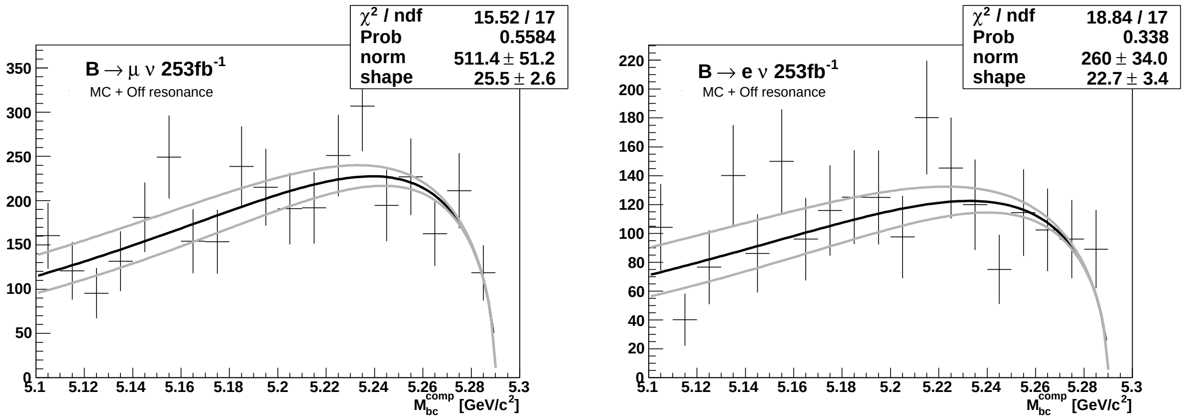


Figure A.2: Fitted M_{bc} distributions for the MC + off-resonance samples with the looser selection criteria of the muon mode (left) and the electron mode (right). Solid line represents the fitting by the ARGUS function. Light lines correspond to errors of the shape parameters.

The comparison table of $S_{\text{signal region}}/S_{\text{sideband}}$ is shown in Table A.3 and A.4 for the muon mode and the electron mode, respectively.

Then the ratios of the $S_{\text{signalregion}}/S_{\text{sideband}}$ between the MC + off-resonance data sample and the on-resonance data sample with the looser selection criteria, respectively:

$$\frac{R_{\text{on}}}{R_{\text{MC+off}}} = 0.990^{+0.102}_{-0.093} \quad \text{Muon mode} \tag{A.2}$$

$$\frac{R_{\text{on}}}{R_{\text{MC+off}}} = 0.896^{+0.115}_{-0.105} \quad \text{Electron mode} \tag{A.3}$$

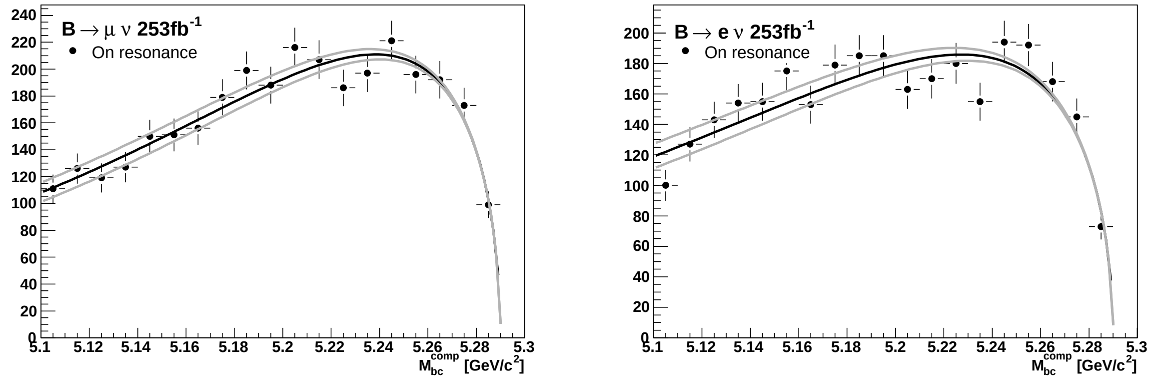


Figure A.3: Fitted M_{bc} distribution for the on-resonance sample with the looser selection criteria. Solid curve represents the fitting by the ARGUS function. Gray curves correspond to errors of the shape parameters.

