Search the decay of the B meson to two leptons at the Belle Experiment

Kimberly M. Williams

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Physics

Leo E. Piilonen, Chair

Patrick Huber

Mark L. Pitt

John H. Simonetti

December 9, 2015

Blacksburg, Virginia

Search the decay of the B meson to two leptons at the Belle Experiment

Kimberly M. Williams

(ABSTRACT)

The set of decays known as $B^0 \to \ell^+\ell^-$ are exceedingly rare. This thesis details procedures developed to optimize signal extraction and improve the upper limit on the measured branching fraction. When applied to 85 million $B\bar{B}$ events collected at the Belle experiment in Tsukuba, Japan, the resulting upper limits on branching fractions were $\mathcal{B}(B^0 \to e^+e^-) < 8.8 \times 10^{-8}$, $\mathcal{B}(B^0 \to \mu^+\mu^-) < 1.0 \times 10^{-7}$, and $\mathcal{B}(B^0 \to e^\pm\mu^\mp) < 8.5 \times 10^{-8}$ at the 90% confidence level.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Rare particle decays that are either forbidden within the Standard Model (SM) of particle physics or are expected to have very small branching ratios provide excellent signatures with which to look for New Physics and allow us to probe subatomic processes that are beyond the reach of direct searches. This study focuses on the set of rare decays known collectively as $B^0 \to \ell^+\ell^-$ where ℓ is an electron (e) or a muon (μ). The charge conjugate modes are implicitly included. The B meson decays $B^0 \to e^+e^-$ and $B^0 \to \mu^+\mu^-$ decays could proceed at unmeasurably low rates through the Flavor-Changing Neutral Current (FCNC). The decay mode $B^0 \to e^\pm\mu^\mp$ is forbidden within the SM, in which lepton number and lepton flavor are conserved.

1.1 The Standard Model

The Standard Model of particle physics (SM) encapsulates science's best understanding of particles and their interactions. According to the SM, all matter is constituted of fractional spin particles called fermions that interact through the exchange of integer spin particles called gauge bosons. Strong, electromagnetic, and weak particle interactions are described in the SM. Quarks are subject to all four fundamental forces (strong, electromagnetic, gravitational and weak). Leptons interact via electroweak interactions. To date, six quarks have been discovered (u, d, c, s, t, b) which strongly interact via the exchange of eight gauge bosons known as gluons. There are six leptons: e, μ, τ and their corresponding neutrinos ν_e, ν_μ, and ν_τ, which interact weakly by exchanging the gauge bosons W^+, W^-, and Z^0. Electromagnetic interactions among the electrically charged particles occur by exchanging photons [1]. With confirmation of its existence by CERN in 2014, the neutral Higgs boson completes the Standard Model particle zoo [2]. Properties of the known elementary particles are listed in Table 1.1.

1.2 Physics of $B \to \ell\ell$

The Standard Model does not allow FCNC at the tree-level, and decays of this type are predicted to have very small branching fractions. Rare decays such as these are therefore interesting as tests of the SM. If the measured branching fraction exceeds the predicted

Table 1.1: Properties of known elementary particles. For each fermion, an antifermion with opposite charge exists.

Particle	Symbol	Mass (MeV/c^2)	Charge	Spin
QUARKS				
up	u	$1.7 - 3.3$	2/3	1/2
down	d	$4.1 - 5.8$	-1/3	1/2
charm	c	1270	2/3	1/2
strange	s	101	-1/3	1/2
top	t	172	2/3	1/2
bottom	b	4200	-1/3	1/2
LEPTONS				
electron	e	0.511	-1	1/2
neutrino e	ν_e	$< 7 \times 10^{-6}$	0	1/2
muon	μ	105.7	-1	1/2
neutrino μ	ν_μ	< 0.26	0	1/2
tau	τ	1777	-1	1/2
neutrino τ	ν_τ	< 31	0	1/2
GAUGE BOSONS				
W	$W^\pm$	80×10^3	± 1	1
Z	Z	91×10^3	0	1
photon	γ	0	0	1
gluons	g	0	0	1

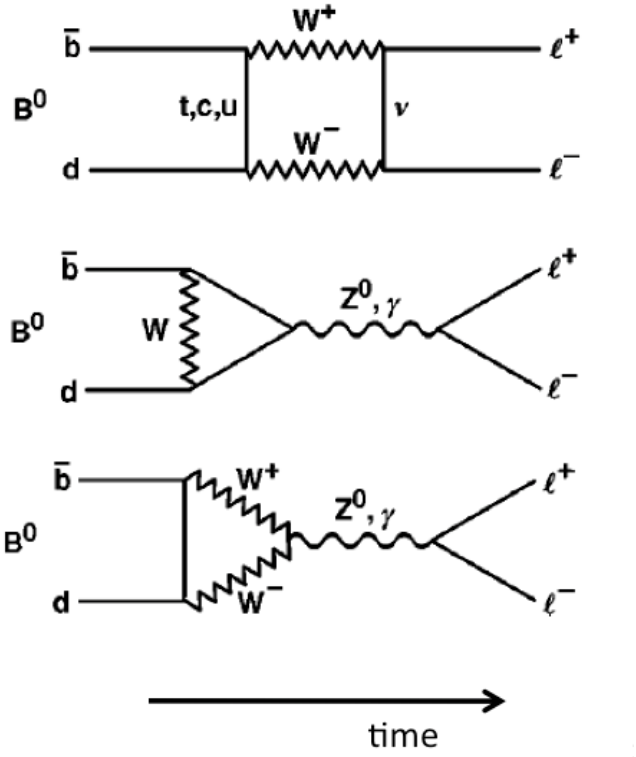
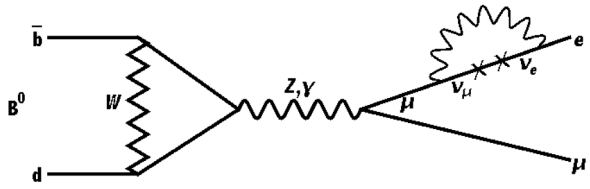


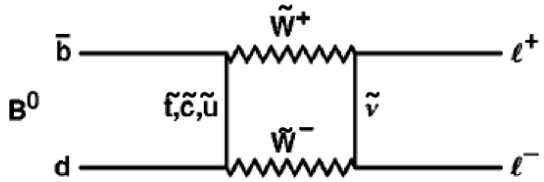
Figure 1.1: Representative Feynman diagrams for $B^0 \to \ell^+\ell^-$ decays allowed in the Standard Model

value, the result would be indicative of new physics which could include supersymmetric particles [3, 4] or the Pati-Salam leptoquark [5].

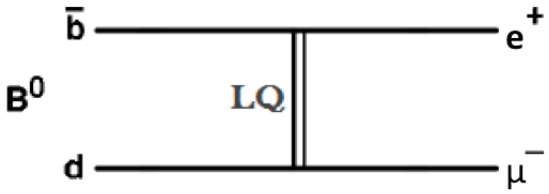
In the Standard Model, $B^0 \to e^+e^-$ and $B^0 \to \mu^+\mu^-$ decays proceed through penguin or box diagrams similar to those shown in Figure 1.1. These contributions are highly suppressed since they have $b \to d$ transition and require an internal quark annihilation. The decays are also helicity suppressed by a factor of $(m_\ell/m_B)^2$, where m_ℓ is the mass of the lepton and m_B is the mass of the B meson [6]. Helicity suppression contributes a factor of 10^{-8} to



(a) Neutrino oscillation



(b) Supersymmetry



(c) Pati-Salam Leptoquark

Figure 1.2: Representative Feynman diagrams for $B^0 \to \ell^+\ell^-$ decays beyond the Standard Model

$B^0 \to e^+e^-$ mode and 4×10^{-4} to $B^0 \to \mu^+\mu^-$ mode.

Furthermore, $B^0 \to e^\pm\mu^\mp$ decays involve a final state that violates lepton flavor conservation; so, they are forbidden in the SM except when accounting for the exceedingly small probability ($\approx 10^{-50}$) of neutrino oscillation [7]. A signal in this analysis mode would not only support the existence of new physics but could also help discriminate among different new physics models [8, 9, 10]. Figure 1.2 shows theoretical processes through which $B^0 \to e^\pm\mu^\mp$

could proceed [11].

1.3 Branching Fraction

In the context of the SM, the predicted decay rate for $B \to \ell\ell$ decays can be predicted with small error because of the limited impact of long distance hadronic corrections [12, 13]. There is only one SM dimension six operator that contributes to $B \to \ell\ell$ [14]. It is:

$$Q = (\bar{b}d)_{V-A} \cdot (\bar{\ell}\ell)_{V-A} = (\bar{b}_L \gamma^\mu d_L) \cdot (\bar{\ell}\gamma_\mu\gamma_5\ell) \tag{1.1}$$

From this starting point, the effective hamiltonian can be written as [15, 16]

$$\mathcal{H}_{eff} = -\frac{G_F}{\sqrt{2}} \frac{\alpha}{2\pi \sin^2 \Theta_W} V_{tb}^* V_{td} \cdot Y(x_t) \cdot Q + \text{h.c.} \tag{1.2}$$

where

- The Fermi coupling constant, G_F, is defined by

$$\frac{G_F}{(\hbar c)^3} = \frac{\sqrt{2}}{8} \frac{g^2}{m_W^2}, \tag{1.3}$$

  where g is the coupling constant of the weak interaction and m_W is the mass of the $W^\pm$ boson.

- α is the fine structure constant,
- Θ_W is the Weinberg angle,
- V_{tb}^* and V_{td} are CKM matrix elements,
- $Y(x_t)$ accounts for the QCD correction [17]

$$Y(x) = Y_0(x) + \frac{\alpha_s}{4\pi} Y_1(x), \tag{1.4}$$

- Q is the dimension-six operator defined in Eq. 1.1, and
- h.c. is the hermitian conjugate.

The Standard Model prediction of the branching fractions then can be calculated by solving the equation [18, 19]

$$\mathcal{B}(B^0 \to \ell^+\ell^-) = \tau(B)\frac{G_F^2}{\pi}\eta_Y^2 \left(\frac{\alpha}{4\pi \sin^2(\Theta_W)}\right)^2 F_B^2 m_\ell^2 m_B \sqrt{1 - 4\frac{m_\ell^2}{m_B^2}} \left|V_{tb}^* V_{td}\right|^2 Y^2(x_t) \tag{1.5}$$

where the terms are as follows [20]:

- $\tau(B)$ is the mean lifetime of the B meson,
- G_F is the Fermi coupling constant,
- α is the fine structure constant,
- Θ_W is the Weinberg angle,

Table 1.2: The values used in the calculation of $B^0 \to \ell^+\ell^-$ branching fraction.

Variable	Value
$\tau(B)$	1.519(7) ps [21]
G_F	1.16638×10^{-5} GeV^{-2} [21]
α^{-1}	127.937 [22]
$\sin^2\Theta_W$	0.22290 [20]
F_B	0.186(4) GeV [23]
m_e^2	0.510999 MeV [24]
m_μ^2	105.6584 MeV [21]
$\lvert V_{tb}^* V_{td}\rvert$	0.0087(2) [25, 26]

- F_B is the decay constant, defined by

$$\langle 0|\,(\bar{b}d)_{V-A,\ell}\,|B(p)\rangle = iF_B p_\ell. \tag{1.6}$$

- m_ℓ is the mass of the lepton,

- m_B is the mass of the B meson,

- $|V_{tb}^* V_{td}|$ is the product of CKM matrix elements, and

- $Y^2(x_t)$ is the QCD correction, given by

$$Y^2(x_t) = Y_0^2(x_t) + \frac{\alpha}{2\pi} Y_0(x_t) Y_1(x_t). \tag{1.7}$$

The values used to calculate the $B^0 \to \ell^+\ell^-$ branching fractions are provided in Table 1.2. The resulting SM predicted branching fractions are $\mathcal{B}(B^0 \to e^+e^-) = (2.34 \pm 0.33) \times 10^{-15}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.05 \pm 0.05) \times 10^{-10}$ [23].

Table 1.3: The expected branching fractions in the Standard Model and the best upper limits (UL) at 90% C.L. as reported by the PDG. Values marked by an asterisk represent the best measured upper limit of the branching fraction at the 90% credibility level.

Decay mode	$B \to e^+e^-$	$B \to \mu^+\mu^-$	$B \to e^\pm\mu^\mp$
SM Prediction [18, 19]	2.3×10^{-15}	1.0×10^{-10}	0
BABAR [27]	$< 11.3 \times 10^{-8}$	$< 5.2 \times 10^{-8}$	$< 9.2 \times 10^{-8}$
Belle [28]	$< 1.9 \times 10^{-7}$	$< 1.6 \times 10^{-7}$	$< 1.7 \times 10^{-7}$
CDF [29]	$< 8.3 \times 10^{-8}$ *	$< 3.8 \times 10^{-9}$	$< 6.4 \times 10^{-8}$
CLEO [30]	$< 8.3 \times 10^{-7}$	$< 6.1 \times 10^{-7}$	$< 15 \times 10^{-7}$
CMS [31]	-	$< 9.2 \times 10^{-10}$	-
LHCb [32, 33]	-	$< 8.1 \times 10^{-10}$ *	$< 2.8 \times 10^{-9}$ *

1.4 Previous Studies

The rare decays $B^0 \to \ell^+\ell^-$ have been studied by multiple collaborations including Belle, Babar, CDF, and the LHC experiments: CMS and LHCb. Table 1.3 presents the upper limits of branching fraction as measured by each experiment. This section will discuss the work previously done at Belle by M.-C. Chang and the studies which currently present the strictest upper limit on branching fraction for each of the three analysis modes which are marked in the table with an asterisk.

1.4.1 $B \to \ell\ell$ at Belle

The first study of $B^0 \to \ell^+\ell^-$ at the Belle experiment was completed by M.-C. Chang in 2003. Using a sample of 78 fb^{-1} of data collected at Belle which corresponds to 85 million $B\bar{B}$ pairs, Chang set upper limits on the branching fractions at $\mathcal{B}(B^0 \to e^+e^-) < 1.9 \times 10^{-7}$,

$\mathcal{B}(B^0 \to \mu^+\mu^-) < 1.6 \times 10^{-7}$, and $\mathcal{B}(B^0 \to e^\pm\mu^\mp) < 1.7 \times 10^{-7}$ [34, 28].

1.4.2 $B \to \ell\ell$ at CDF

CDF, operating at the Fermilab Tevatron, analyzed data corresponding to $2fb^{-1}$ of integrated luminosity of $p\bar{p}$ collisions collected at $\sqrt{s} = 1.96$ TeV. In 2007, they reported a branching fraction limit of $\mathcal{B}(B^0 \to e^+e^-) < 8.3 \times 10^{-8}$ at the 90% credibility level [29] which holds as the best limit tabulated by the PDG [35].

1.4.3 $B \to \ell\ell$ at LHCb

LHCb is one of four particle detector experiments located at CERN's Large Hadron Collider. In 2013, LHCb published an upper limit of $\mathcal{B}(B^0 \to \mu^+\mu^-) < 1.0 \times 10^{-9}$ at the 95% confidence level and $\mathcal{B}(B^0 \to \mu^+\mu^-) < 8.1 \times 10^{-10}$ at the 90% confidence level which is currently recognized as the best limit by the Particle Data Group [35]. This limit on branching fraction is the strictest upper limit seen for $B^0 \to \mu^+\mu^-$ and exceeds the possible sensitivity of the Belle data sample. The combined data set analyzed corresponds to 1 fb^{-1} of pp collisions collected with a center of mass energy of 7 TeV and 2 fb^{-1} of pp collisions collected with a center of mass energy of 8 TeV [32]. In the summer of 2015, the CMS (Compact Muon Solenoid) and LHCb collaborations performed a joint analysis and published the first

evidence of $B^0 \to \mu^+\mu^-$. With a 3 sigma significance, they reported a branching fraction of $(1.06 \pm 0.09) \times 10^{-10})$ [36].

In a separate study published in 2013, LHCb reported the upper limit on $\mathcal{B}(B^0 \to e^{\pm}\mu^{\mp}) < 2.8 \times 10^{-9}$ at the 90% confidence level which is currently the best limit for the lepton-flavor violating mode. The data sample used for this study was 1 fb^{-1} of pp collisions collected with a center of mass energy of 7 TeV [33] .

Chapter 2

The Belle Experiment

In their classic 1973 paper, Kobayashi and Maskawa presented the possibility of CP violation within the Standard Model [37]. To investigate the B-meson system in detail, two so-called B-factories were constructed: BABAR/PEPII at SLAC in the USA and Belle/KEKB at KEK in Japan. Both collided electrons and positrons counter-circulating in high-energy storage rings. The KEKB accelerator in Japan provided high luminosity particle beams for the Belle detector which was installed at the single interaction point of KEKB [38].

2.1 The $\Upsilon(4S)$ Resonance

The $\Upsilon(nS)$ resonances are bound states of b and $\bar{b}$ quarks with quantum numbers $J^{PC} = 1^{--}$. These mesons can be produced directly from an e^+e^- collision. The cross-section for e^+e^- annihilations was studied by the CLEO collaboration and is shown in Figure 2.1. Here, the

cross section was measured across an energy scan from 9.45 GeV to 11.15 GeV. The peaks in the data represent the family of Υ resonances. The mass of the $\Upsilon(4S)$ resonance is just above the threshold of $B\bar{B}$ production. More than 96% of the time, $\Upsilon(4S)$ decays into $B\bar{B}$ pairs [39, 40].

The flat area below the cross section line represents non-resonant light-quark production $e^+e^- \rightarrow qq$ where $q = u, d, s, c$. This background is commonly referred to as "continuum" background. The cross section for continuum production is roughly three times that of $\Upsilon(4S)$ hadron production [41]. Although Belle operates at the $\Upsilon(4S)$ resonance, it produces three times as many light-quark pairs as $B\bar{B}$ pairs, resulting in a major source of background for the study of B-physics. In order to study the effects of $q\bar{q}$ events, the Belle experiment sometimes takes data with the KEKB accelerator operating at 60 MeV below the $\Upsilon(4S)$ resonance. The $\sim$ 10% of the Belle data taken at 60 MeV below the $\Upsilon(4S)$ resonance is referred to as "off-resonance" data.

2.2 The KEKB Accelerator

In order to facilitate the study of CP violation in the decays of the B-mesons in the Υ(4S) decay, the beam momenta were chosen to be asymmetric to convert the time difference between the two B decays to a position separation in the lab frame. The momentum of the

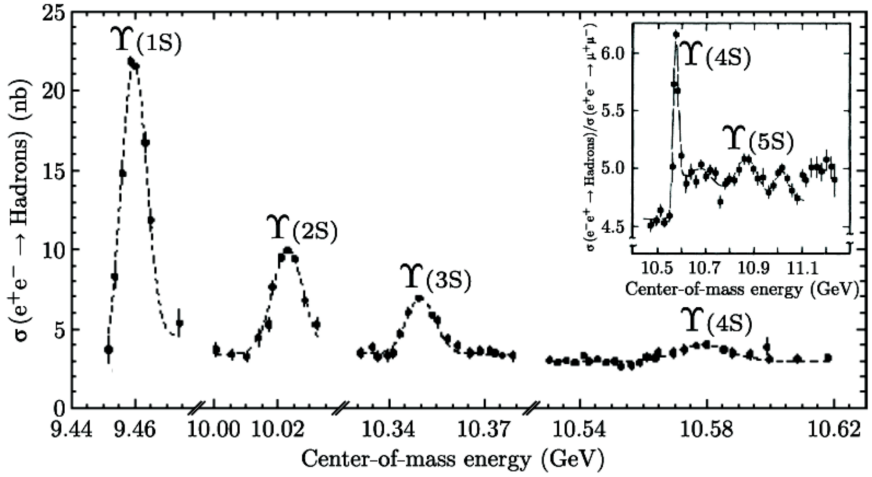


Figure 2.1: Cross section of e^+e^- annihilations in nanobarns [42], [43]

electron beam was 8 GeV/c, whereas that of the positron beam was 3.5 GeV/c. The rest frame of the produced particles in Belle e^+e^- annihilations is therefore boosted along the direction of the e^- beam in the laboratory frame. This circumstance converts the decay-time difference of the B-mesons in the Υ(4S) decay into a measurable decay-position difference along the boost direction. In order to account for this boost in the forward direction, the Belle detector was designed asymmetrically with respect to the interaction point.

The KEKB accelerator system consists of several components: an electron source, several pre-accelerators, a linear accelerator (LINAC), a positron source at one point along the LINAC, and two main storage rings; one for electrons (High Energy Ring - HER) and one for positrons (Low Energy Ring - LER). Electrons and positrons are accelerated to their respective energies, 8 GeV and 3.5 GeV, in the LINAC system, and injected into HER and

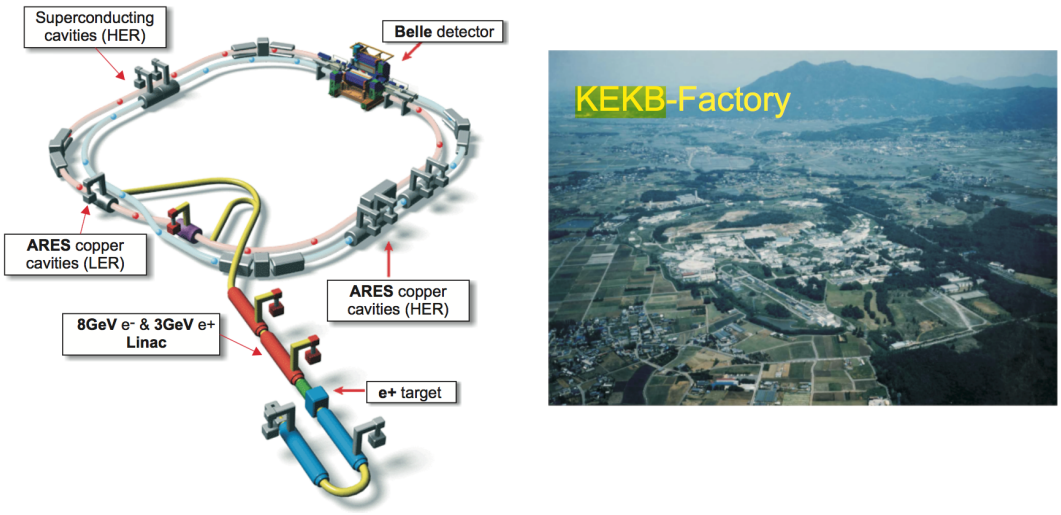


Figure 2.2: The KEKB accelerator machine and its location within KEK in Tsukuba, Japan.

LER [44]. The length of both rings is about 3 km and roughly 5000 particle bunches circulate in each ring with a bunch spacing of 59 cm [45]. In KEKB the designed beam currents are 1.1A and 2.6A for HER and LER, respectively, and the beam size at the interaction point is 90 μm in the horizontal and 1.9 mm in the vertical direction [46]. A schematic picture of the accelerator system is shown in Fig. 2.2.

The detector was operated from early 1999 through mid-2010. CP-violation in the B-meson-system was reported at Belle by 2001 [47]. In addition to the main physics goals of CP violation discovery, the detector performance and the high statistics available make it possible to carry out precision measurements in hadron spectroscopy and QCD fragmentation physics [48]. A world-record peak luminosity of 2.11×10^{34} cm^{-2}s^{-1} was achieved in 2009 [49].

2.3 The Belle Detector

The Belle detector is a multilayered general purpose detector surrounding the interaction point (IP) to detect the particles produced in e^+e^- collisions. The Belle detector is configured within a 1.5 T superconducting solenoidal magnet and iron structure. It is located at the Tsukuba interaction region of the KEKB beams and consists of the following subdetectors: a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL) comprised of CsI(Tl) crystals located inside the superconducting solenoid coil. Outside the magnet, multi-levels of resistive plate counters form the KLM. The barrel KLM was designed and built by Virginia Tech. Figure 2.3 shows the side-view of the Belle detector. A detailed explanation of the Belle detector can be found in Ref. [50].

The Cartesian detector coordinate system is centered at the nominal electron-positron interaction point and additionally specified by setting the positive z direction opposite the positron beam. The polar angle θ is defined with respect to the z axis. The +z direction is also referred to as the forward direction, as opposed to the backward direction along the negative z axis. (Because of the small non-zero crossing angle at the IP, the electron beam is not quite parallel to the positron beam.) The polar angle region from about $34^\circ - 120^\circ$ is labeled as the barrel region. Five charged particle species are considered stable within the

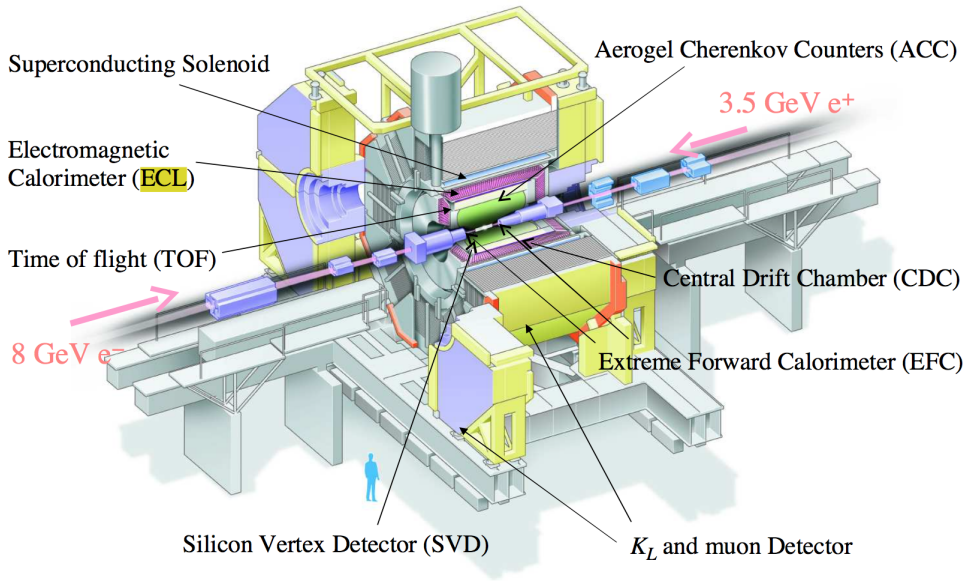


Figure 2.3: Cut away view of the Belle detector

volume of the Belle detector: these are electrons (e), muons (μ), pions (π), kaons (K), and protons (p.)

2.3.1 Beam Pipe

Although not an active component of the Belle detector, the beam pipe is the first material through which particles must traverse before reaching the detector. Since Coulomb scattering affects track resolution, it is important to minimize the impact of the beam pipe on particle trajectories by choosing a thin material with a low atomic number. In addition, track quality is reduced as the detector is radially farther from the IP. Therefore, the diam-

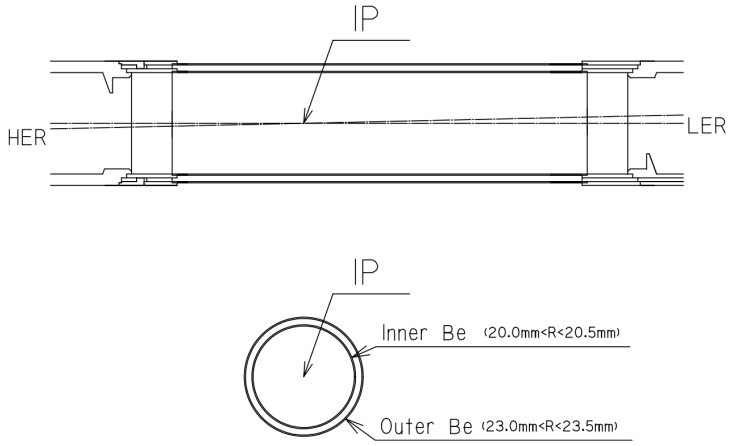


Figure 2.4: Schematic of the Belle beam pipe [50]

eter of the beam pipe must also be as small as possible.

With these considerations in mind, a beryllium beam pipe was chosen. As shown in Fig. 2.4, the beam pipe is a dual layer cylinder whose radii are 20.0 mm and 23.0 mm respectively, where each cylinder has a 0.5 mm thickness. The 2.5 mm gap between the cylinders is used as a helium gas channel which acts as a coolant. This coolant prevents overheating induced by the beam and minimally interferes with tracks due to its low atomic number. The outer layer is covered with a 20 μm thick gold sheet to reduce low energy X-ray background from the HER. The total thickness of the beam pipe corresponds to 0.9% of a radiation length. During the upgrade of the SVD detector in 2003, the beam pipe was redesigned, and its inner radius was decreased to 15 mm. The main motivation for the change was to place a new (fourth) SVD layer closer to the interaction point.

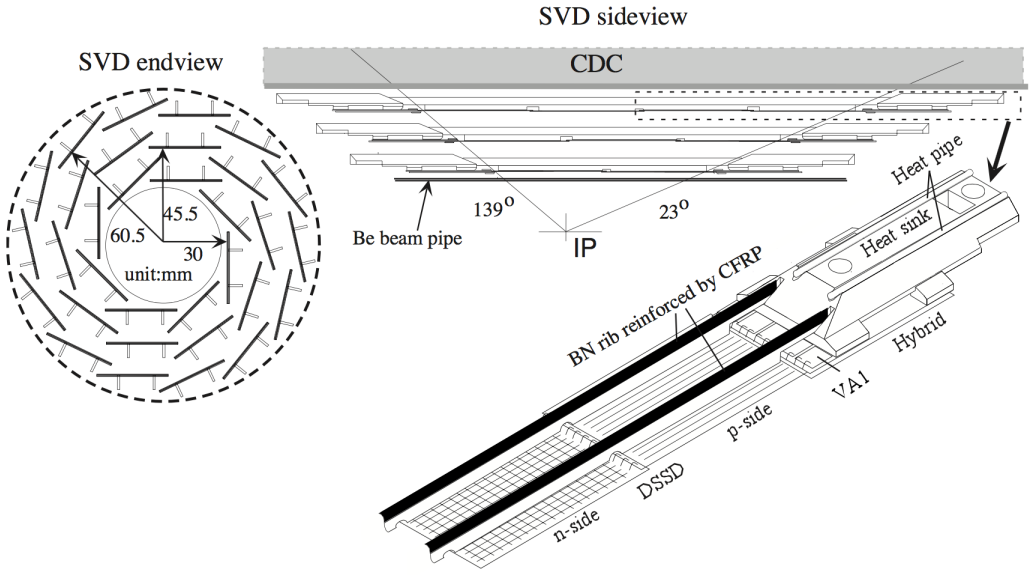


Figure 2.5: Schematic of the SVD [51]

2.3.2 Silicon Vertex Detector

The main goal of the Belle experiment was to study CP violation in B-decays, which requires a precise vertex resolution to effectively measure the difference in z-vertex positions for the B meson pairs. The Silicon Vertex Detector (SVD) was designed to provide the required z-resolution for B vertices of 100 μm or better. The SVD also contributed to the reconstruction of charged particles and helps improve the resolution of low momentum particles.

Figure 2.5 shows the side and endcap view of the first version of the vertex detector, SVD1. It consisted of three cylindrical detection layers of ladders, where each ladder was composed of two, three, or four double-sided silicon strip detectors (DSSD). SVD1 covered a solid angle

$23^\circ < \theta < 139^\circ$. The three layers of SVD1 consisted of 8, 10, and 14 ladders for inner, middle, and outer layers, respectively. Moving from the center outward, the ladder radii were 3.0 cm, 4.55 cm, and 6.05 cm. Each ladder was made of two half-ladders which were electrically independent and contained one or two DSSDs supported by boron-nitride (BN) ribs sandwiched with carbon-fiber reinforced plastic (CFRP) [52]. In total, there were 32 ladders and 102 DSSDs. The DSSDs were produced by Hamamatsu Photonics and were originally designed for the DELPHI micro-vertex detector [53]. They were essentially reverse-biased diode-strip detectors. A charged particle passing through the junction would liberate electrons from the valence band into the conduction band creating electron-hole pairs. These pairs create currents in the p^+ and n^+ strips aligned along and perpendicular to the beam axis respectively. Two-dimensional hit positions were determined from the charge distributions on these orthogonally segmented strips. Each DSSD had 1280 sense strips and 640 readout pads on both side. The size of the DSSDs was 57.5×33.5 mm^2. Signal from DSSDs were read out by 128 channel VA1 chips [54] placed on both sides of the ladder. The total number of readout channels was 81920 [55, 56].

As mentioned before, the main purpose of the SVD is to determine the B meson decay vertex and to improve the charged particle tracking. For studies of time-dependent CP asymmetries, the z-axis distance of the two B vertices for an $\Upsilon(4S)$ decay must be measured with the precision of about 100 μm. As shown in Fig. 2.6, the momentum and angular

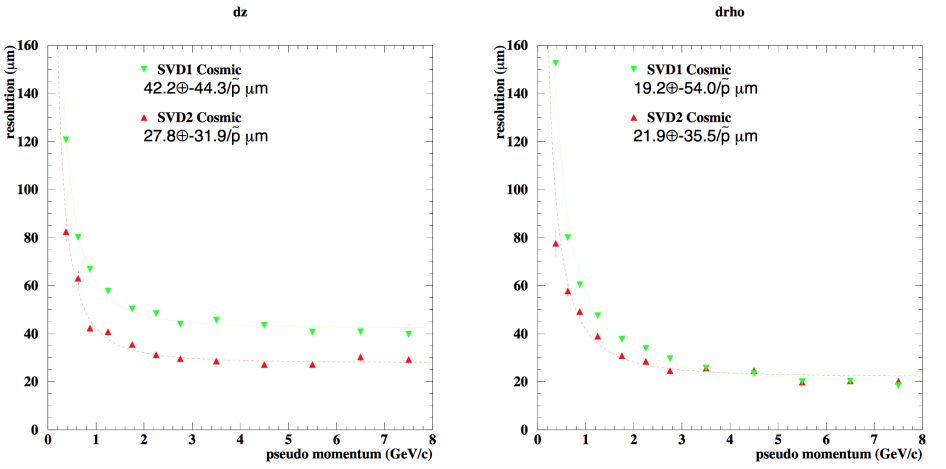


Figure 2.6: Impact parameter resolution of the SVD [57]

dependences of the impact parameter (closest approach of tracks to the IP) resolution give:

$$\sigma_{xy} = 19 \oplus \frac{49}{p\beta \sin^{3/2}\theta} \ \mu\text{m}, \tag{2.1}$$

$$\sigma_{z} = 36 \oplus \frac{42}{p\beta \sin^{5/2}\theta} \ \mu\text{m}, \tag{2.2}$$

where $\oplus$ indicates a quadratic sum, p represents the track momentum, β is speed over c, and θ is the polar angle. The impact parameter resolution for a 1 GeV/c perpendicular ($\theta = 90°$) track is around 55 μm. The SVD1 readout electronics had limited radiation tolerance [58] and had to be upgraded four times to the latest radiation tolerant readout chips without changing the three-layer mechanical structure. In the summer of 2003, the SVD1 was replaced by a four-layer structure SVD2, covering a range of $17° < \theta < 150°$ for improved tracking, better vertexing, and momentum resolution [59, 60]. The innermost layer is closer to the IP at a distance of 2 cm instead of 3 cm in SVD1. This is possible with

a significantly smaller beam pipe. The fourth layer is accommodated by a redesign of inner region of the CDC. The four layers of SVD2 contains 6, 12, 18, and 18 full ladders from inside to outside. Each half ladder consists of 1, 2, or 3 DSSDs. Both SVDs used a common double-sided silicon strip detector (DSSD) design. The DSSDs for SVD2 have 512 readout channels in both r - z and r - ϕ, or in total 110,592 readout channels [61]. In Ref. [62], the impact parameter resolution of the SVD2 detector perpendicular and along the beam axis, respectively, is reported as

$$\sigma_{xy} = 21.9 \oplus \frac{35.5}{p\beta \sin^{3/2}\theta} \ \mu\text{m}, \tag{2.3}$$

$$\sigma_z = 27.8 \oplus \frac{31.9}{p\beta \sin^{5/2}\theta} \ \mu\text{m}, \tag{2.4}$$

which are considerably improved over the SVD1 resolutions [63].

2.3.3 Central Drift Chamber

The Central Drift Chamber (CDC) is a cylindrical wire drift chamber and one of the most important subdetectors in the Belle experiment. The CDC is immersed in a 1.5 T magnetic field produced by the solenoid coil. Therefore, a charged track follows a helical trajectory in the CDC. The CDC provides the following three measurements:

- Precise determination of three-dimensional trajectories, providing measurements of charged particle momentum vectors,
- Measurement of charged particle energy loss in the chamber gas (dE/dx) for particle

identification,

- Provision of fast-track information for discriminating interesting physics events at the hardware-trigger level.

The structure of the CDC is shown in Figure 2.7. The CDC has a total of 8400 drift cells and each drift cell has a maximum drift distance between 8 mm to 10 mm. The sense wires are gold-plated tungsten wires of 30 μm diameter while the field wires are unplated aluminum of 126 μm diameter. When a charged particle travels through a cell, it ionizes the gas atoms. The released electrons cause the release of more electrons ("avalanching") while drifting to the sense wire. They are collected by the sense wires, and a hit is recorded by the CDC electronics. Three z-coordinate measurements at the inner-most radii are provided by cathode strips as shown in Fig. 2.8. The cathode strip having width of 7.4 mm is divided into eight segments in the ϕ direction and has an 8.2 mm pitch in the z-direction. There are a total of 1792 cathode channels.

The path of a charged particle in the constant magnetic field is a helix which is defined with five independent parameters measured by the CDC: the signed curvature of the helix, the slope of the helix, and the 3D coordinates of the helix reference point. The curvature radius is proportional to p_t, the slope is proportional to p_z, and particle charge is identified by the sign of the curvature. The reference point is called a pivot and chosen as the position of the innermost hit in the CDC.

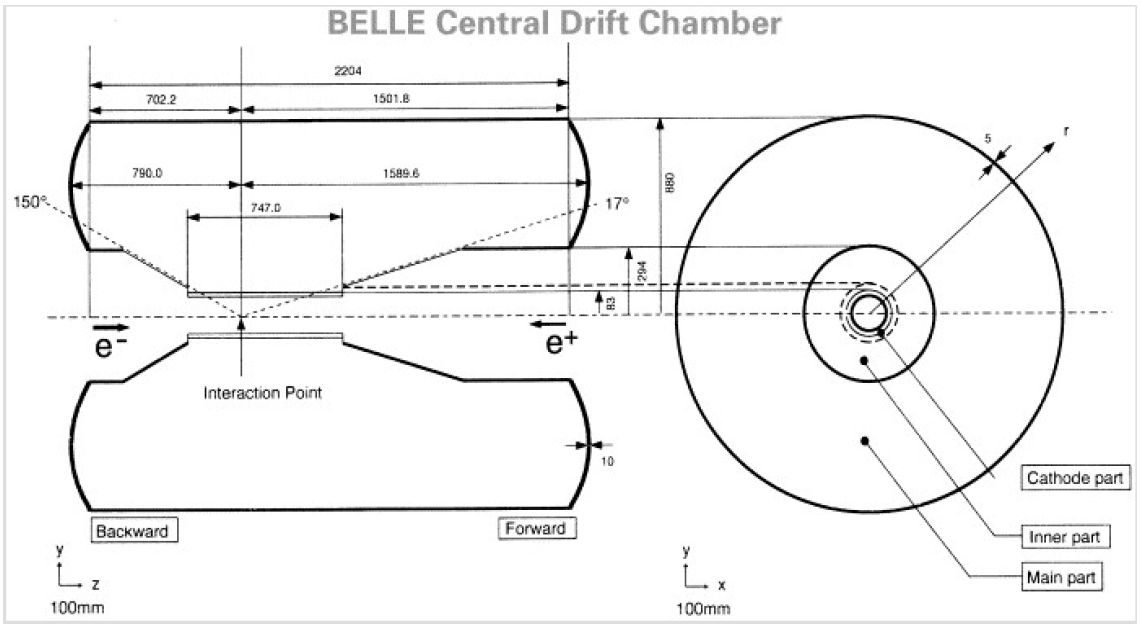


Figure 2.7: Schematic of the CDC [64]

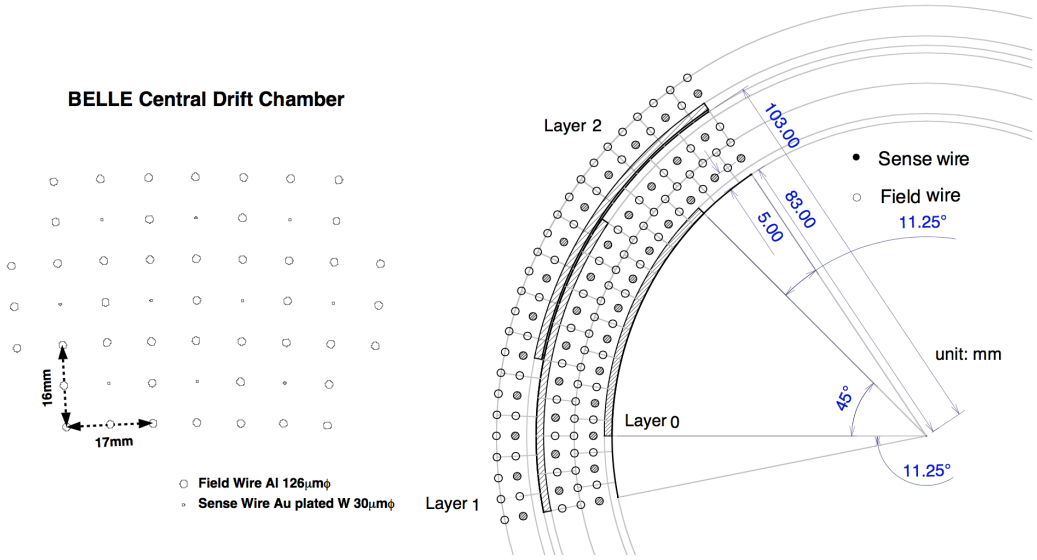


Figure 2.8: Cell structure of the CDC. Cathode sector configuration is also shown in the right figure.