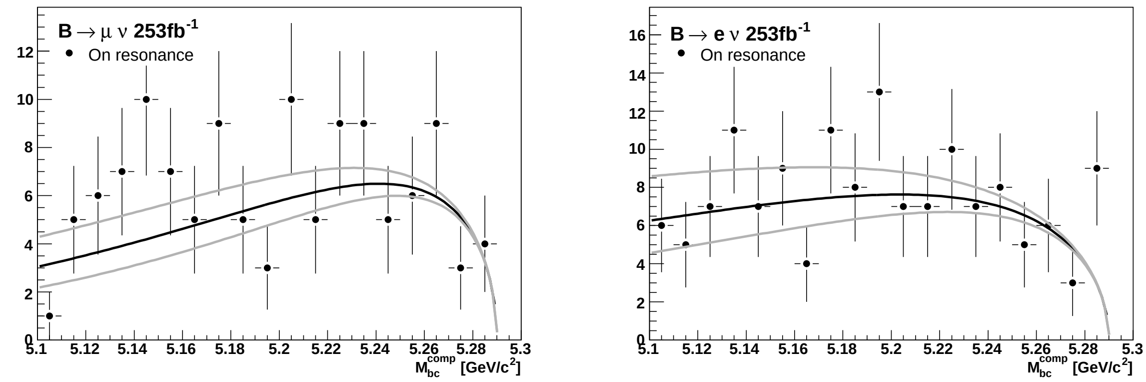


Figure A.4: Fitted M_{bc} distribution for the on-resonance sample with the standard selection criteria. Solid curve represents the fitting by the ARGUS function. Gray curves correspond to errors of the shape parameters.

Data sample	Selection	$S_{\mathrm{signal\ region}}/S_{\mathrm{sideband}}$
MC+off-resonance	Looser	$0.195^{+0.019}_{-0.017}$
On-resonance	Looser	0.193 ± 0.006
On-resonance	Standard	$0.205^{+0.037}_{-0.032}$

Table A.3: Ratio of the integration values of the signal region and the sideband region with the looser and the standard selection criteria for the muon mode.

Data sample	Selection	$S_{\text{signal region}}/S_{\text{sideband}}$
MC+off-resonance	Looser	$0.183^{+0.020}_{-0.018}$
On-resonance	Looser	0.164 ± 0.006
On-resonance	Standard	$0.163^{+0.022}_{-0.020}$

Table A.4: Ratio of the integration values of the signal region and the sideband region with the looser and the standard selection criteria for the electron mode.

where R is defined by $R = \frac{S_{\text{signal region}}}{S_{\text{sideband}}}$. This comparison corresponds to the 1st check and indicates that the M_{bc} distributions of the the MC + off-resonance sample and the on-resonance data are almost same under the looser selection. We confirm the equivalence of the behavior under the looser selection criteria between the MC + off-resonance and the on-resonance sample. Secondly we check the difference between the looser selection and the standard selection by the same method as foregoing.

$$\frac{R_{\text{looser}}}{R_{\text{all}}} = 1.063^{+0.195}_{-0.169} \quad \text{Muon mode} \tag{A.4}$$

$$\frac{R_{\text{looser}}}{R_{\text{all}}} = 0.994^{+0.139}_{-0.127} \quad \text{Electron mode} \tag{A.5}$$

Because the center values are almost equal to 1.0, we regard the background estimation have not to be corrected. The errors are regarded as systematic uncertainties. The systematic uncertainties for the background estimation by using the MC + off-resonance with the looser selections are evaluated and shown in Table A.5.