To minimize multiple Coulomb scattering contributions to the momentum resolution of the charged particles, Belle uses a gas having low Z. A mixture of 50% helium (He) and 50% ethane (C_2H_6) gas is used in our experiment. This mixture has a long radiation length (640 m) and a drift velocity that saturates at 4 cm/μs at a relatively low electric field. A good dE/dx resolution is provided by the large C_2H_6 component, which increases the electron density and thereby improves the ionization energy loss measurement resolution.

Figure 2.9 shows the spatial resolution as a function of the drift distance. The spatial resolution is approximately $\sigma_{r\phi} = 130\mu$m. The transverse momentum resolution measured from the cosmic ray data is:

$$\sigma_{p_t}/p_t = \sqrt{(0.20p_t)^2 + (0.29/\beta)^2}\%, \tag{2.5}$$

where p_t is the transverse momentum measured in GeV/c and β is velocity in units of the speed of light.

The CDC is involved in particle identification for the tracks with $p < 0.8$ GeV/c and $p > 2.0$ GeV/c through the measurement of dE/dx. A charged particle's energy loss due to ionization (dE/dx) in the drift cell is determined using the hit amplitude recorded on the sense wire. Since the energy loss depends on the particle velocity at a given momentum, dE/dx will

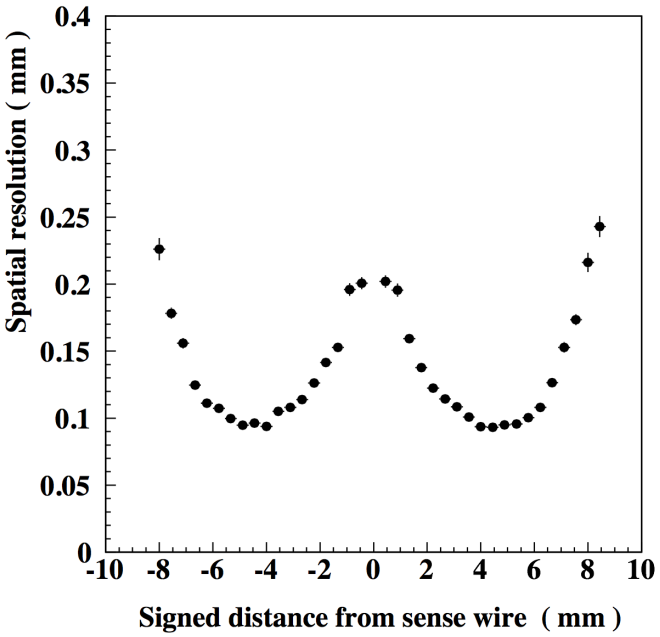


Figure 2.9: Signed distance from sense wire (mm).

vary according to the particle mass. Figure 2.10 shows the scatter plot of the measured dE/dx and the particle momentum p. The expected relation for π, K, p and e are shown by the solid curves in Fig. 2.10. The separation between the different particles can be clearly seen [65].

2.3.4 Aerogel Cherenkov Chamber

An important component of the Belle particle identification system (PID) is represented by an array of threshold silica Aerogel Cherenkov Counters (ACC). The ACC efficiently extends the PID momentum coverage beyond the common reach of dE/dx measurements in the CDC and the time-of-flight measurements by the TOF (described in the next section).

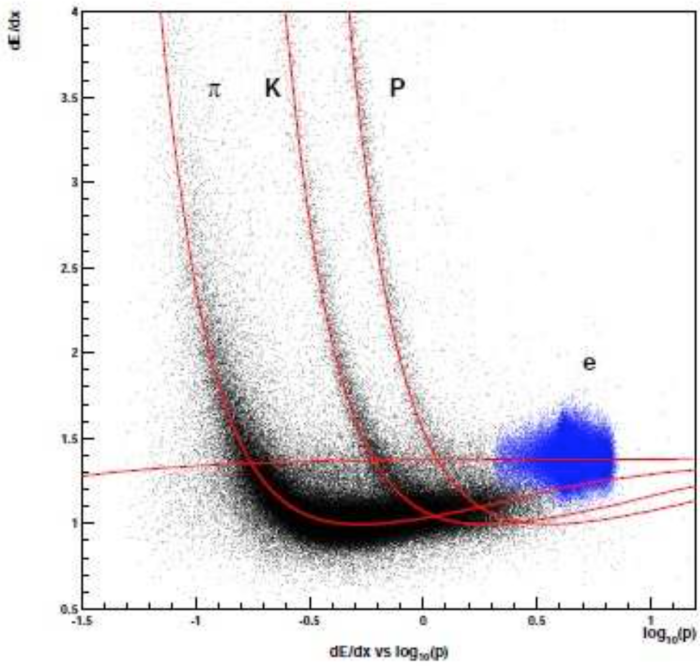


Figure 2.10: Scatter plot for momentum versus dE/dx. Expected relations for π, K, p, and e are shown in the solid curves. The momenta are given in units of GeV/c

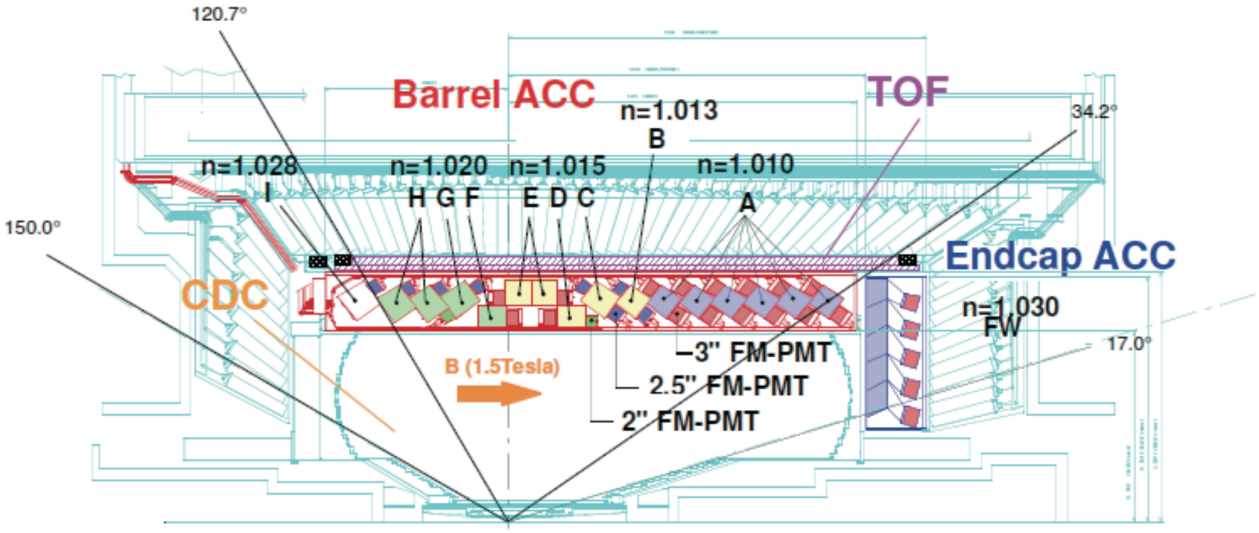


Figure 2.11: Schematic of the ACC showing the varied index of refraction [50]

Regarding the ACC sensitivity range, it was primarily designed to cover the PID of kaons and pions with momentum over 1 GeV/c and to provide a good separation between them up to momenta of 3.5 GeV/c [66].

If a particle passes through a medium faster than the speed at which photons propagate in the material, the particle emits Cherenkov photons. The speed of light in the medium is related to the refractive index of the material, n, via $c_{mat} = c/n$. For the given n, the threshold momentum for Cherenkov emission, p, is proportional to a particle mass, m, and inversely proportional to the refractive index: $p > mc/\sqrt{n^2 - 1}$. (The ACC does not measure the production cone angle of the Cherenkov light.) Therefore, the efficient way to separate $\pi^{\pm}/K^{\pm}$ in high momenta region is to choose a suitable refractive index for the radiator material.

Table 2.1: ACC threshold momenta corresponding to the choice of different refractive indices for the charged particles considered stable at Belle. The values for the refractive indices are taken from Ref. [50]

Refractive Index	e [MeV/c]	μ [GeV/c]	π [GeV/c]	K [GeV/c]	p [GeV/c]
1.010	3.6	0.75	0.98	3.48	6.62
1.013	3.2	0.65	0.86	3.05	5.80
1.015	2.9	0.61	0.80	2.84	5.40
1.020	2.54	0.53	0.69	2.46	4.67
1.028	2.14	0.44	0.59	2.07	3.94

Since the final-state particles are produced with a forward boost in the lab frame, the values of refractive index have to follow the forward-backward asymmetry of this boost. That can be efficiently achieved by a θ-dependent gradation of ACC modules, with various refractive indices from 1.010 - 1.028 which are shown in Fig. 2.11. The ACC blocks in the forward end-cap region use silica aerogel with $n = 1.030$. The higher index is chosen to incorporate the particle identification of low momenta particles into the PID system. The ACC is the only PID detector providing particle identification in the most forward end-cap region [67].

The ACC consists of 960 counter modules positioned in sets by 60 in ϕ direction for the barrel part and 228 modules arranged in 5 concentric layers for the forward end-cap part. The total polar angle covered by the ACC is $17° \leq \theta \leq 127°$. Emitted Cherenkov photons are measured by fine-mesh photomultipliers, which are able to work in the 1.5 T magnetic field [68].

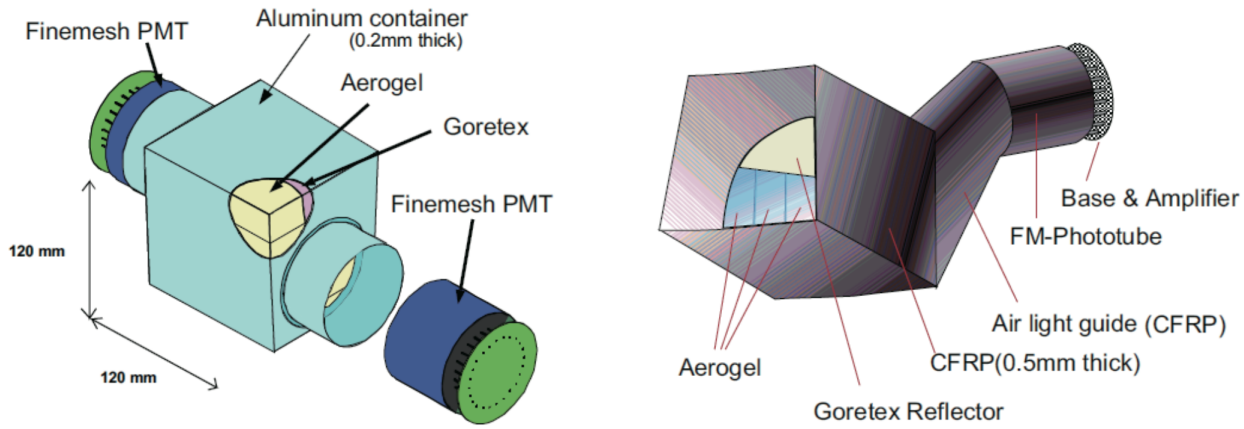


Figure 2.12: ACC modules: On the left is a barrel module. On the right is an endcap module.

2.3.5 Time of Flight Detector

The Time of Flight (TOF) is a particle identification device designed to provide precise time measurement for the event trigger and to identify charged particles over the region 0.8 GeV/c < p < 1.2 GeV/c. The system measures the time difference between the e^+e^- collision and the passage of a charged particle through the TOF counter. To obtain a 3σ separation between kaons and pions at a track length of about 1.2 m, the timing resolution has to be of the order of 100 ps. The mass of the particle m can be determined from the time-of-flight T measured with the TOF and the momentum p measured with the CDC as follows:

$$T = \frac{L}{c\beta} = \frac{L}{c}\sqrt{1 + \left(\frac{m}{p}\right)^2}, \tag{2.6}$$

where L is the length of the track.

The TOF counter employs a fast plastic scintillator with an attenuation length larger than 2m. Since the system lies in a 1.5 T magnetic field, fine mesh dynode photomultipliers are used to measure the scintillating photons. The timing resolution of the system is about 100 ps. Additionally, the TOF detector is enhanced by a thin Trigger Scintillation Counter (TSC) that provides an additional input to the trigger system. This detector component is comprised of 128 individual TOF counters and 64 TSC counters with 2 TOF and 1 TSC counter forming one module. The system covers a polar angle of $34^\circ < \theta < 120^\circ$ [69].

2.3.6 Electromagnetic Calorimeter

The main purpose of the electromagnetic calorimeter (ECL) is the detection of photons from B meson decays with high efficiency and good resolutions of both energy and position. Since most of these photons are end products of cascade decays, they have relatively low energies and, thus, good performance below 500 MeV is especially important. However, important modes with a high energy photon, such as $b \to s\gamma$, produce energies up to 4 GeV and high resolution is needed to reduce backgrounds. Good electromagnetic energy resolution results in better hadron rejection. High momentum π^0 detection requires the separation of two nearby photons and a precise determination of their opening angle. Therefore, a fine-grained segmentation is required for the ECL. Furthermore, the ECL is an important sub-detector for electron identification. In the ECL, energy from an photon or electron is deposited in electromagnetic showers produced by Bremsstrahlung and pair production, while other charged particles deposit a small amount of energy by dE/dx ionization. Consequently,

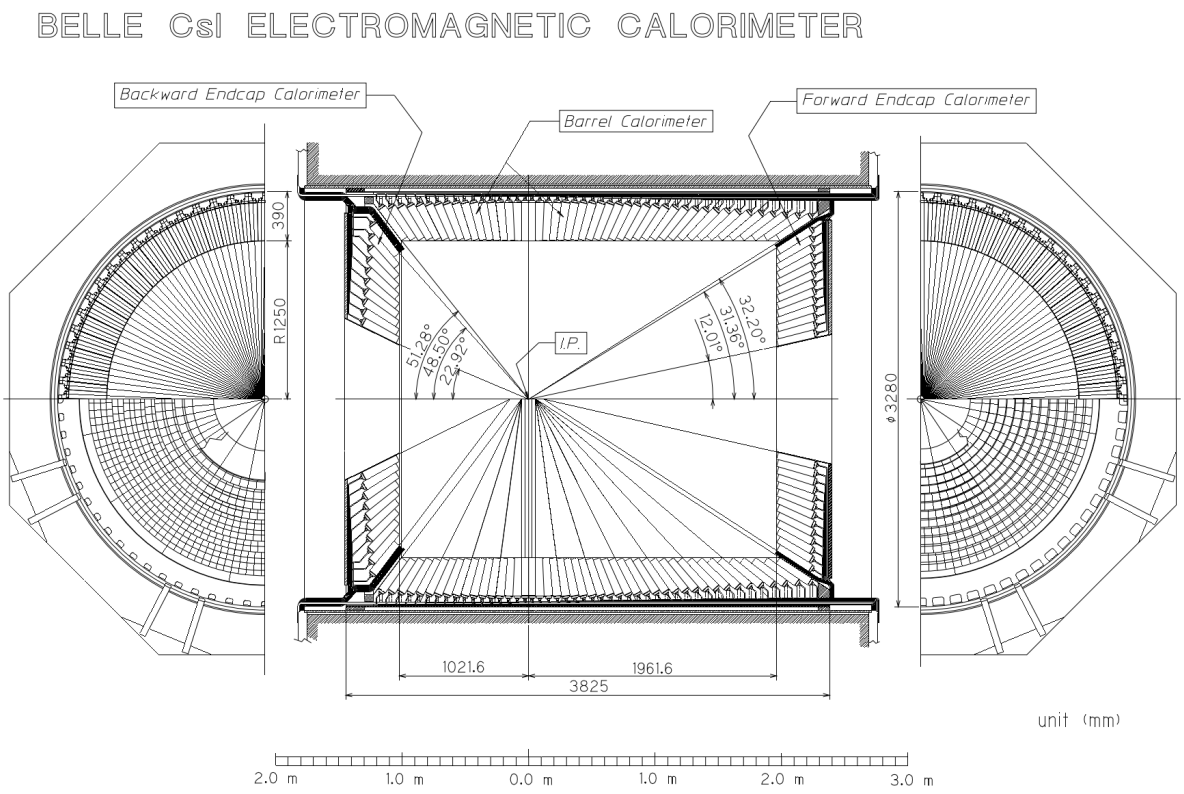


Figure 2.13: ECL Schematic.

the ratio of the cluster energy measured by the ECL to the charged track momentum measured by the CDC, E/p, is close to one for a electron and smaller than one for other particles.

In order to satisfy these requirements, a highly segmented array of CSI(Tl) crystals was assembled. CsI(Tl) crystals have various features such as a large photon yield, weak hygroscopicity, mechanical stability, and moderate price [70].

Figure 2.13 shows the overall configuration of the ECL, which contains 8736 crystals. As

Table 2.2: Geometrical parameters of ECL.

Item	Θ coverage	Θ segments	ϕ segments	No. of crystals
Forward Endcap	$12.4^\circ - 31.4^\circ$	13	48 - 144	1152
Barrel	$32.2^\circ - 128.7^\circ$	46	144	6624
Backward Endcap	$130.7^\circ - 155.1^\circ$	10	64 - 144	960

itemized in table 2.2, the ECL consists of three sections: the forward endcap section consists of 1152 crystals and covers $12.4^\circ < \theta < 31.4^\circ$, the barrel section has 6624 crystals and covers $32.2^\circ < \theta < 128.7^\circ$, and the backward section has 960 crystals and covers $130.7^\circ < \theta < 155.1^\circ$ [71].

The size of a crystal in the $\theta - \phi$ direction is determined so that one crystal constrains approximately 80% of the total energy deposit by a photon injected at the center of its front face. The typical dimension of a crystal is 55 mm $\times$ 55 mm at the front face and 65 mm $\times$ 65 mm at the rear face for the barrel part. The radial thickness is 30 cm, which corresponds to 16.2 radiation lengths. This thickness is enough to avoid deterioration of the energy resolution at high energy due to the shower leakage. The total weight of the crystals is about 43 tons. The light from each crystal is read out by two photodiodes mounted at the back face of each crystal [72].

The energy dependence of the average position resolution is estimated by MC and can be

approximated by

$$\sigma(\text{mm}) = 0.27 + \frac{3.4}{\sqrt{E}} + \frac{1.8}{\sqrt{E}}, \tag{2.7}$$

where the energy E is in GeV. The results are in good agreement with the beam test [73]. The energy resolution given by the beam test is

$$\frac{\sigma_E}{E} = \sqrt{\frac{0.066^2}{E} + \frac{0.81^2}{\sqrt[4]{E}} + 1.34^2}. \tag{2.8}$$

In addition to the measurement of the energy of photons and electrons, the ECL plays an important role in electron identification. The electron identification is performed by combining the following information:

- Matching between the position of the charged track measured by the CDC and that of the energy cluster measured by the ECL,
- E/p, the ratio of energy measured by the ECL to momentum measured by the CDC,
- E_9/E_{25} at the ECL, the ratio of ECL shower energy in an array of 3×3 crystals to the energy in an array of 5×5 crystals,
- dE/dx in CDC,
- Light yield in the ACC.

The probability density functions (PDFs) for the above parameters are made and then a likelihood ratio for every track is calculated.

2.3.7 Superconducting Solenoid

A crucial ingredient for a proper operation of the Belle tracking system is a superconducting solenoid. This coil provides a magnetic field of 1.5T in a cylindrical volume of 3.4 m in diameter and 4.4 m in length and consists of a single layer of niobium-titanium-copper alloy embedded in a high purity (99.99%) aluminum stabilizer. In order to contain magnetic flux and improve the field uniformity in the CDC volume, an iron structure of the Belle detector (iron yoke) is used. This segmented structure is also designed to effectively support the overall detection system and work as an absorber of the KLM, described in the next section.

2.3.8 K_L and Muon Detector

The K_L^0 and muon detector (KLM) provides muon identification for charged particles that possess enough momentum to reach the KLM, $p_t > 0.6$ GeV/c. It can also detect the neutral K_L^0. Since these are highly penetrating particles, a long radial depth is needed to identify them efficiently [74]. The KLM barrel detector was designed and built by Virginia Tech. The same technology was adopted for the endcaps and Virginia Tech's expertise in RPCs was transferred to our colleagues at Tohoku University, Tohoku Gakuin University and Osaka City University for the endcap KLM's fabrication.