Mode	Systematic uncertainty[%]
Muon	18.3
Electron	14.0

Table A.5: Uncertainties for the background estimation from the statistics uncertainties of the on-resonance sample.

A.2 Counting Method

We can also extract an upper limit on the branching fraction at 90 % confidence level by the method of Feldman and Cousins by using Pole.f module [51]. This method requires the expected background events, the number of the observed events in the signal region, the signal efficiency and the systematic uncertainties for all of them. The systematic uncertainties of the background estimation is in Section A.1.1. The systematic uncertainties of the signal efficiency are described in Section A.1.1 and shown in the Table 5.12.

The upper limits of the branching fractions are extracted by the counting method considering the systematic uncertainties as follows:

$$\mathcal{B}(B \to \mu\nu) \quad < \quad 2.1 \times 10^{-6} \quad \text{at } 90\,\% \text{ C.L.} \tag{A.6}$$
$$\mathcal{B}(B \to e\nu) \quad < \quad 1.8 \times 10^{-6} \quad \text{at } 90\,\% \text{ C.L.} \tag{A.7}$$

A.3 Unbinned Fit in the Signal Region

Using the constraints on the expected number of the background events in the signal region, the upper limits on the branching fraction at 90 % confidence level can be extracted by an unbinned maximum likelihood fit. The likelihood function, Equation 5.18 is extended to

$$\mathcal{L} = \frac{1}{\sqrt{2\pi}\sigma_b} e^{-(n_b+N_b)^2/2\sigma_b^2} \frac{e^{-(n_s+n_b)}}{N!} \prod_{i=1}^{N} (n_s f_s(i) + n_b f_b(i)), \tag{A.8}$$

where N_b is expected the number of the background events in the signal region. The signal region ($5.26\,\mathrm{GeV}/c^2 < M_{\mathrm{bc}} < 5.29\,\mathrm{GeV}/c^2$) is used for the fitting. The PDFs are the same as the PDFs defined in Section 5.3.1. However the number of the background events is already included in Equation A.8 and we also consider the systematic uncertainty of the background estimation.

The likelihood distribution of the branching fraction is shown in Figure A.5, and the distribution of confidence level vs branching fraction are shown in Figure A.6.

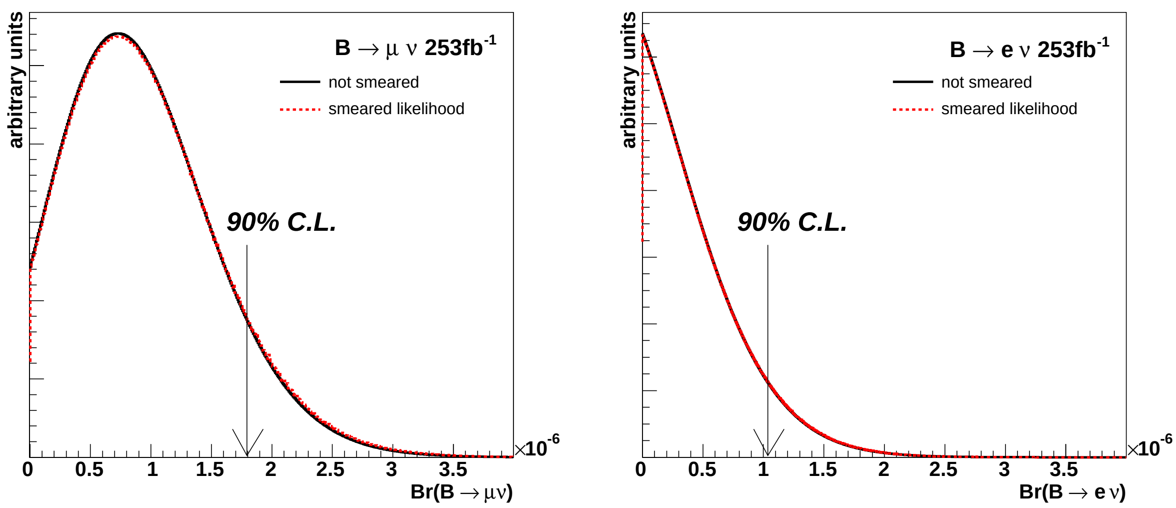


Figure A.5: Likelihood vs branching fraction for the muon mode(left) and the electron mode(right). Dashed curves are likelihood distributions of the branching fraction. The 90% *C.L.* arrows indeicate the branching fraction up to which the integration is 0.9.