The KLM consists of alternating layers of charged particle detectors and 4.7 cm thick iron plates. There are 15 resistive plate counter (RPC) superlayers and 14 iron layers in the barrel

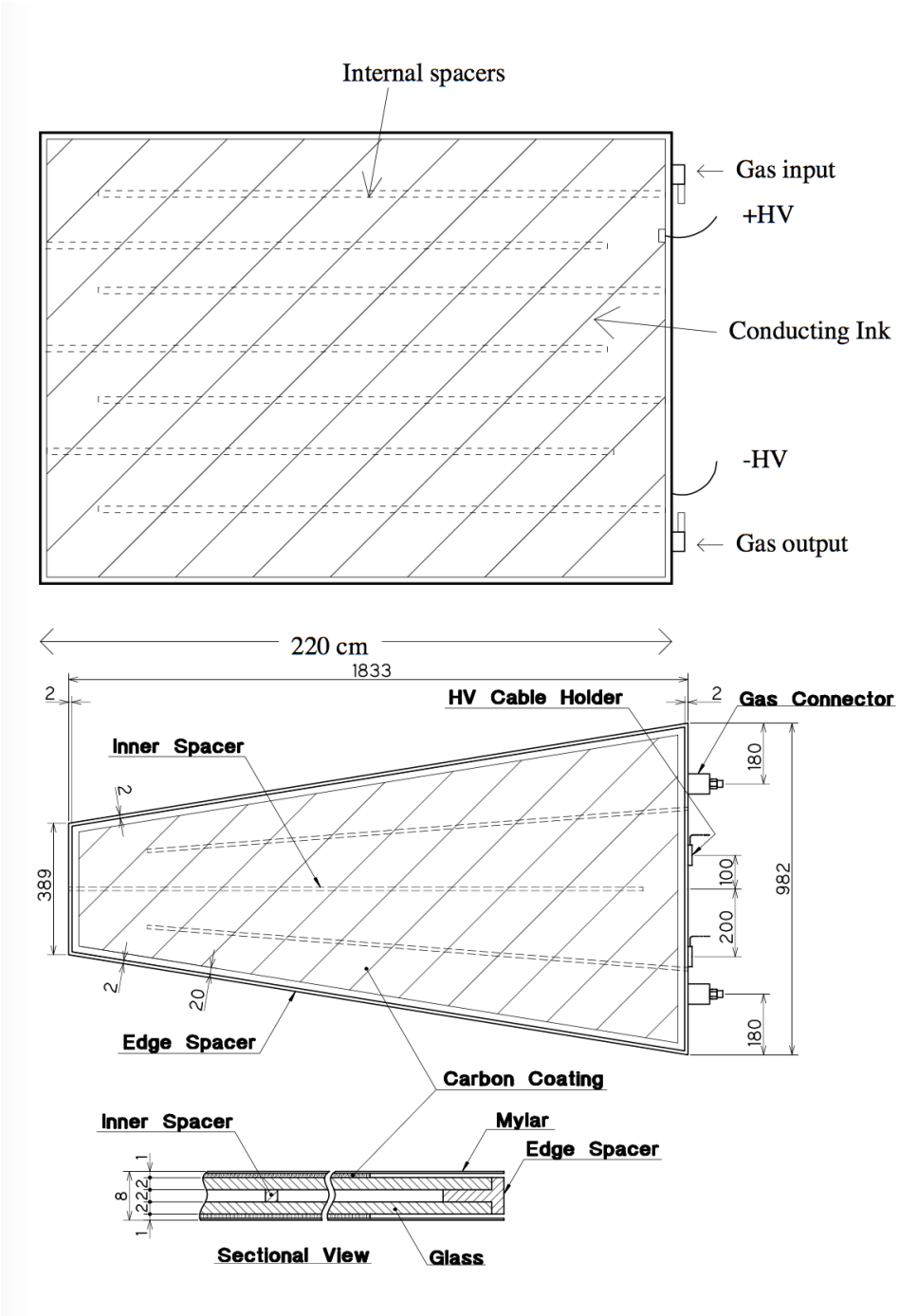


Figure 2.14: RPCs in the barrel (top) and endcap (middle); side view (bottom).

region and 14 RPC superlayers in each endcap, covering the polar region, $20° < \theta < 155°$. The iron layers also serve as a return yoke for the magnetic flux provided by the superconducting solenoid. The configuration of an RPC is shown in Fig. 2.14. Each RPC superlayer consists of two RPC modules to provide 2-dimensional $\theta - \phi$ information. The cross-section of an RPC superlayer is shown in Figure 2.15. Hadrons interacting with the iron plates produce a shower of ionizing particles that are detected by the RPC layers. The result is a cluster of hits deposited in the KLM. A K_L^0 candidate can be distinguished from another charged hadron because it will not leave an associated track in the CDC. A muon, on the other hand, does leave a charged track in the CDC. However, muons can still be distinguished from charged and neutral hadrons because they do not feel the strong interaction. Hadrons are more rapidly absorbed and deflected by strong interactions with iron, resulting in wide clusters that are stopped within a few layers of iron. Muons only experience electromagnetic multiple scattering and energy loss. Therefore, their KLM hit patterns tend to be thinner and have far greater penetration depth [75].

2.4 Trigger and Data Acquisition

The Belle trigger system is used to select signal events and reject background events as efficiently as possible. Moreover, it works hand in hand with the Belle data acquisition system, responsible for the data read-out from all sub-detector components. The trigger

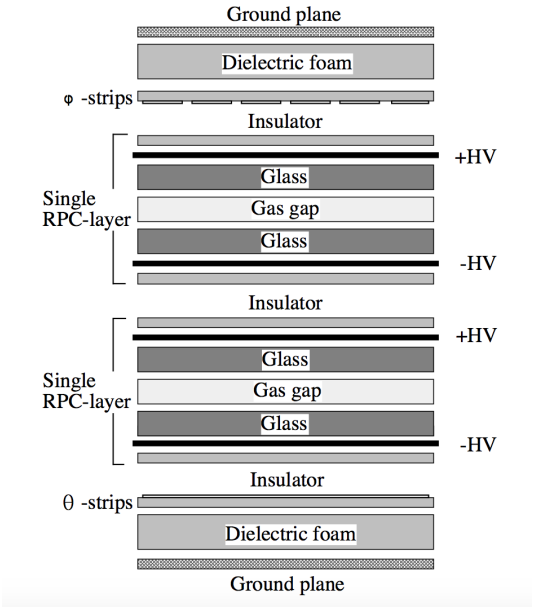


Figure 2.15: KLM RPC Superlayer Cross section

system reduces data event rate in three sequential steps: online data triggering with hardware Level-1 trigger, a software Level-3 trigger, and a more refined Level-4 trigger; the last is based on event reconstruction and is performed offline. The Level-1 trigger can be schematically described as following: the CDC and SVD readout electronics provide charged track signals, the TOF trigger system gives event timing information, the ECL gives trigger signals for both neutral and charged particles, based on their deposited energy and ECL hits cluster counting, and finally, the KLM provides an additional information on muons. All these trigger signals are controlled by a Global Decision trigger Logic circuit (GDL) [76], which makes the final global decision and issues the trigger signal to record all of the data from the various subdetectors for the event of interest. In January 2001, the second step in the triggering design was introduced, the Level-3 trigger, which uses fast track fitting algorithms

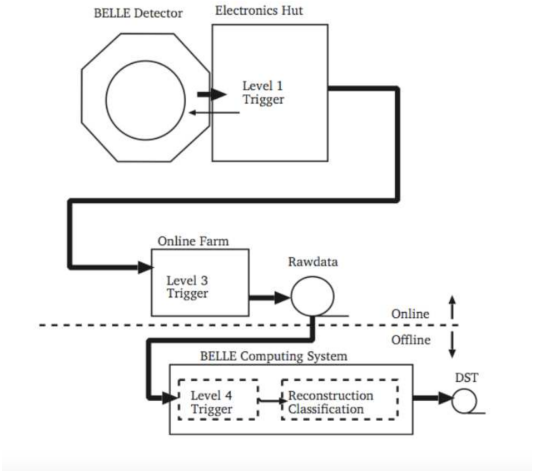


Figure 2.16: The Belle trigger system [77].

to suppress background tracks not coming from the interaction point. The third level, Level-4, is performed off-line on the recorded data and represents a prerequisite for any data analysis [77].

2.5 Belle Data Set

The list of all $\Upsilon(4S)$ data subsets acquired with Belle detector and available for data analysis is shown in Table 2.3. The numbers of $B\bar{B}$ pairs corresponding to each subset (so-called experiment) and used further in branching ratio calculations are listed in the same table. The full set of Belle data corresponds to 772×10^6 events. (About 30% more data were recorded at other energies but are not used in this analysis.)

Table 2.3: The list of Belle data subsets recorded at the $\Upsilon(4S)$ resonance's energy [78].

Experiment number	$N(B\bar{B}$-pairs)$[\times 10^6]$
7	6.4587 + 0.1615 - 0.0976
9	4.7597 + 0.0286 - 0.0473
11	8.8509 + 0.0517 - 0.0518
13	11.6998 + 0.2393 - 0.2392
15	13.5679 + 0.0963 - 0.1055
17	12.4588 + 0.3301 - 0.3301
19	27.1705 + 0.1676 - 0.1676
21	4.3371 + 0.0540 - 0.0676
23	6.4755 + 0.0675 - 0.0989
25	28.0008 + 0.3329 - 0.1605
27	28.1814 + 0.2110 - 0.1516
SVD1	151.961±1.241
31	19.6587 + 0.3045 - 0.3031
33	19.3022 + 0.3000 - 0.2987
35	18.5262 + 0.2861 - 0.2855
37	67.1819 + 1.0326 - 1.0319
39	47.0818 + 0.7265 - 0.7246
41	64.0134 + 0.9863 - 0.9857
43	61.5614 + 0.9493 - 0.9474
45	14.3538 + 0.2218 - 0.2215
47	41.2186 + 0.6406 - 0.6393
49	29.7271 + 0.4648 - 0.4634
51	41.8919 + 0.6605 - 0.6590
55	80.2472 + 1.2462 - 1.2439
61	37.4460 + 0.5624 - 0.5617
63	35.6231 + 0.5297 - 0.5291
65	41.7867 + 0.6317 - 0.6309
SVD1 & SVD2	771.581±10.566

Chapter 3

Signal and Background Monte Carlos

To prevent any bias, this research was performed as a blind study. All methods are tested and optimized using Monte Carlo simulated data. After the methods of analysis are fully validated, the optimized event selection and signal extraction procedure are applied to the full set of Belle data. This chapter describes the variables used for analysis as well as the Monte Carlo data used in this study.

3.1 Analysis Variables

3.1.1 B Meson Kinematics

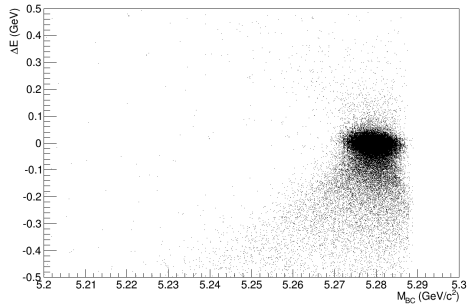
Throughout this analysis, many variables were used to separate the signal from the background events. The most powerful separation arises from the beam constrained mass

$$M_{bc} = \sqrt{\left(\frac{E_{beam}}{c^2}\right)^2 - \left(\frac{|p_B|}{c}\right)^2} \tag{3.1}$$

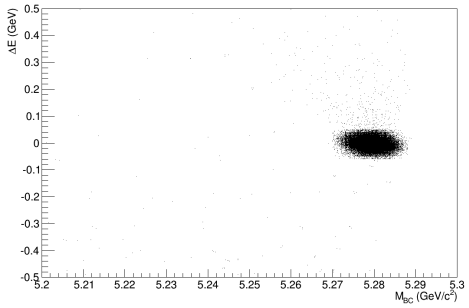
and the energy difference

$$\Delta E = E_B - E_{beam}. \tag{3.2}$$

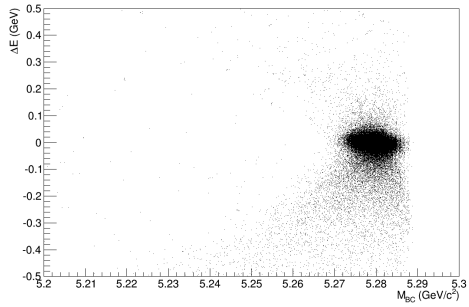
Here E_{beam} represents the beam energy, p_B is the momentum of the reconstructed B meson, and E_B is the energy of the reconstructed B meson. All measurements are made in the e^+e^- center of mass frame. The region of interest for this study is bounded by 5.2 GeV/c^2 $\leq M_{bc} \leq$ 5.3 GeV/c^2 and $|\Delta E| \leq 0.5$ GeV and is referred to as the fitting region. Figure 3.1 shows the fitting region for signal and selected background modes. The dominant $B\bar{B}$ backgrounds arise from the allowed two-body decays $B^0 \to \pi^+\pi^-$, π^+K^-, and K^+K^-, which share the same kinematics as the signal events.



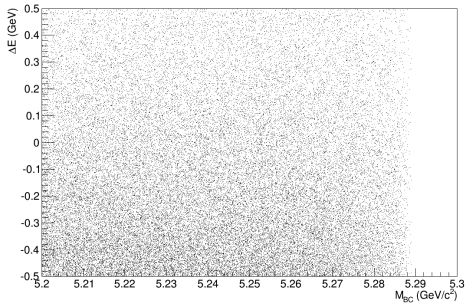
(a) $B \to e^+e^-$



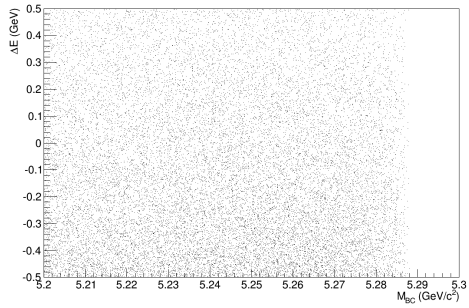
(b) $B \to \mu^+\mu^-$



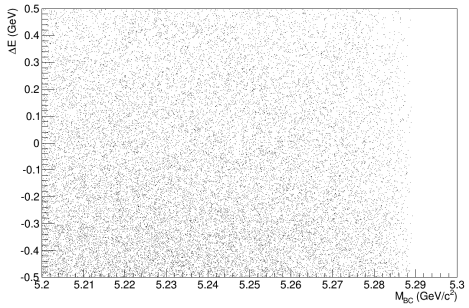
(c) $B \to e^\pm\mu^\mp$



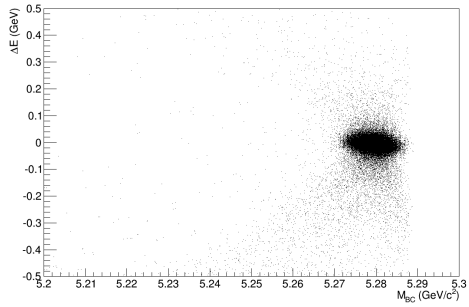
(d) $e^+e^- \to c\bar{c}$



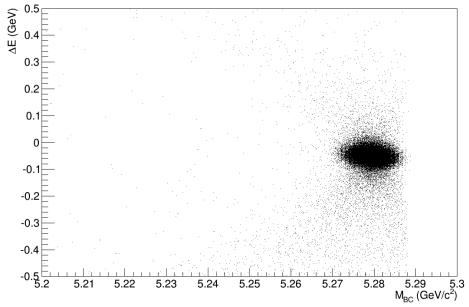
(e) $e^+e^- \to u\bar{u}$, $d\bar{d}$, $s\bar{s}$



(f) τ events



(g) $B \to \pi^+\pi^-$



(h) $B \to K^-\pi^+$

Figure 3.1: M_{bc} versus ΔE for signal modes and selected background modes

3.1.2 Event Shape

The dominant non-$B\bar{B}$ background for this study comes from $e^+e^- \to q\bar{q}$ (where $q = u, d, s,$ or c) continuum events. Three shape variables were identified as useful for continuum suppression. They are the Kakuno-Super-Fox-Wolfram (KSFW) discriminant, the cosine of the angle between the beam axis and the B meson flight direction in the $\Upsilon(4S)$ rest frame ($\cos\theta_B$), and the distance between the reconstructed vertices of the candidate B meson and other B meson (Δz).

The topology for continuum events is jet-like while signal events are spherical in the $\Upsilon(4S)$ rest frame. The Kakuno-Super-Fox-Wolfram (KSFW) method is an established technique in the Belle experiment which utilizes the difference in event shape to separate signal from background events [79]. The particles comprising one event are split into two categories: the daughter particles of the candidate B meson and the other particles in the event. The other particles in the event are assumed to be the daughters of the remaining B meson of the putative $B\bar{B}$ pair.

The KSFW discriminant is defined in terms of the event-shape moments as

$$\text{KSFW} \equiv \sum_{l=0}^{4} \sum_{m=c,n,v} \alpha_{l,m} (R_l^{so})_m + \sum_{l=0}^{4} \beta_l R_l^{oo} + \gamma \sum_{n=1}^{N_t} |P_{t,n}| \tag{3.3}$$

where P_t is the transverse momentum, N_t is the number of tracks in a event, and α, β, and γ are Fisher coefficients. R_l^{so} is defined as

$$R_l^{so} = \begin{cases} \frac{\sum_{j,k} Q_j Q_k |p_k| P_l(\cos\theta_{jk})}{E_{e^+e^-} - E_B} & (l = 1, 3) \\ \frac{\sum_{j,k} |p_j| P_l(\cos\theta_{jk})}{E_{e^+e^-} - E_B} & (l = 0, 2, 4) \end{cases} \tag{3.4}$$

- $P_l(\cos\theta_{jk})$ are Legendre polynomials of the cosine of the angle between the j^{th} and i^{th} tracks, where j and k run over signal B and other B tracks, respectively.
- $p_j(Q_j)$ and $p_k(Q_k)$ are the momentum (charge) of corresponding particles.
- $E_{e^+e^-}$ and E_B are the energy in the center of mass frame of e^+e^- and B, respectively.

R_l^{so} is calculated by dividing the particles from the other B meson into three categories: charged, neutral, and missing. This leads to the following set of conditional calculations

- $(R_l^{so})_{charged}$ (for $l = 0$ to 4) uses only the charged tracks of the other B meson
- $(R_l^{so})_{neutral}$ (for $l = 0$, 2, 4) uses only the photons of the other B meson
- $(R_l^{so})_{missing}$ (for $l = 0$, 2, 4) uses only the missing momentum of the other B meson

R_l^{so} produces a total of 11 parameters when summed over $l = 0$, 1, 2, 3, and 4.

R_l^{oo} is calculated similarly to R_l^{so} but the summations over j and k both include charged tracks not forming the signal B candidate. The resulting calculation becomes

$$R_l^{oo} = \begin{cases} \frac{\sum_{j,k} Q_j Q_k |p_j||p_k| P_l(\cos\theta_{jk})}{E_{e^+e^-} - E_B} & (l = 1, 3) \\ \frac{\sum_{j,k} |p_j||p_k| P_l(\cos\theta_{jk})}{E_{e^+e^-} - E_B} & (l = 0, 2, 4) \end{cases} \tag{3.5}$$

which creates 5 parameters corresponding to R_l^{oo}.

$\sum_{n=1}^{N} |P_{t,n}|$ is the scalar sum of the transverse momenta P_t of all particles in the event. The n represents particle index, and N is the number of particles in the event. This summation adds one additional parameter.

This calculation has a total of 17 parameters (α, β and γ) that must be obtained by prior training to distinguish best between signal and background. Empirically, the value of KSFW strongly depends on the square of the missing momentum in the event, defined as

$$mm^2 \equiv (E_{\Upsilon(4S)} - \sum_i E_i)^2 - \sum_i |p_i|^2. \tag{3.6}$$

The α, β and γ coefficients are optimized to distinguish between signal and continuum background separately in each of seven disjoint bins of mm^2 [80].

The ROOKSFW package is used to calculate the KSFW moments using signal and continuum Monte Carlo data and to optimize the separation coefficients using training samples of pure signal and continuum events. After completion of the training, the moments in an event under study are then combined into a likelihood ratio which gives a probability of the

event being signal or background. The KSFW likelihood ratio has a range of zero to one, where zero represents events jet-like in shape and one represents events spherical in shape. Therefore, an event falling on the left-hand side of the plot near zero is likely background, and an event with a KSFW value near one is likely a signal event.

3.1.3 Upsilon Decay Kinematics

The B flight direction determines the variable $\cos\theta_B$. The trajectory of the candidate B meson is measured to find the angle it makes with the beam axis. The cosine of this angle is known as $\cos\theta_B$. For this study, the absolute value of $\cos\theta_B$ will be used, and the absolute value operation is always implied. Figure 3.2 illustrates $\cos\theta_B$.

During reconstruction, all daughter particle tracks are combined to form the parent particles: the candidate B meson and the remaining B meson, referred to as the tagged B meson. The decay points, or vertices, of the two reconstructed B mesons are compared, and the distance of their separation along the beam axis is referred to as Δz. A schematic diagram for Δz is included in Figure 3.2.

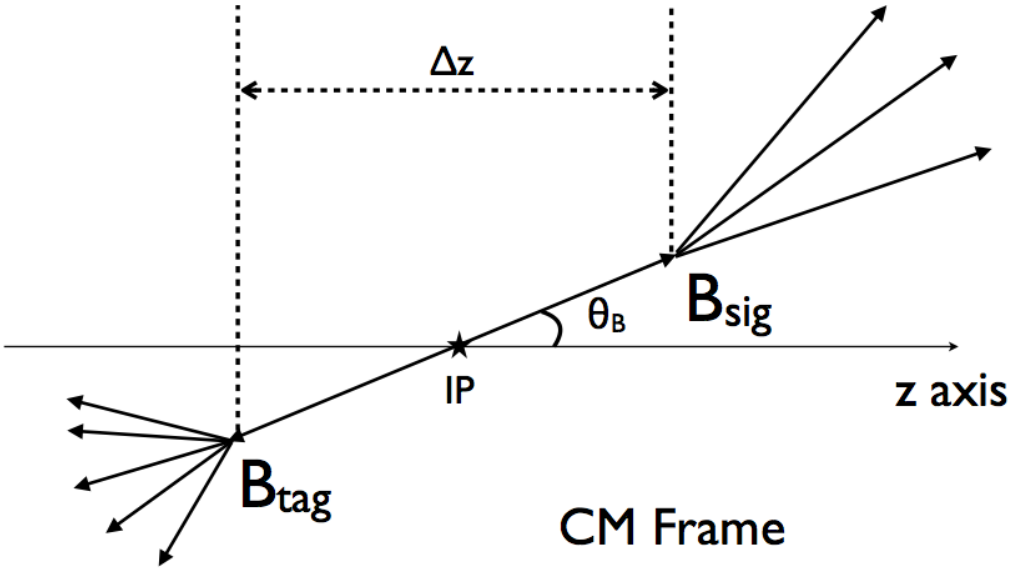


Figure 3.2: θ_B and Δz in the center of mass frame

3.1.4 Particle Identification

Another group of useful variables is the particle identification values: the lepton identification number (ℓID) for both electrons and muons, and the particle identification number of pions versus kaons. The lepton identification number is the expectation that a given particle is a lepton as opposed to a hadron. This variable ranges from zero to one, where zero indicates that the particle is not lepton-like, and one indicates that the particle in question is lepton-like.

The electron identification number, also known as eID, is designed to separate electrons from hadrons. The discriminants used to achieve separation include the ratio of measured

energy to track momentum (E/p), the transverse shower shape ($E9/E25$), the rate of energy loss (dE/dx), and the light yield [81]. The combination of these variables creates an efficiency of 95% for electrons with high momentum and a hadron fake rate of less than one percent [82, 83]. The Belle PID group recommends using cut values of 0.1, 0.5, 0.6, 0.8, or 0.9 for electron ID and has provided systematic errors for these values [82].

The muon identification number (μID) is primarily determined using tracking in the KLM detector. A charged track, reconstructed in the CDC, can be matched with nearby hits in the KLM as the track is extrapolated outward from the CDC [84]. For particles with high momentum, the muon detection efficiency is 95% and the pion fake rate is under 1.3% [82]. The PID group's recommended cut values for muon identification are 0.1, 0.8, 0.9, 0.95, and 0.97 [85], for which they provide systematic uncertainties.

Charged pions and kaons are reconstructed as tracks in the CDC and SVD. The individual likelihoods of particle identification ($\mathcal{L}_\pi$ and $\mathcal{L}_K$) are determined using information from the ACC, the TOF and the CDC and are used to establish a comparison ratio

$$\mathcal{L}_{\pi/K} = \frac{\mathcal{L}_\pi}{\mathcal{L}_\pi + \mathcal{L}_K}. \tag{3.7}$$

The measurement $\mathcal{L}_{\pi/K}$ then separates particles that behave like pions from particles that behave like kaons. The variable ranges from zero to one where zero represents kaon-like

behavior and one represents pion-like behavior. The PID group recommended cut values for pion-kaon separation are 0.1 through 0.9 in increments of 0.1 [86].

3.1.5 Charged Track Multiplicity

The total number of charged tracks (N_{track}) counts all charged tracks from both the candidate and tag side of an event. Some non-$B\bar{B}$ backgrounds, especially the tau production mode $ee \to \tau\tau$, consists of low numbers of charged tracks per event. This feature can be used to reduce these backgrounds.

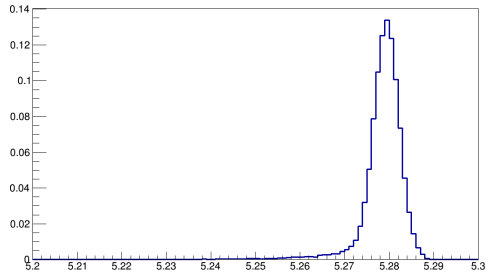
3.1.6 Other Variables

Variables considered but not used include $q \cdot r$ (the flavor of the B meson combined with the quality of its determination), the cosine of the angle between the thrust of the B candidate and the rest of the event ($\cos\theta_{thr}$), and dr and dz which were only used to eliminate badly reconstructed events. Variables such as the missing energy (E_{miss}) and the missing momentum (p_{miss}) are included in the KSFW moments calculation and were not incorporated separately.

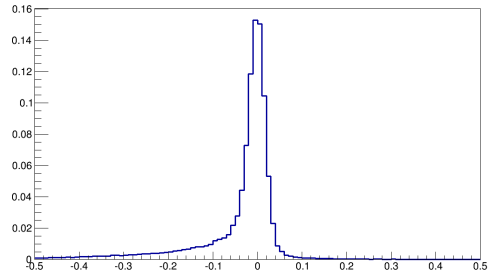
3.2 Signal Monte Carlo

This analysis focuses on three different signal modes: $B^0 \to e^+e^-$, $B^0 \to \mu^+\mu^-$, and $B^0 \to e^\pm\mu^\mp$. For each decay mode, two million luminosity-weighted, experiment- and run-dependent Monte Carlo events were generated. Using the cross-section of the $ee \to \Upsilon(4S) \to B^0\bar{B}^0$ interaction [87] and the limits on branching fraction provided by the Particle Data Group [5], the number of generated events represents 3×10^4 streams for $B^0 \to e^+e^-$, 4×10^6 streams for $B^0 \to \mu^+\mu^-$, and 9×10^5 streams for $B^0 \to e^\pm\mu^\mp$. One stream of data is defined as the number of events that corresponds to the integrated luminosity of Belle at the $\Upsilon(4S)$ resonance, assuming the expected branching fraction of the data type in question. The simulated Monte Carlo events were generated using the EvtGen package [88, 89]. The distributions of the analysis variables for the three signal modes are described below and can be seen in Figures 3.3, 3.4, and 3.5. The histograms are normalized to an area of one.

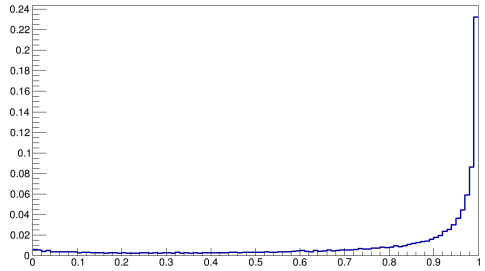
For signal modes, the beam constrained mass (M_{bc}) is a Gaussian distribution centered around 5.28 GeV/c^2. The energy difference (ΔE) for the analysis mode $B^0 \to \mu^+\mu^-$ is also Gaussian in shape and is centered at 0 GeV. $B^0 \to e^\pm\mu^\mp$ contains a skewed Gaussian distribution because of electron bremsstrahlung. In the $B^0 \to e^+e^-$ analysis mode, the long tail on the negative side of ΔE is even more densely populated since both signal tracks may radiate which causes an apparent reduction in energy during event reconstruction.



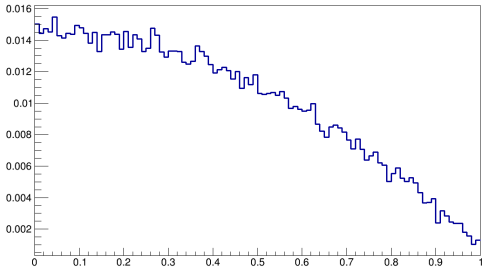
(a) M_{bc}



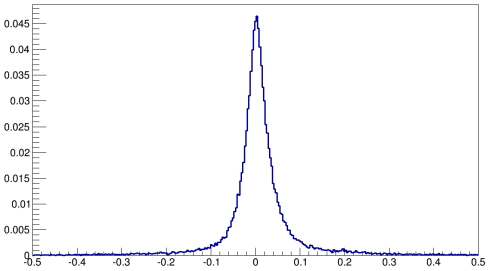
(b) ΔE



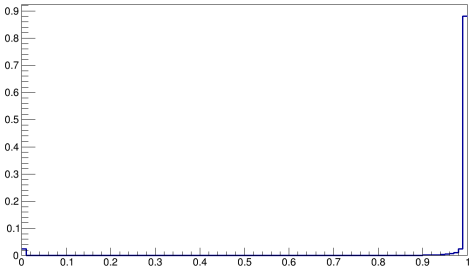
(c) KSFW Likelihood Ratio



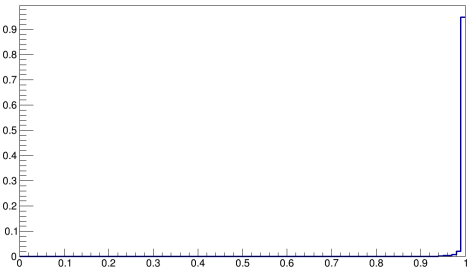
(d) $\cos\theta_B$



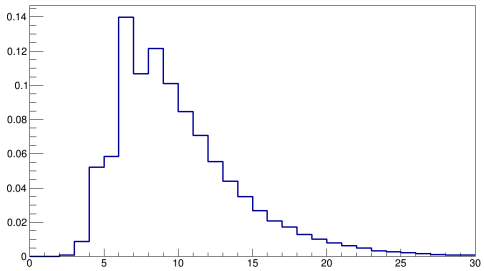
(e) Δz



(f) Electron PID



(g) $\mathcal{L}_{\pi/K}$



(h) Number of Charged Tracks

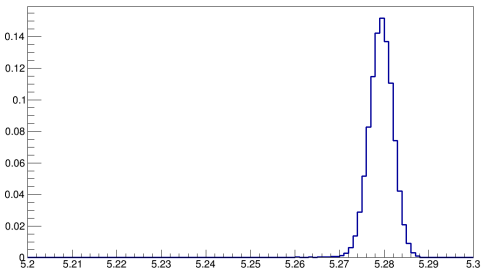
Figure 3.3: Analysis Variables for $B \to e^+e^-$ MC

For all signal modes, the KSFW likelihood ratio is heavily weighted toward the spherical side of the graph. The probability density function formed by the KSFW likelihood ratio is a flat, near-zero line until it reaches a value of 0.8 and begins to slope sharply upward.

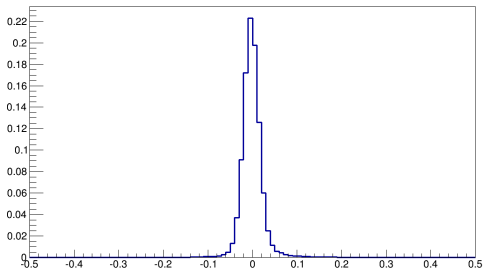
The $\Upsilon(4S)$ bottomonium resonance is a spin-1 resonance that decays into two spin-0 B mesons. This results in a $1-\cos^2(\theta_B)$ angular distribution with respect to the beam axis for true $ee \to \Upsilon(4S) \to B^0\bar{B}^0$ events. Because the absolute value of $\cos\theta_B$ is used, the range for this variable is zero to one.

The signal probability density function for Δz is a wide Gaussian distribution. The width is determined by the lifetime of the B meson. Eight percent of signal events fall outside the limit $|\Delta z| \leq 0.2$ cm. The majority of events with high $|\Delta z|$ values contained failed vertex reconstructions for one or both B mesons for which the value of Δz defaulted to 999 cm. Events with $|\Delta z| > 0.2$ cm are analyzed without this variable.

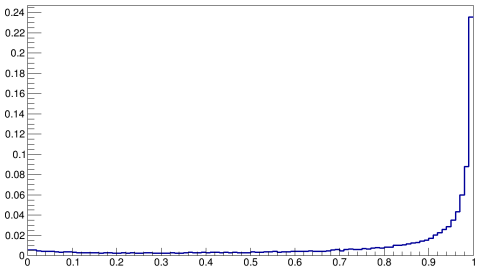
Figure 3.3f shows the eID values for the daughter tracks of a signal B meson. Figure 3.4f shows the μID values for signal tracks. Because $B^0 \to e^{\pm}\mu^{\mp}$ contains both an electron and a muon, figure 3.5f shows either the eID or the μID values for the signal tracks. For all signal modes, the ℓID PDFs are densely populated near one indicating that the daughter particles of the candidate B meson are considered leptons.



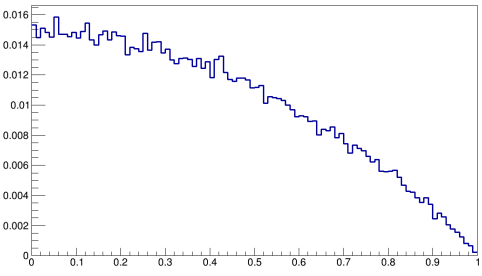
(a) M_{bc}



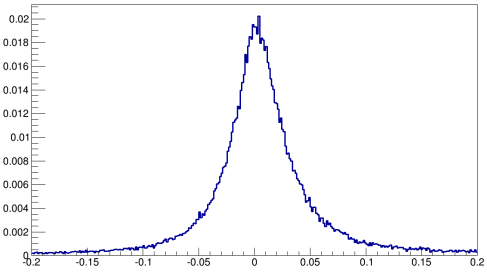
(b) ΔE



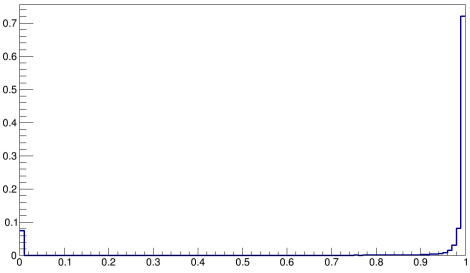
(c) KSFW Likelihood Ratio



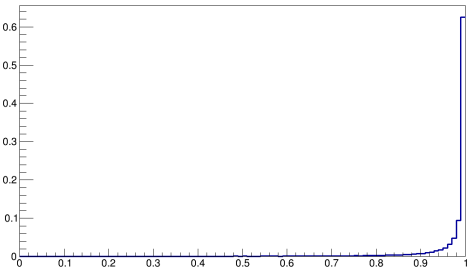
(d) $\cos\theta_B$



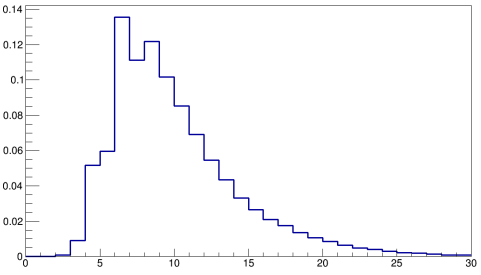
(e) Δz



(f) Muon PID

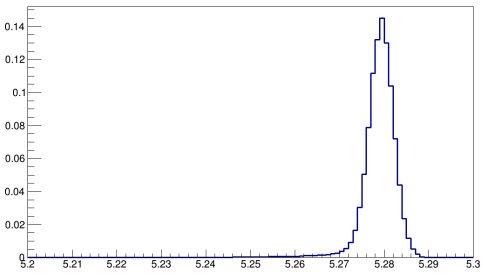


(g) $\mathcal{L}_{\pi/K}$

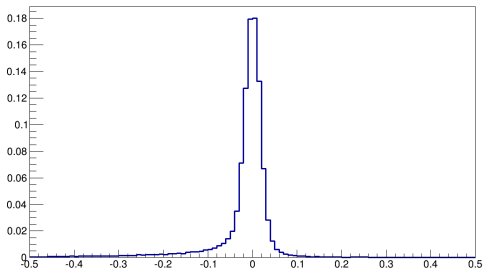


(h) Number of Charged Tracks

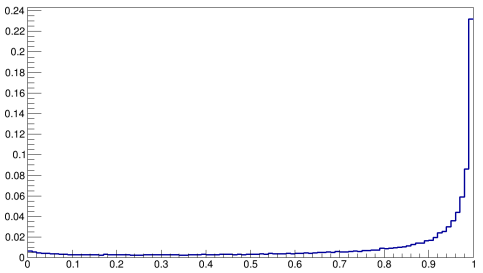
Figure 3.4: Analysis Variables for $B \to \mu^+\mu^-$ MC



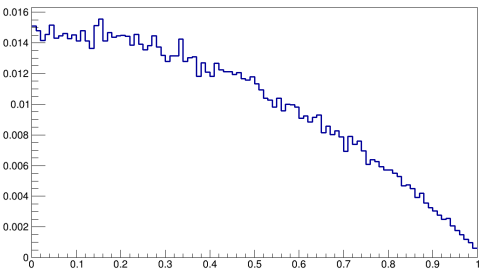
(a) M_{bc}



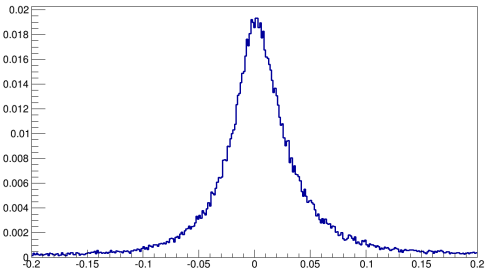
(b) ΔE



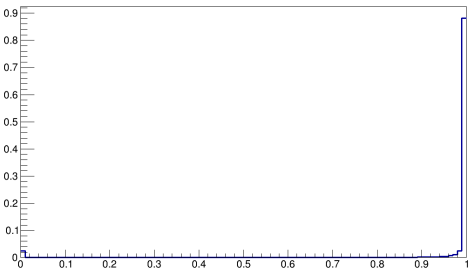
(c) KSFW Likelihood Ratio



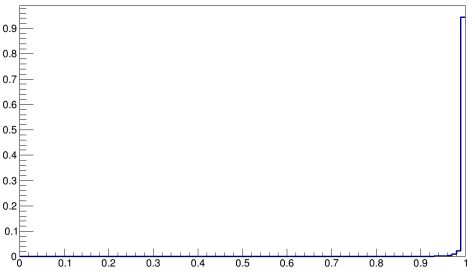
(d) $\cos\theta_B$



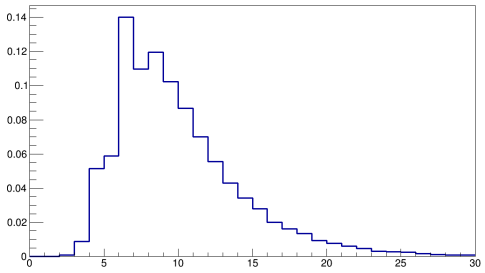
(e) Δz



(f) Lepton PID



(g) $\mathcal{L}_{\pi/K}$



(h) Number of Charged Tracks

Figure 3.5: Analysis Variables for $B \to e^{\pm}\mu^{\mp}$ MC

The measurement $\mathcal{L}_{\pi/K}$ is shown for each signal mode in subfigure (g). Since leptons are light, this plot is heavily populated near one. Consequently, placing a cut on this variable can be used to eliminate background modes involving kaons without significantly impacting the signal efficiency.

Each signal event contains two charged tracks on the signal side plus some number of charged tracks on the tag B side, depending on its generic decay mode. The number of charged tracks (N_{track}) in a signal event peaks at seven in the simulation. The frequency slopes downward as the number of charged tracks increases. The frequency of events with fewer than seven charged tracks drop drastically as the number of charged tracks decreases.

3.3 Continuum Monte Carlo

For the continuum study, one full stream of each charm and uds background Monte Carlo were used. Additionally, data sets specifically identified as important to this study were analyzed. They include tau-pair production events ($e^+e^- \to \tau\bar{\tau}$) and selected other QED events ($e^+e^- \to e^+e^-e^+e^-$ and $e^+e^- \to e^+e^-\mu^+\mu^-$).