The upper limits of the branching fraction are extracted considering the systematic uncertainties as follows:

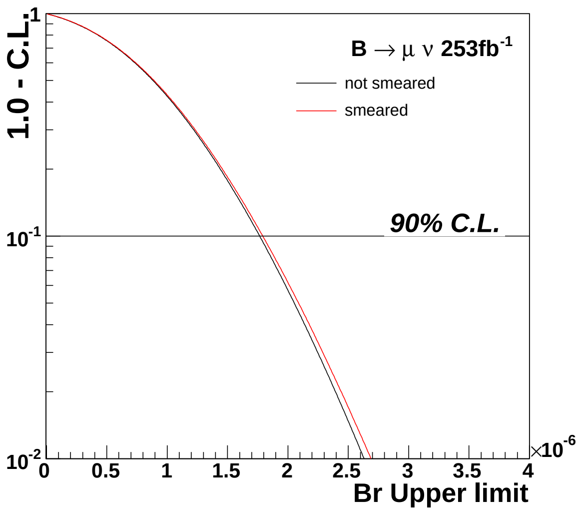


Figure A.6: Confidence level vs upper limit of the branching fraction for the muon mode (left) and the electron mode (right). Solid curves and light curves represent the curves without and with smearing by systematic uncertainty.

$$Br(B \to \mu\nu) \quad < \quad 1.8 \times 10^{-6} \quad \text{at } 90\,\% \text{ C.L.} \tag{A.9}$$

$$Br(B \to e\nu) \quad < \quad 1.0 \times 10^{-6} \quad \text{at } 90\,\% \text{ C.L.} \tag{A.10}$$

A.4 Sensitivities on Upper Limits

We examined three methods to extract the upper limit on the branching fraction. We study the toy MC sample to understand sensitivities for these methods in the same manner as described in Section 5.5.1. We show the method with the best sensitivity by 3000 sets of the toy MC samples.

Counting Method

The toy MC generator for the counting method generate the observed events according to Poisson distribution which has the mean values of the number of the expected background events. Then the 90 % confidence level the upper limit distributions by the results of 3000 trials are shown in Figure A.7. The uncertainties for the counting method are considered in these results. The sensitivities for this method can be evaluated by the mean values of the distributions of the branching fraction at 90 % confindence level as $\mathcal{B}(B^+ \to \mu^+\nu) < 1.1 \times 10^{-6}$ and $\mathcal{B}(B^+ \to e^+\nu) < 1.3 \times 10^{-6}$.