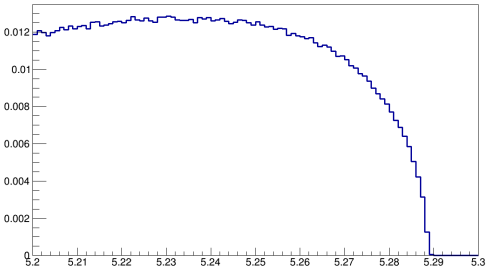
3.3.1 $e^+e^- \to q\bar{q}$ Monte Carlo

The set of data referred to as continuum data represents the interactions $e^+e^- \to q\bar{q}$ where $q = u, d, s, c$. In Belle, these are split into two groups known as charm continuum ($e^+e^- \to c\bar{c}$) and uds continuum ($e^+e^- \to q\bar{q}$ where $q = u, d, s$).

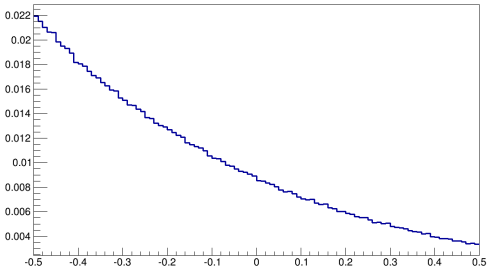
The probability density functions (PDFs) for charm and uds continuum are shown in Figures 3.6 and 3.7, respectively. For continuum events, the beam constrained mass (M_{bc}) is modeled as an ARGUS distribution [90] as opposed to the Gaussian distribution of the signal Monte Carlo. Instead of peaking about zero, the energy difference (ΔE) is a featureless shape that can be modeled by a polynomial function.

Due to the jet-like event topology, the KSFW likelihood ratio is heavily populated near zero for both the charm and uds continuum. The difference in the KSFW likelihood ratio distributions for signal and continuum events provides a powerful tool for continuum suppression. Furthermore, the continuum events emerge uniformly into the solid angle which creates a flat distribution for $\cos\theta_B$. The Δz distribution for continuum events forms a Gaussian, but without the lifetime of the B meson as a contributing element, it exhibits a sharper peak than for the signal.

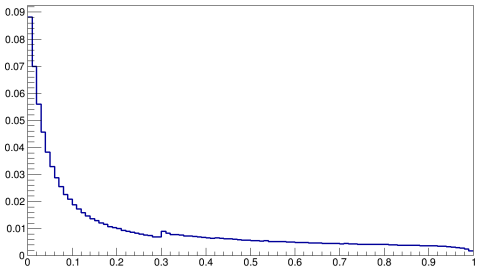
The lepton particle identification number classifies most particles in the continuum decays



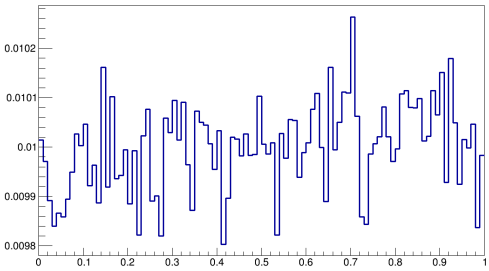
(a) M_{bc}



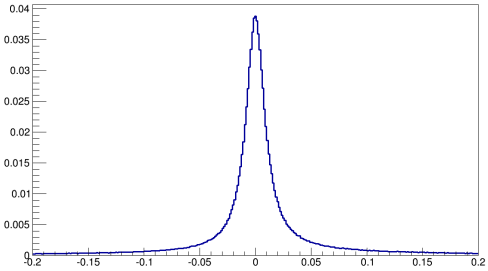
(b) ΔE



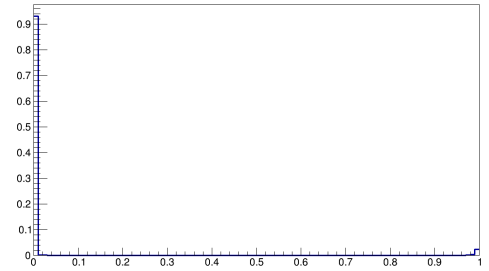
(c) KSFW Likelihood Ratio



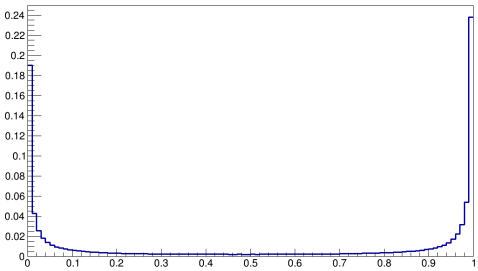
(d) $\cos\theta_B$



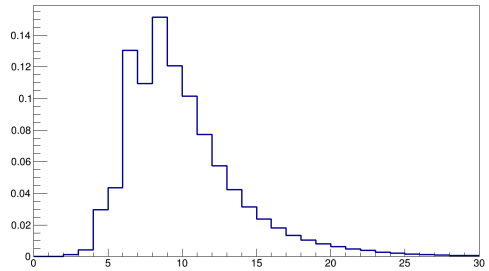
(e) Δz



(f) Lepton PID

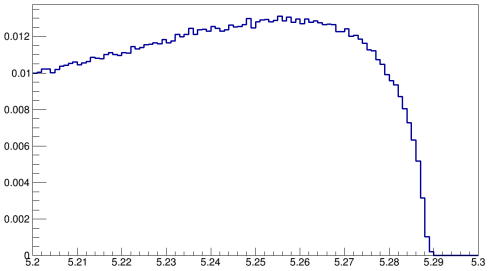


(g) $\mathcal{L}_{\pi/K}$

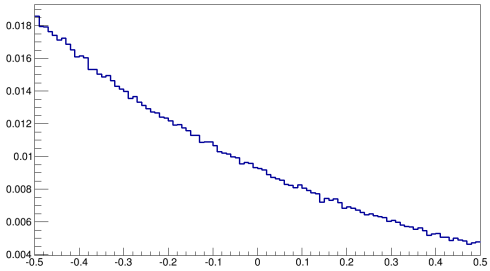


(h) Number of Charged Tracks

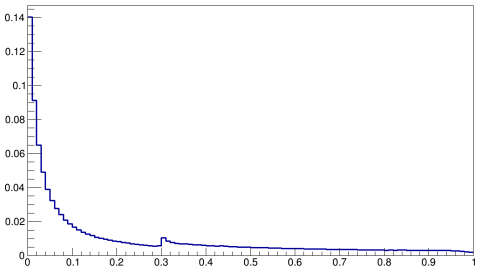
Figure 3.6: Analysis Variables for $e^+e^- \to c\bar{c}$ MC



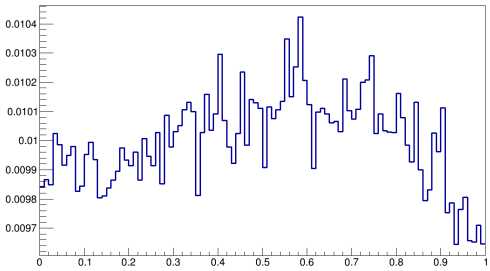
(a) M_{bc}



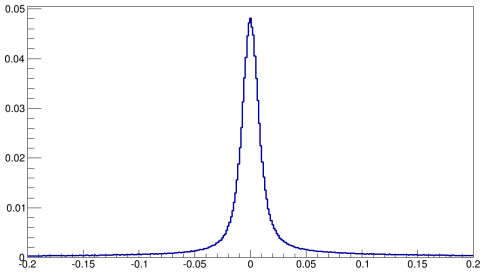
(b) ΔE



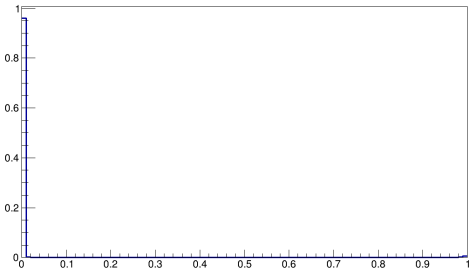
(c) KSFW Likelihood Ratio



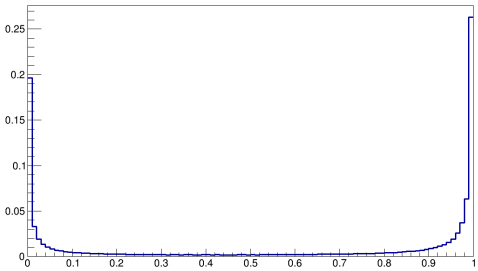
(d) $\cos\theta_B$



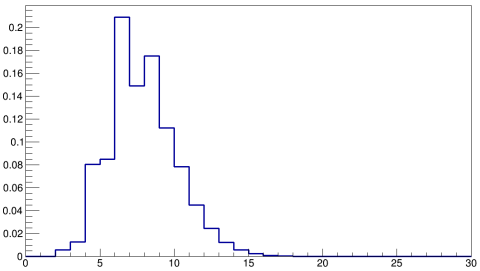
(e) Δz



(f) Lepton PID



(g) $\mathcal{L}_{\pi/K}$



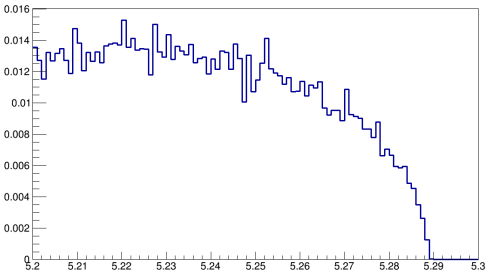
(h) Number of Charged Tracks

Figure 3.7: Analysis Variables for $e^+e^- \to q\bar{q}$ $(q = u, d, s)$ MC

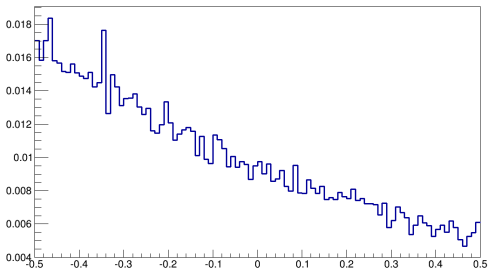
as hadrons. $\mathcal{L}_{\pi/K}$ shows pion-like and kaon-like particles in nearly equal proportions while the signal Monte Carlo events were placed on the pion-like end of the scale. The charged track distribution shows that most events have at least four charged tracks.

3.3.2 Tau-pair Monte Carlo

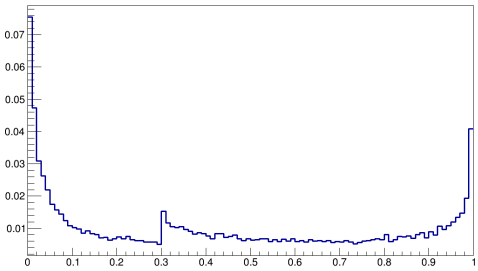
The tau production Monte Carlo data show variables similar to the continuum quark production events, but with a few significant exceptions. The KSFW likelihood ratio has a significant portion of events landing in the spherically shaped region, meaning that this variable will be less effective at suppressing the tau-pair background. The lepton identification also shows a number of events with high values of ℓID. Since the fake rate for identification is low, it is likely that these are leptons arising from a secondary decay of the tau particles. $\mathcal{L}_{\pi/K}$ shows a distribution similar to that of signal events. Finally, the number of charged tracks is lower for tau production events, peaking at only four charged tracks per event. The PDFs of the tau production variables are shown in Figure 3.8. One stream of tau-pair Monte Carlo was used in this analysis.



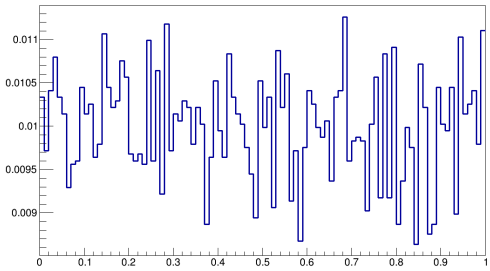
(a) M_{bc}



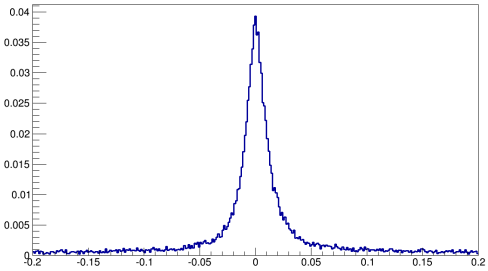
(b) ΔE



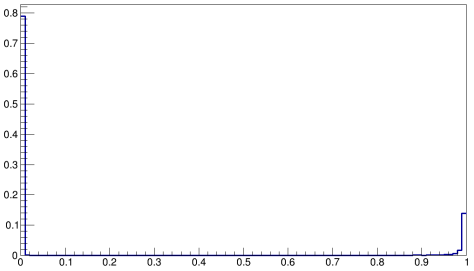
(c) KSFW Likelihood Ratio



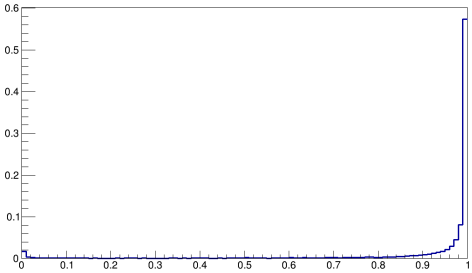
(d) $\cos\theta_B$



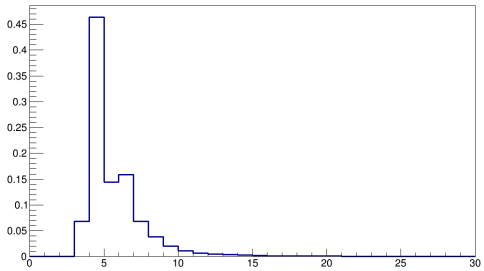
(e) Δz



(f) Lepton PID



(g) $\mathcal{L}_{\pi/K}$



(h) Number of Charged Tracks

Figure 3.8: Analysis Variables for $q\bar{q} \to \tau\bar{\tau}$ MC

3.3.3 QED Monte Carlo

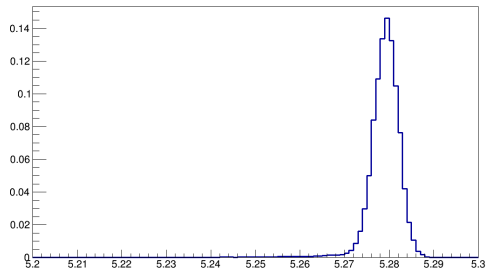
The QED Monte Carlo events of interest include the interactions $e^+e^- \to e^+e^-e^+e^-$ and $e^+e^- \to e^+e^-\mu^+\mu^-$. Ten streams of $e^+e^- \to e^+e^-e^+e^-$ and five streams of $e^+e^- \to e^+e^-\mu^+\mu^-$ were analyzed. Because the candidate B mesons don't satisfy the ΔE selection cuts for this analysis, no significant background contribution was found from the QED events.

3.4 $B\bar{B}$ Background Monte Carlo

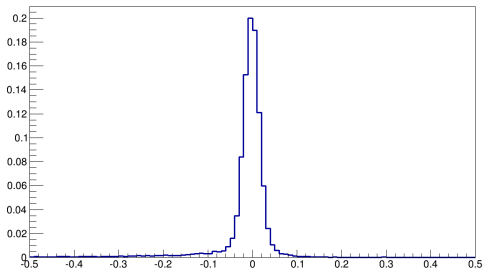
3.4.1 $B \to hh$ Monte Carlo

Three $B\bar{B}$ decay modes were identified as being significant backgrounds: $B \to \pi^+\pi^-$, $B \to K^-\pi^+$, and $B \to K^+K^-$. One million events for each decay mode were used in the analysis. This corresponds to 250 streams of $B \to \pi^+\pi^-$, 66 streams of $B \to K^-\pi^+$, and 10,000 streams of $B \to K^+K^-$ when compared with the Belle data set.

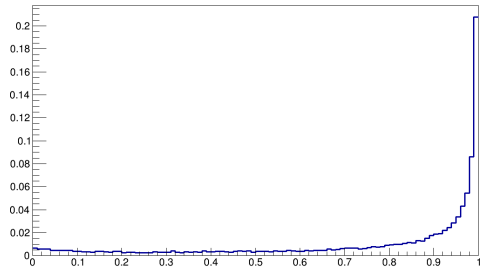
$B \to \pi^+\pi^-$ analysis variables are shown in Figure 3.9. $B \to \pi^+\pi^-$ is a two-body decay with kinematics similar to the signal modes. Additionally, pions and muons have similar masses. Consequently, $B \to \pi^+\pi^-$ has many PDFs in common with signal Monte Carlo. M_{bc} is a Gaussian distribution centered at 5.28 GeV/c^2. ΔE is a Gaussian centered at 0 GeV. Like all $B\bar{B}$ events, the KSFW likelihood ratio strongly favors a spherical shape. Δz, $\cos\theta_B$,



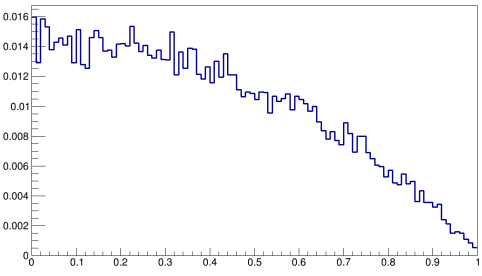
(a) M_{bc}



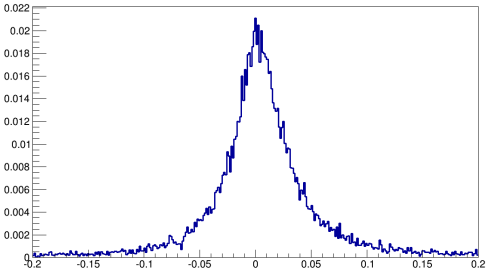
(b) ΔE



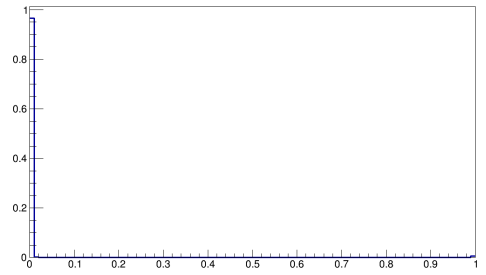
(c) KSFW Likelihood Ratio



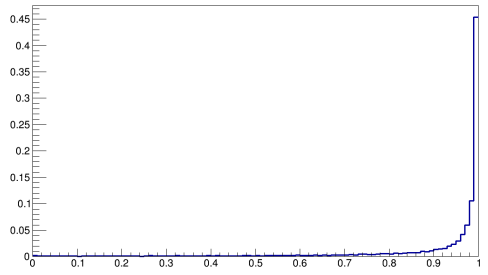
(d) $\cos\theta_B$



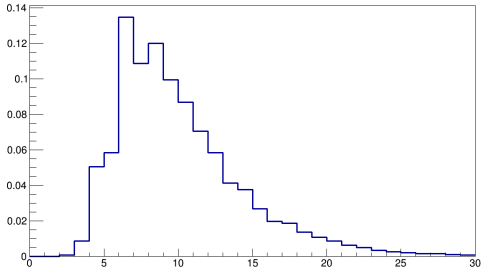
(e) Δz



(f) Lepton PID



(g) $\mathcal{L}_{\pi/K}$



(h) Number of Charged Tracks

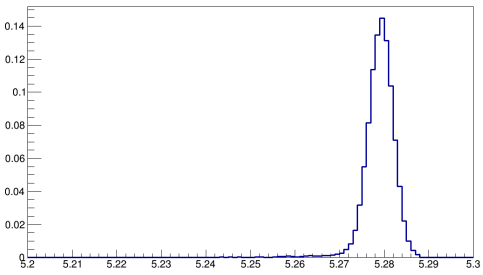
Figure 3.9: Analysis Variables for $B \to \pi^+\pi^-$ MC

$\mathcal{L}_{\pi/K}$, and N_{track} are all similar to the probability density functions of the signal events, particularly $B \to \mu^+\mu^-$. The only variable significantly different from signal MC is the lepton identification number. Due to the low muon-pion fake rate, $B \to \pi^+\pi^-$ corresponds to low values of ℓID.

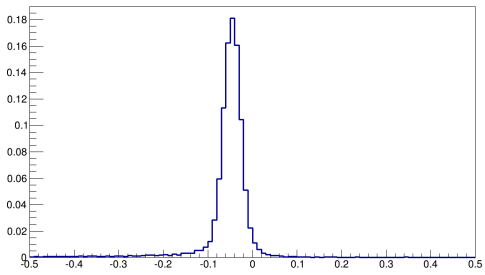
$B \to K^-\pi^+$ behaves similarly to $B \to \pi^+\pi^-$, but the kaon does provide key differences. The differences in mass cause the ΔE plot to be shifted to the left because the kaon is misassigned a muon or electron mass. Also, the $\mathcal{L}_{\pi/K}$ shows equal particles that are pion-like or kaon-like. The $B \to K^+K^-$ decay mode takes these differences one step further. The ΔE Gaussian is shifted farther leftward. Virtually all candidate B meson daughters fall on the kaon-like end of the $\mathcal{L}_{\pi/K}$ ratio.

3.4.2 Rare Decay Monte Carlo

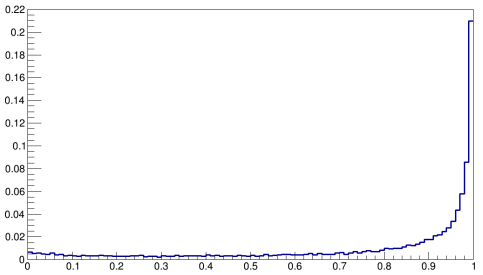
To determine if any other $B\bar{B}$ modes would play a significant role in the background, the mixed-neutral rare Monte Carlo and charged rare Monte Carlo data sets were studied. The charged rare MC contains data equivalent to 50 streams of B^+B^- events but with only the rare decay modes included. The total simulated branching fraction is 3×10^{-3}. The mixed-neutral rare MC contains 50 streams of rare $B^0\bar{B}^0$ events (incorporating B–$\bar{B}$ mixing) representing a total simulated branching fraction of 4×10^{-3}. Following the application



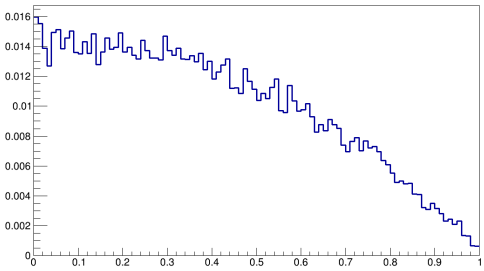
(a) M_{bc}



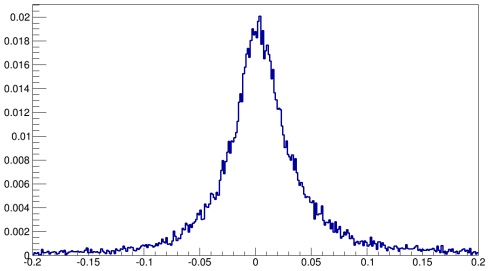
(b) ΔE



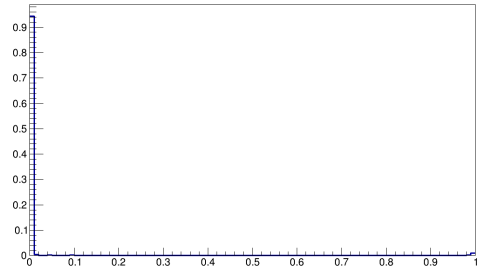
(c) KSFW Likelihood Ratio



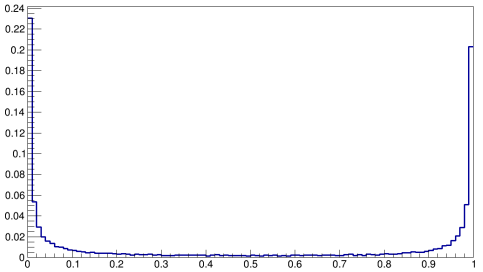
(d) $\cos\theta_B$



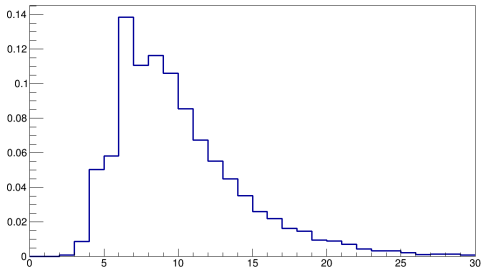
(e) Δz



(f) Lepton PID

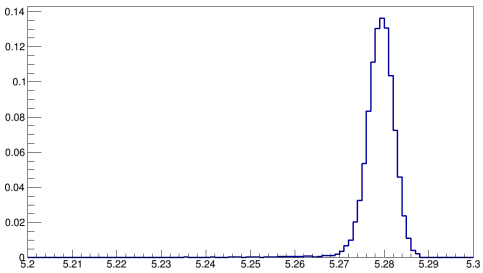


(g) $\mathcal{L}_{\pi/K}$

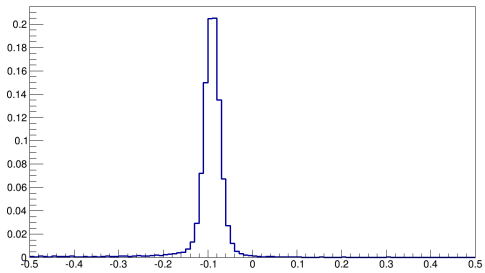


(h) Number of Charged Tracks

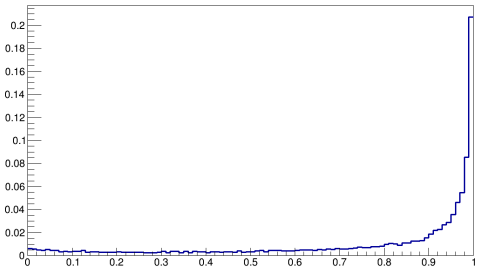
Figure 3.10: Analysis Variables for $B \to K^-\pi^+$ MC



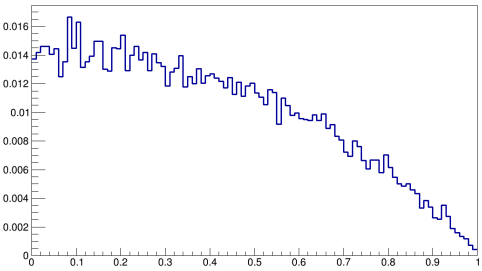
(a) M_{bc}



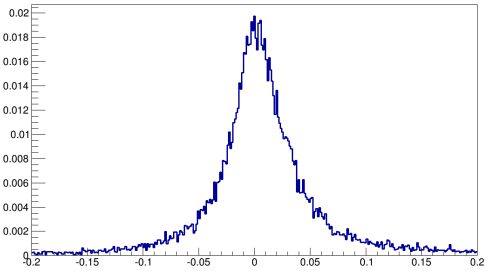
(b) ΔE



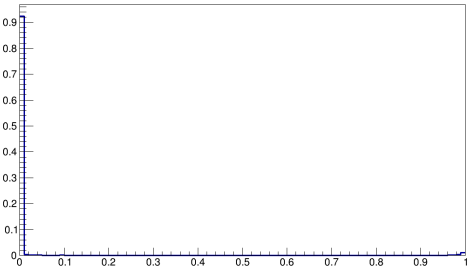
(c) KSFW Likelihood Ratio



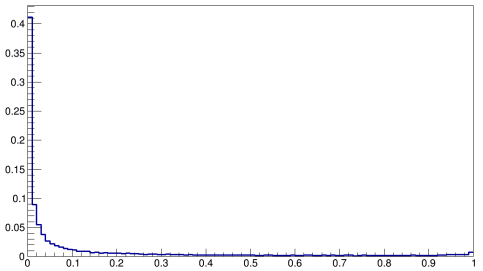
(d) $\cos\theta_B$



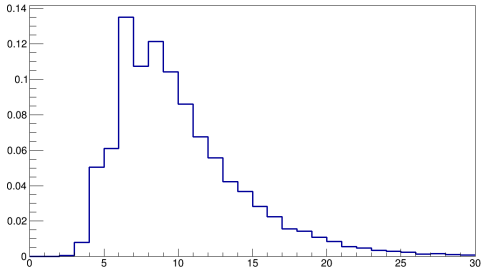
(e) Δz



(f) Lepton PID



(g) $\mathcal{L}_{\pi/K}$



(h) Number of Charged Tracks

Figure 3.11: Analysis Variables for $B \to K^+K^-$ MC

of cuts, the charged rare MC showed a contribution of less than 0.02 events for $B \to e^+e^-$ mode, less than 0.1 event for $B \to \mu^+\mu^-$ mode, and less than 0.06 events for $B \to e^\pm\mu^\mp$ mode. Due to the low number of events expected, the charged rare events were not included in further studies.

The mixed rare MC is more applicable to this study. It contains the two-body final states $B \to \pi^+\pi^-$, $B \to K^-\pi^+$, and $B \to K^+K^-$ which were already identified as important background modes. However, it does not contain the signal $B \to \ell^+\ell^-$ decay modes. Once the optimized cut values were applied to the mixed rare MC, the passing events matched the results for $B \to \pi^+\pi^-$ and $B \to K^-\pi^+$ in both the number of events and the variable shapes. From this result, it was determined that the significant $B\bar{B}$ background modes were already identified and treated in Section 3.4.1.

3.5 Conclusions

Monte Carlo data sets are used to optimize and validate analysis methods. The amount of Monte Carlo data used for each signal and background decay mode is itemized in Table 3.1. The amount of data used is measured either in millions of events or in streams.

Table 3.1: Monte Carlo Data Samples Used.

Mode	Number of Events	Number of Streams
$B \to e^+e^-$	2×10^6 events	3×10^4 streams
$B \to \mu^+\mu^-$	2×10^6 events	4×10^6 streams
$B \to e^\pm\mu^\mp$	2×10^6 events	9×10^5 streams
Charm Continuum		1 stream
UDS Continuum		1 stream
Mixed Rare		50 streams
Charged Rare		50 streams
$B \to \pi\pi$	1×10^6 events	250 streams
$B \to K\pi$	1×10^6 events	66 streams
$B \to KK$	1×10^6 events	1×10^4 streams
Tau Events		1 stream
$ee \to eeee$		10 streams
$ee \to ee\mu\mu$		5 streams

Chapter 4

Data Skimming

The purpose of skimming data is to create a subset of data that retains the highest signal efficiency possible while eliminating a large portion of the background events. Since the skim code must be applied to a large amount of data, it should be simple and efficient. It is applied to the full set of Belle data collected at the $\Upsilon(4S)$ resonance, which corresponds to 772×10^6 $B\bar{B}$ pairs, and to all of the MC data sets.

4.1 Skimming Criteria

This analysis begins with the HadronB data set (for real data) or the HadronB-selected events (for MC). See Appendix B for more details on the HadronB selection criteria. For the purposes of the $B \to \ell^+\ell^-$ skim, all charged particles are assumed to be leptons. Each

event is simultaneously analyzed in the modes $B \to e^+e^-$, $B \to \mu^+\mu^-$, and $B \to e^\pm\mu^\mp$. Two oppositely charged tracks are combined into a reconstructed B meson. To meet the skim criteria, the reconstructed B candidate must satisfy cuts on two kinematic variables: the beam constrained mass, M_{bc}, and the energy difference, ΔE.

For the analysis modes $B \to \mu^+\mu^-$ and $B \to e^\pm\mu^\mp$, any event with a candidate B meson that falls into the fitting range ($5.2\ \mathrm{GeV}/c^2 \leq M_{bc} \leq 5.3\ \mathrm{GeV}/c^2$ and $|\Delta E| \leq 0.5$ GeV) is selected.

For the $B \to e^+e^-$ analysis mode, any event with a B candidate falling in the range $5.2\ \mathrm{GeV}/c^2 \leq M_{bc} \leq 5.3\ \mathrm{GeV}/c^2$ and $-1.0\ \mathrm{GeV} \leq \Delta E \leq 0.5$ GeV is selected. The asymmetry in the ΔE range allows for electron bremsstrahlung. An event that is selected by any (or all) of the final-state modes is written to a ROOT file for additional analysis.

4.2 Bremsstrahlung Treatment

Because of their radiative behavior, electrons lose energy in the form of emitted photons. To help offset this effect and retain the highest signal efficiency possible, radiated photons were recombined with electrons, and the combined track was used for the remainder of the analysis. To be considered a radiated photon, the candidate photon must be within 50 milliradians of the candidate electron track as shown in Figure 4.1.

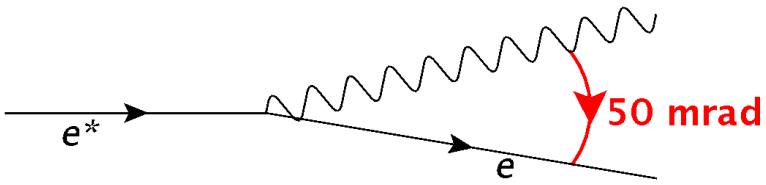


Figure 4.1: Diagram of photon criteria for bremsstrahlung treatment

A total of eleven percent of electrons showed some recovered energy. Following the bremsstrahlung treatment, the analysis range on ΔE was reduced to match the fitting region. Two percent of the signal events were moved from the rejected region (-1.0 GeV $\leq \Delta E \leq -0.5$ GeV) to the fitting region ($|\Delta E| \leq 0.5$ GeV).

Additionally, the combined electron-photon track was treated as one signal particle during the calculations of the Kakuno-Super-Fox-Wolfram (KSFW) moments. By combining the electron and photon into a single track, it provides a two-body final state in the signal group and removes the bremsstrahlung photon from the "other" group. Figure 4.2 shows one of the KSFW bins before and after this bremsstrahlung treatment. By using the bremsstrahlung treatment the signal shape was shifted slightly right, providing slightly more separation between the background and signal events.

The skim efficiencies for each real and MC data set are shown in Figure 4.3. For the modes with electron(s), the efficiencies were calculated after the bremsstrahlung treatment and the

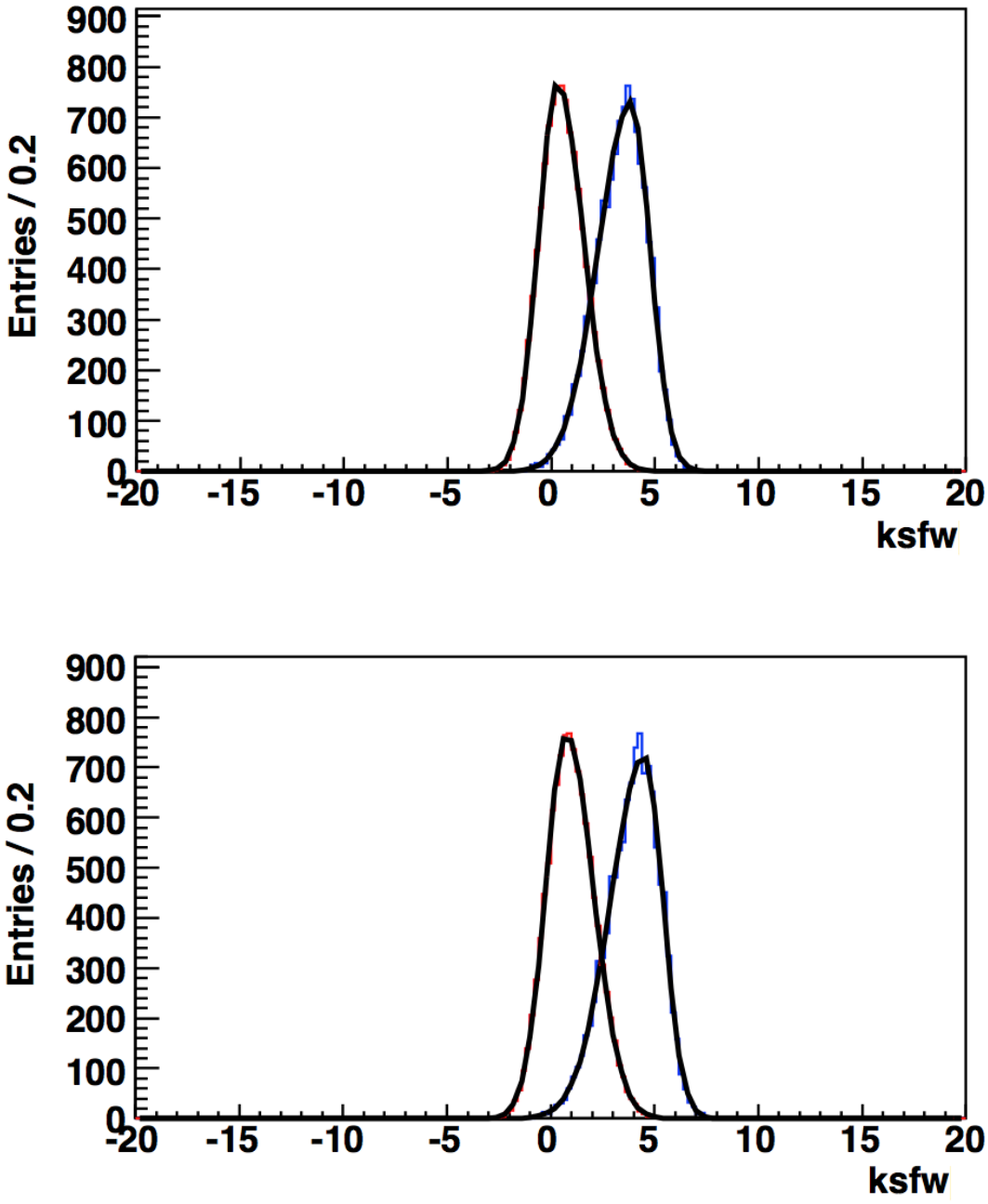


Figure 4.2: KSFW for one bin ($-0.5 \leq mm^2 < 0.3$) before (top) and after (bottom) bremsstrahlung treatment for the e^+e^- final state.

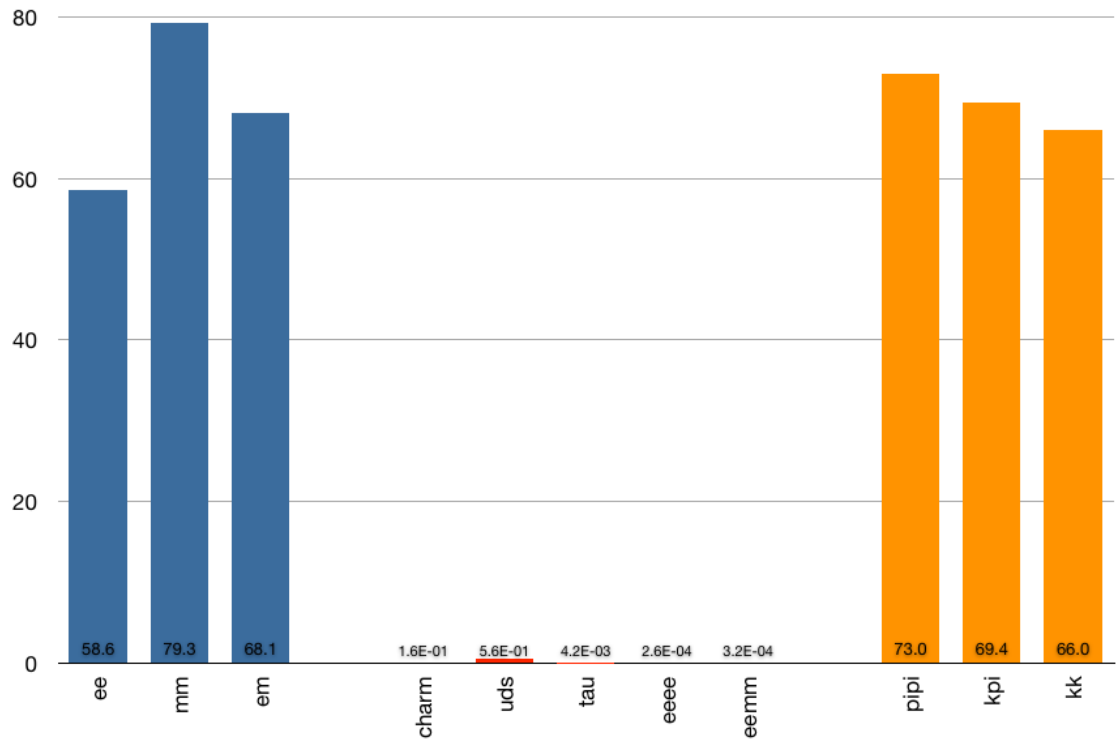


Figure 4.3: Skim efficiencies of signal and background modes

subsequent imposition of the fitting range. Because both signal tracks are required to carry very high momentum, most continuum background events are eliminated by the cuts on M_{bc} and ΔE.

The skim efficiencies for signal modes are 59% for $B \to e^+e^-$, 79% for $B \to \mu^+\mu^-$, and 68% for $B \to e^\pm\mu^\mp$ reflecting the lower passing rate for electrons relative to muons. The background $B\bar{B}$ modes of $B \to \pi\pi$, $B \to K\pi$, and $B \to KK$ had efficiencies of 73%, 69%, and 66%, respectively, closely mimicking $B \to \mu^+\mu^-$ which has similar kinematics and mass values. Continuum background shows efficiencies of 0.16% for charm and 0.56% for uds.

Table 4.1: Accumulated Efficiencies.

Mode	Skim Efficiency (%)
$B \to e^+e^-$	58.7
$B \to \mu^+\mu^-$	79.3
$B \to e^\pm\mu^\mp$	68.1
Charm Continuum	0.156
UDS Continuum	0.559
Tau Events	4.23×10^{-3}
$ee \to eeee$	2.59×10^{-4}
$ee \to ee\mu\mu$	3.2×10^{-4}
$B \to \pi\pi$	73.0
$B \to K\pi$	69.4
$B \to KK$	66.0
Mixed Rare	11.3
Charged Rare	3.7

Only 4×10^{-3} % of τ-pair events and 3×10^{-4} % of QED events pass since these decay modes do not satisfy the kinematics of $B \to \ell^+\ell^-$ decays.