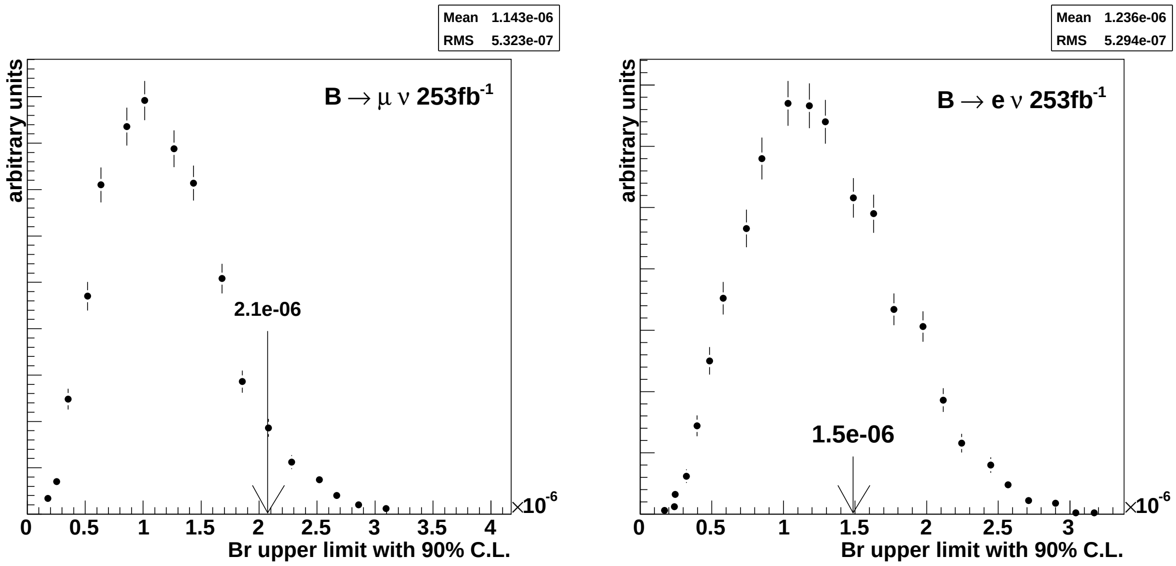


Figure A.7: Toy MC study for the counting method : The trial is 3000 times.

Unbinned Fit in the Signal Region

For this method, events in the fit region ($5.26 < M_{\rm bc} < 5.29\,GeV/c^2$) are generated according to the ARGUS function. Then the number of the generated events are determined according to Poisson distribution which has the mean value of the number of the expected background events. The extraction of the upper limit of the branching fraction is the same method as described in Section 5.5.1. Then 3000 times trials are done for each mode. In all the trials the background PDF is the same ARGUS function which is determined by the off-resonance data and the $B\bar{B}$ MC events with the looser selection. The signal PDF is also the same as described in Section 5.3.1. All considered uncertainties are also same as described in Section 5.3.1.

The distributions of the upper limit of the branching fraction with the 90 % confidence level is shown in the Figure A.8.

The sensitivities for this method can be evaluated by the mean values of the distributions of the branching fraction at 90 % confidence level as $\mathcal{B}(B^+ \to \mu^+\nu) < 1.1 \times 10^{-6}$ and $\mathcal{B}(B^+ \to e^+\nu) < 1.2 \times 10^{-6}$.

Method	Muon mode	Electron mode
Unbinned fit in the fit region	1.0×10^{-6}	1.1×10^{-6}
Unbinned fit in the signal region	1.1×10^{-6}	1.2×10^{-6}
Counting Method	1.1×10^{-6}	1.3×10^{-6}

Table A.6: Sensitivities of the upper-limit-extraction methods.

Table A.6 shows the sensitivity for every method. The method extracting our results is selected among these methods so as to get the best sensitivity.

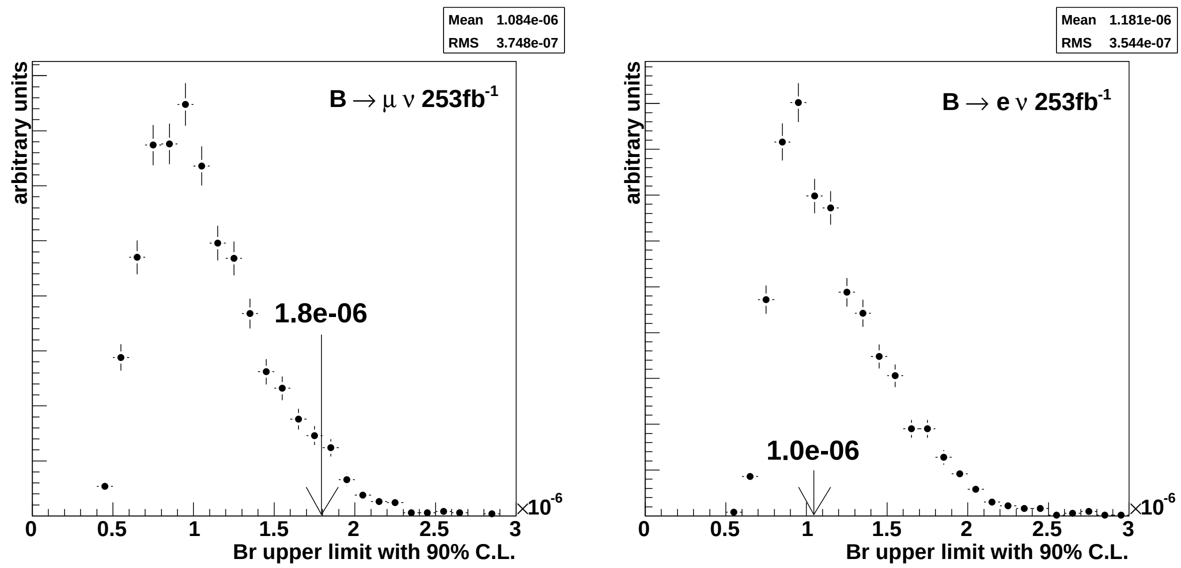


Figure A.8: Distributions of the expected upper limits of the branching fractions for the unbinned maximum likelihood fit in the signal region.

Bibliography

[1] N. Cabibbo, “Unitary Symmetry And Leptonic Decays,” Phys. Rev. Lett. **10**, 531 (1963).

[2] M. Kobayashi and T. Maskawa, “CP Violation In The Renormalizeable Theory Of Weak Interaction,” Prog. Theor. Phys. **49**, 652 (1973).

[3] A. B. Carter and A. I. Sanda, “CP Violation In B Meson Decays,” Phys. Rev. D **23**, 1567 (1981).

[4] Q. H. Kim, W. Y. Pham, “Elementary Particles and Their Interactions: Concepts and Phenomena,” Springer Verlag Berlin Heidelberg. (1998).

[5] L. L. Chau and W. Y. Keung, “Comments On The Parametrization Of The Kobayashi-Maskawa Matrix,” Phys. Rev. Lett. **53**,1802 (1984).

[6] L. Wolfenstein, “Parameterization Of The Kobayashi-Maskawa Matrix,” Phys. Rev. Lett. **51**, 1945 (1983).

[7] S. Eidelman *et al.*, [Particle Data Group], “Review of particle physics,” Phys. Lett. B **592** 1 (2004).

[8] M. Battaglia *et al*., “The CKM matrix and the unitarity triangle,” [arXiv:hep-ex/0304132].

[9] CKMfitter Group (J. Charles *et al.*), Eur. Phys. J. C41, 1-131 (2005) [hep-ph/0406184].

[10] A. Abashian *et al*., (Belle Collaboration), “Measurement of the CP violation parameter $\sin(2\phi_1)$ in B_d^0 meson decays, Phys. Rev. Lett. **86**, 2509 (2001) [arXiv:hep-ex/0102018].

[11] B. Aubert *et al*., (BaBar Collaboration), “Measurement of CP violating asymmetries in B^0 decays to CP eigenstates,” Phys. Rev. Lett. **86**, 2515 (2001) [arXiv:hep-ex/0102030].

[12] A. Gray *et al.* (HPQCD Collaboration), “The B meson decay constant from unquenched lattice QCD,” Phys. Rev. Lett. **95**, 212001 (2005). See also S. Hashimoto, “Recent results from lattice calculations,” Int. J. Mod. Phys. **A 20**, 5133 (2005), which quotes an average value of 0.189 ± 0.027 GeV. QCD sum rules predict $f_B = 210 \pm 19$ MeV, see M. Jamin and B. O. Lange, “f(B) and f(B/s) from QCD sum rules,” Phys. Rev. D **65**, 056005 (2002).

[13] M. Neubert, “QCD Based Interpretation Of The Lepton Spectrum In Inclusive Anti-$B \to X(U)$ Lepton Anti-Neutrino Decays,” Phys. Rev. D **49**, 3392 (1994) [arXiv:hep-ph/9311325].

[14] I. I. Y. Bigi *et al.*, “On The Motion Of Heavy Quarks Inside Hadrons: Universal Distributions And Inclusive Decays,” Int. J. Mod. Phys. A **9**, 2467 (1994) [arXiv:hep-ph/9312311].

[15] C. W. Bauer *et al.*, “A model independent determination of $|V_{ub}|$,” Phys. Lett. B **479**, 395 (2000) [arXiv:hep-ph/0002161].

[16] A. Bornheim *et al.*, (CLEO Collaboration), “Improved measurement of $|V_{ub}|$ with inclusive semileptonic B decays,” Phys. Rev. Lett. **88**, 231803 (2002) [arXiv:hep-ex/0202019].

[17] B. Aubert *et al.*, (BaBar Collaboration), “Measurement of the inclusive electron spectrum in charmless semileptonic B decays near the kinematic endpoint and determination of $|V_{ub}|$,” [arXiv:hep-ex/0408075].

[18] A. Limosani *et al.*, (Belle Collaboration), “Measurement of inclusive charmless semileptonic B-meson decays at the endpoint of the electron momentum spectrum,” Phys. Lett. B **621**, 28 (2005) [arXiv:hep-ex/0504046].

[19] I. Bizjak *et al.*, (Belle Collaboration), “Measurement of the inclusive charmless semileptonic partial branching fraction of B mesons and determination of $|V_{ub}|$ using the full reconstruction tag,” Phys. Rev. Lett. **95**, 241801 (2005) [arXiv:/hep-ex/0504046]; B. Aubert *et al.*, (BaBar Collaboration), “Measurement of the inclusive charmless semileptonic branching ratio of B mesons and determination of $|V_{ub}|$,” Phys. Rev. Lett. **92**, 071802 (2004) [arXiv:hep-ex/0307062].

[20] M. Okamoto *et al.*, “Semileptonic $D \to \pi/K$ and $B \to \pi D$ decays in 2+1 flavor lattice QCD,” Nucl. Phys. Proc. Suppl. **140**, 461 (2005) [arXiv:hep-lat/0409116].

[21] J. Shigemitsu *et al.*, ‘Semileptonic B decays with $N(f) = 2 + 1$ dynamical quarks,” Nucl. Phys. Proc. Suppl. **140**, 464 (2005) [arXiv:hep-lat/0408019].

[22] P. Ball and R. Zwicky, ‘Semileptonic B decays with N(f) = 2+1 dynamical quarks,” Phys. Rev. D **71**, 014015 (2005) [arXiv:hep-ph/0406232].

[23] R. Kowalewski and T. Mannel, “Determination of V_{cb} and V_{ub},” in W. M. Yao *et al.* [Particle Data Group], J. Phys. G **33**, 867 (2006).

[24] M. Battaglia and L. Gibbons, “Determination of $|V_{ub}|$,” in S. Eidelman *et al.*, [Particle Data Group], Phys. Lett. B **592**, 793 (2004).

[25] K. Ikado *et al.* (Belle Collaboration), “Evidence of the purely leptonic decay $B^- \to \tau^- \bar{\nu}_\tau$,” Phys. Rev. Lett. **97**, 251802 (2006).

[26] B. Aubert *et al.* (BaBar Collaboration), "Search for the rare leptonic decay $B^+ \to \mu^+\nu_\mu$," Phys. Rev. Lett. **92**, 221803 (2004).

[27] M. Artuso *el al.* (CLEO Collaboration), "A Search for $B \to$ lepton anti-lepton-neutrino," Phys. Rev. Lett. **75**, 785 (1995).

[28] K. Abe *et al.* (Belle Collaboration), "Search for $B^+ \to \mu^+\nu$ and $B^+ \to \ell^+\nu\gamma$ decays," hep-ex/0408132 (2004).

[29] B. Aubert *et al.* (BaBar Collaboration), "A search for the decays $B^+ \to e^+\nu_e$ and $B^+ \to \mu^+\nu_\mu$ using hadronic-tag reconstruction," hep-ex/0607110 (2006).

[30] W. S. Hou, "Enhanced charged Higgs boson effects in $B^- \to \tau\bar{\nu}_\tau, \mu^-\bar{\nu}_\mu$ and $B \to \tau\nu_t au + X$," Phys. Rev. D **48**, 2342 (1993).

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

[32] A. Abashian *et al.*, (Belle Collaboration), "The Belle detector," Nucl. Instrum. Meth. A **479**, 117 (2002).

[33] H. Aihara, "The Belle silicon vertex detector," Nucl. Instrum. Meth. A **466**, 268 (2001). [arXiv:hep-ex/0011028].

[34] H. Hirano *el at.*, "A high resolution cylindrical drift chamber for the KEK B-factory," Nucl. Instrum. Meth. A **455**, 294 (2000).

[35] T. Sumiyoshi *et al.*, "Silica aerogel Cherenkov counter for the KEK B-factory experiment," Nucl. Instrum. Meth. A **433**, 385 (1999).

[36] T. Iijima *et al.*, "Aerogel Cherenkov counter for the BELLE detector," Nucl. Instrum. Meth. A **453**, 321 (2000).

[37] H. Kichimi *et al.*, "The BELLE TOF system," Nucl. Instrum. Meth. A **453**, 321 (2000).

[38] A. Abashian *et al.*, (Belle Collaboration) "The K(L)/mu detector subsystem for the BELLE experiment at the KEKB factory," Nucl. Instrum. Meth. A **449**, 112 (2000).

[39] The CLEO QQ generator, see http://www.lns.cornell.edu/public/CLEO/soft/QQ/.

[40] D. J. Lange, "The EvtGen particle decay simulation package," Nucl. Instrum. Meth. A **462**, 152 (2001).

[41] T. Sjostrand, "PYTHIA 5.7 and JETSET 7.4: Physics and manual," arXiv:hep-ph/9508391.

[42] T. Sjöstrand, P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna and E. Norrbin, "High-energy-physics event generation with PYTHIA 6.1," Computer Physics Commun. **135** 238 (2001).

[43] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand, "Parton Fragmentation And String Dynamics," Phys. Repot. **97**, 31 (1983).

[44] R. Brun, R. Hagelberg, M. Hanstroul, and J. C. Lasalle, "Geant: Simulation Program For Particle Physics Experiments. User Guide And Reference Manual," CERN-DD-78-2-REV ; R. Brun *et al*., GEANT 3.21, CERN Report DD/EE/84-1 (1984).

[45] Tracking Group (Belle Collaboration), "Charged Particle Tracking in Belle," Belle Note 327 (2000).

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

[47] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima and T. Tsukamoto, "Electron identification in Belle," Nucl. Instrum. Meth. A**485**, 490 (2002) arXiv:hep-ex/0108044.

[48] H. Kakuno, "Measurement of $|V_{ub}|$ through Charmless Semileptonic B-meson Decay," Tokyo Institute of Technology, Doctor Thesis (2003).

[49] 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).

[50] K. Abe *et al.*, (Belle Collaboration), "A measurement of the branching fraction for the inclusive $B \to X_s\gamma$ decays with Belle," Phys. Lett. B **511**, 151 (2001).

[51] J.Conrad *et al.*, "Including systematic uncertainties in confidence interval construction for Poisson statistics," Phys. Rev. D **67**, 012002 (2003).