4.3 Conclusions

The efficiencies to this point are summarized in Table 4.1. They include skim efficiency plus bremsstrahlung recovery for electrons. The skim criteria prove most effective at eliminating QED and tau-pair backgrounds but also significantly reduce the number of continuum background Monte Carlo events.

Chapter 5

Event Reconstruction

Each reconstructed event that passes the skim criteria for one or more of the final-state modes is written to a ROOT file for analysis. Here, relevant variables are saved to a custom TTree and can be optimized easily by reanalyzing the file repeatedly in ROOT. ROOT is an object-oriented framework for large scale data analysis developed at CERN specifically for the needs of particle physics analysis. Using a ROOT file has several advantages over a traditional MDST file, including keeping all particles in an event, saving multiple analysis modes as different branches for a single event, and allowing more convenient optimization of cuts using ROOT rather than basf for the subsequent analysis. ROOT also lets the user run fitting on complex data sets, build toy Monte Carlo simulations, and create analysis plots within one program [91].

5.1 Event Reconstruction

Reconstruction of the mdst-format events is performed once during the skimming stage using the cabs3 module in the basf framework [92]. The reconstructed events are saved in ROOT format for subsequent analysis.

For each event, cabs3 assembles lists of muons, electrons, and gammas (including their charge conjugates) from the mdst tables [93]. Every charged track is saved as both an electron and a muon. Relevant variables related to the individual charged tracks are the charge, the χ^2 from the track reconstruction, the number of degrees of freedom in this reconstruction, the distance of closest approach of the track to the interaction point along the beam axis (dz) and in the transverse plane (dr), and the particle identification information: eID, μID, $\mathcal{L}_{\pi/K}$, $\mathcal{L}_{\pi/p}$. Variables available for photons are the momentum vector, the cluster's reconstruction quality, width, and the $E9/E25$ energy spread. Oppositely charged track pairs are combined into neutral "vees." If the parent vee appears to be photon-like and meets bremsstrahlung criteria, it may be treated as a photon (i.e., moved to the gamma list) and recombined with the electron track.

Once the electron, muon, and gamma particle lists have been created, cabs3 combines oppositely charged tracks into B candidates that satisfy the loose cuts of 5.2 GeV/$c^2 \leq M_{bc} \leq$ 5.3 GeV/c^2 and $|\Delta E| \leq 0.5$ GeV. The resulting B candidate list typically contains a given track

pair that is treated in multiple modes: $B \to e^+e^-$, $B \to \mu^+\mu^-$, and $B \to e^\pm\mu^\mp$. The event is saved to the ROOT file with each of these B candidates. Following B candidate reconstruction of charged track pairs, the bremsstrahlung treatment is implemented for electron tracks. Nearby photons are combined with their partner electrons and kinematic variables including M_{bc}, ΔE, four-momentum, and KSFW event shape are recalculated.

Then, several other calculations are performed. The signal and tagged B meson vertices are reconstructed from their charged tracks and used to determine the Δz value between the two vertices. Thrust and $q \cdot r$ are calculated. Finally, for Monte Carlo events, the generated decay mode details are obtained from the mdst table and compared to the cabs3 reconstructed B candidates. Signal Monte Carlo events are given a flag referred to as isTrueB, which is true if the B candidate decay mode matches the generated mode. This flag is used to select only true (simulated) signal events during cut optimization and while developing the selection criteria, preventing fake signal events from influencing the optimization of the selection criteria.

5.2 ROOT File

The output from the reconstruction code is a ROOT file. Since the reconstruction code performs no optimization and no cuts beyond the mild skimming, all selection criteria are executed on the ROOT file. This permits quick re-analysis and turnaround as each cut is

optimized. The ROOT file is structured so that these optimizations can be performed as efficiently as possible [94].

ROOT data is stored in TTree format. A TTree has a header with a name and title and a user-defined number of branches. Each branch can house structures, variables, objects, arrays, or data clones. It can also contain an independent number of subbranches which themselves accommodate multiple types of information [95]. The final level of data held beneath the subbranches is referred to as leaves. When writing a TTree, the branch buffers are filled with leaf data until the buffer is full, then the data is written to disk in bunches. This practice optimizes the TTree class to reduce disk space and enhance access speeds [96].

Figure 5.1 shows the TTree structure and headers for each event. The basic information saved for each event includes whether the event is simulated or real, the experiment number, the run number, the event number, the beam energy, the center of mass boost vector, and the number of particles saved to the cabs3 particle lists. The second branch lists every final-state particle saved and contains information about each track, gamma, and vee.

The most important branch written to the ROOT file contains information regarding each candidate B meson. The first leaf is the candidate mode. This flag can be used to access events as analyzed using a specific decay mode ($B \to e^+e^-$, $B \to \mu^+\mu^-$, or $B \to e^\pm\mu^\mp$). The inclusion of this flag allows the same reconstruction code to be used simultaneously for

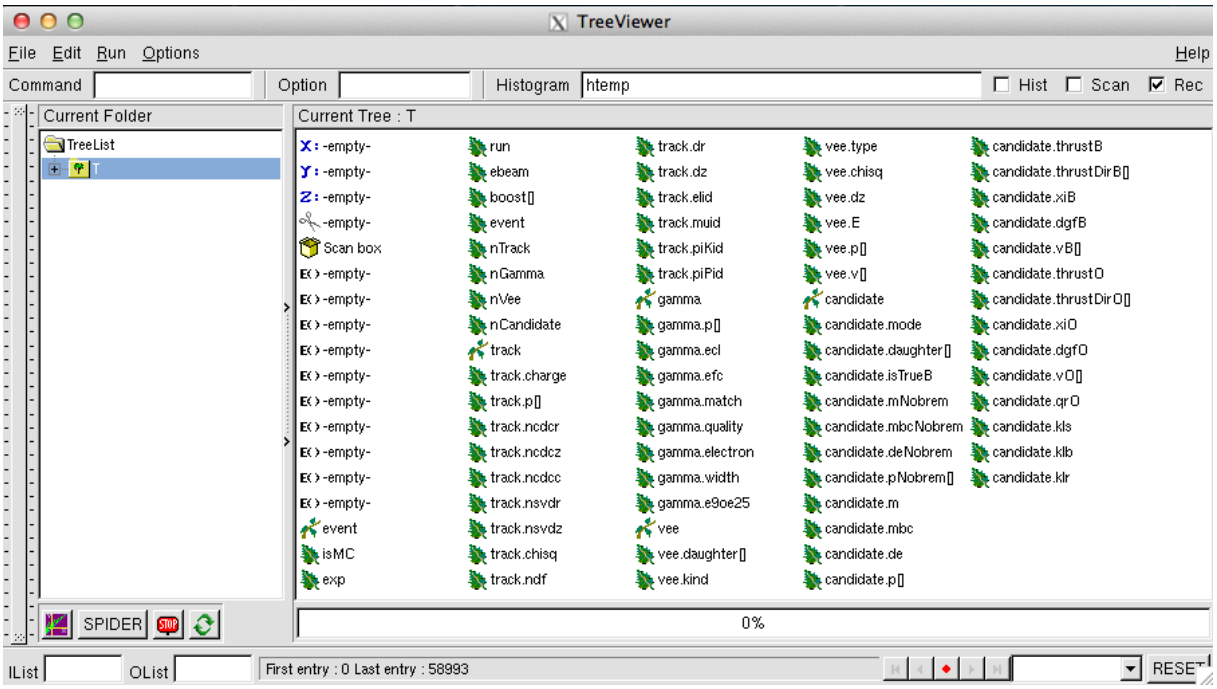


Figure 5.1: Format of ROOT TTree files

multiple modes. The candidate daughter leaves contain pointers to the candidate B meson's daughter tracks located within the track branch. The kinematic variables M_{bc}, ΔE, and momentum are saved prior to and following the bremsstrahlung treatment. This was included so that the bremsstrahlung-recovery method itself could be tested. The variables saved here were used to determine the number of electrons with radiated energy that could be recovered and the recovered energy spectrum. Finally, other variables considered during analysis to optimize selection criteria are saved.

5.3 Conclusions

The cabs3 framework provides a means to reconstruct B candidates with multiple hypotheses efficiently in an event from the Belle mdst files. These are written to ROOT files for subsequent analysis. By using the ROOT analysis package, the user can optimize selection criteria without redoing the slower and less flexible mdst-based analysis. Following the creation of the ROOT files, the next step in analysis is the optimization of the cut values to distinguish signal and background events.

Chapter 6

Optimization and Cuts

In this analysis, we reconstruct the following modes: $B^0 \to e^+e^-$, $B^0 \to \mu^+\mu^-$, $B^0 \to e^\pm\mu^\mp$. Inclusion of the charge conjugate modes is always implied. Before analysis, the relevant variables of each signal and background mode were fit using Monte Carlo events to construct the probability density functions (PDFs) used in this study. The fitted variables include M_{bc}, ΔE, KSFW likelihood ratio, $\cos\theta_B$, Δz, ℓID, $\mathcal{L}_{\pi/K}$, and N_{track}. The full set of variable fits including parameters is available in Appendix C. The selection criteria used to reconstruct candidate B mesons were determined through a series of optimizations performed on Monte Carlo events and a toy Monte Carlo sensitivity study. The sensitivity study simultaneously optimized the lepton identification variable cut values and the continuum suppression likelihood ratio cut value whilst determining the best likely outcome for the branching fraction.

Before the sensitivity study was performed, two variables proved useful for background suppression. They are N_{track} and the pion identification cut, $\mathcal{L}_{\pi/K}$. Here N_{track} represents the total number of charged tracks within an event as combined from both the reconstructed B candidate and the tagged B meson. $\mathcal{L}_{\pi/K}$ is the likelihood that a particle is pion instead of a kaon as defined the ratio by $\mathcal{L}_{\pi/K} = \frac{\mathcal{L}_\pi}{\mathcal{L}_\pi + \mathcal{L}_K}$.

6.1 N_{track} Cut

HadronB requires that an event have a minimum of three charged tracks. By requiring here that a candidate event has a minimum of five charged tracks, a portion of the non-$B\bar{B}$ background was eliminated, particularly in the production of $e^+e^- \rightarrow \tau^+\tau^-$ events. The background contribution from tau-pair events was cut in half. The QED background was also reduced. Meanwhile, ninety-five percent of the signal events passed this cut. The efficiencies for signal and background modes are listed in Table 6.1. For each mode, the first line contains the efficiency of the skim followed by the efficiency of the cut on the number of tracks. Below the N_{track} column header, the accumulated efficiency is the combined efficiency for both the skim and this cut.

Table 6.1: Accumulated Efficiencies after N_{track} Cut.

Mode		**Skim** (%)	N_{track} (%)
$B \to e^+e^-$	Individual ϵ	58.7	94.5
	Accumulated ϵ		55.4
$B \to \mu^+\mu^-$	Individual ϵ	79.3	94.5
	Accumulated ϵ		75.0
$B \to e^\pm\mu^\mp$	Individual ϵ	68.1	94.5
	Accumulated ϵ		64.4
Charm Continuum	Individual ϵ	0.156	97.1
	Accumulated ϵ		0.151
UDS Continuum	Individual ϵ	0.559	92.6
	Accumulated ϵ		0.518
Tau Events	Individual ϵ	4.23×10^{-3}	47.9
	Accumulated ϵ		2.03×10^{-3}
$ee \to eeee$	Individual ϵ	2.59×10^{-4}	97.3
	Accumulated ϵ		2.52×10^{-4}
$ee \to ee\mu\mu$	Individual ϵ	3.2×10^{-4}	44.3
	Accumulated ϵ		1.42×10^{-4}
$B \to \pi\pi$	Individual ϵ	73.0	94.4
	Accumulated ϵ		68.9

Continued on next page

Table 6.1 – *Continued from previous page*

Mode		**Skim** (%)	N_{track} (%)
$B \to K\pi$	Individual ϵ	69.4	94.3
	Accumulated ϵ		65.4
$B \to KK$	Individual ϵ	66.0	94.6
	Accumulated ϵ		62.4
Mixed Rare	Individual ϵ	11.3	93.6
	Accumulated ϵ		10.6
Charged Rare	Individual ϵ	3.70	97.8
	Accumulated ϵ		3.62

6.2 Pion ID Cut

The measurement $\mathcal{L}_{\pi/K}$ separates particles that behave like pions from particles that behave like kaons. Since leptons are light, they fall on the pion end of this plot and this variable can be used to eliminate background modes involving kaons without significantly impacting the signal efficiency. This cut value suppresses the background $B\bar{B}$ modes of $B^0 \to K\pi$ and $B^0 \to KK$.

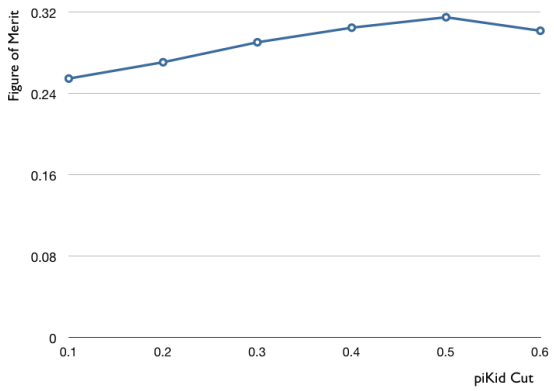


Figure 6.1: Figure of Merit for Pion ID Cut

To determine the optimal cut value, a figure of merit (FOM) was employed. Figures of merit were designed to obtain the smallest statistical error fraction $\delta N_S/N_S$ where N_S represents the number of signal events scaled by the measured limits on branching fraction (see Table 1.3). This is done by maximizing the value $N_S/\sqrt{N_S + N_B}$, where N_B is the number of background events. For each analyzed value of πID, the scaled number of passing events for both signal modes and the background modes was used to calculate the FOM value. The values studied were chosen based on the availability of systematic error tables as calculated by Belle's PID Joint Group.

The results of the figure of merit are shown in Figure 6.1. The peak at 0.5 indicates that this is the optimum value. Because 0.5 is often a default value for the $\mathcal{L}_{\pi/K}$ calculation when the particle identification is indeterminate, it rarely is used as a cut value. Further investigation showed that due to the high momentum required by this study, the selected data do not

exhibit an anomalous peak at 0.5. Therefore, the cut on $\mathcal{L}_{\pi/K}$ was placed here. Both tracks creating the candidate B meson must have $\mathcal{L}_{\pi/K} \geq 0.5$. The resulting efficiencies are listed in Table 6.2.

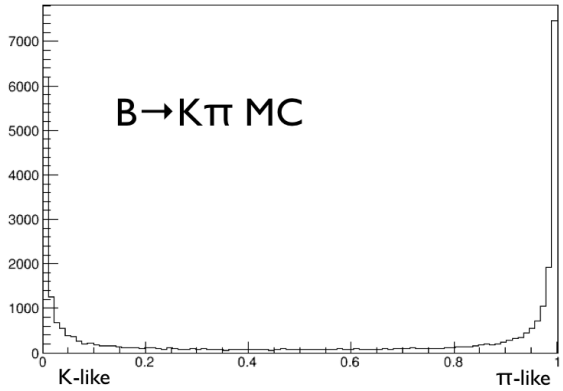


Figure 6.2: πID for $B \to K\pi$ Monte Carlo

The cut on $\mathcal{L}_{\pi/K}$ proved very effective at eliminating both $B^0 \to K\pi$ and $B^0 \to KK$, rejecting 90% of remaining $B^0 \to K\pi$ events and 99% of remaining $B^0 \to KK$ events. The resulting accumulated efficiency of $B^0 \to KK$ mode is less than one percent, rendering its contribution to background events negligible. Additionally, placing a cut on $\mathcal{L}_{\pi/K}$ reduced continuum background events by 64% in the case of charm and 61% for uds.

Table 6.2: Accumulated Efficiencies after Pion ID Cut.

Mode		**Skim** (%)	N_{track} (%)	**Pion ID** (%)
$B \to e^+e^-$	Individual ϵ	58.7	94.5	100
	Accumulated ϵ		55.4	55.4
$B \to \mu^+\mu^-$	Individual ϵ	79.3	94.5	96.2
	Accumulated ϵ		75.0	72.2
$B \to e^\pm\mu^\mp$	Individual ϵ	68.1	94.5	98.5
	Accumulated ϵ		64.4	63.4
Charm Continuum	Individual ϵ	0.156	97.1	36.1
	Accumulated ϵ		0.151	.055
UDS Continuum	Individual ϵ	0.559	92.6	38.6
	Accumulated ϵ		0.518	.200
Tau Events	Individual ϵ	4.23×10^{-3}	47.9	88.6
	Accumulated ϵ		2.03×10^{-3}	1.80×10^{-3}
$ee \to eeee$	Individual ϵ	2.59×10^{-4}	97.3	52.3
	Accumulated ϵ		2.52×10^{-4}	1.32×10^{-4}
$ee \to ee\mu\mu$	Individual ϵ	3.2×10^{-4}	44.3	74.6
	Accumulated ϵ		1.42×10^{-4}	1.06×10^{-4}
$B \to \pi\pi$	Individual ϵ	73.0	94.4	88.1
	Accumulated ϵ		68.9	60.7

Continued on next page

Table 6.2 – *Continued from previous page*

Mode		**Skim** (%)	N_{track} (%)	**Pion ID** (%)
$B \to K\pi$	Individual ϵ	69.4	94.3	10.5
	Accumulated ϵ		65.4	6.90
$B \to KK$	Individual ϵ	66.0	94.6	1.50
	Accumulated ϵ		62.4	0.935
Mixed Rare	Individual ϵ	11.3	93.6	44.5
	Accumulated ϵ		10.6	4.71
Charged Rare	Individual ϵ	3.70	97.8	56.1
	Accumulated ϵ		3.62	2.03

6.3 Continuum Suppression

$B\bar{B}$ events are isotropic in the $\Upsilon(4S)$ rest frame. Background continuum events $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) have a jet-like topology. The differences in shape are utilized to separate the signal and continuum events.

Several variables were examined for their effectiveness in suppressing continuum background. The variables that were found to be the most useful were the KSFW likelihood ratio, the

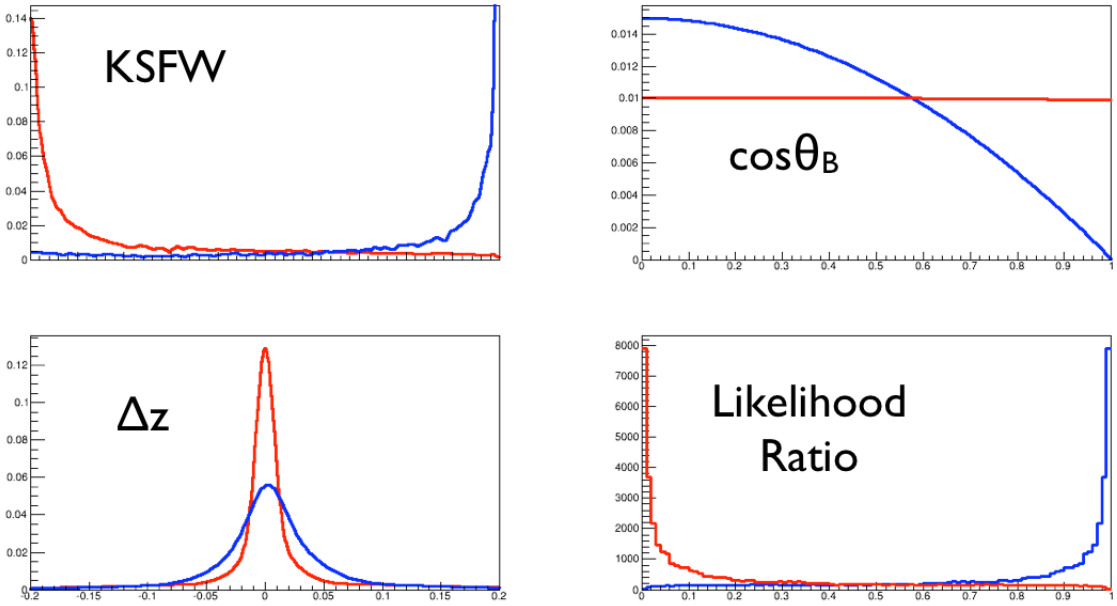


Figure 6.3: Continuum Suppression Variables for Events with a valid Δz

cosine of the angle between the beam axis and the B meson flight direction in the $\Upsilon(4S)$ rest frame ($cos\theta_B$), and the distance along the beam axis between the reconstructed vertices of the candidate B meson and other B meson (Δz). Figure 6.3 shows the variables used for continuum suppression with the signal MC data in blue and the continuum MC data in red. The three variables were then combined into a single likelihood ratio to distinguish signal and continuum events. The bottom right panel in Fig. 6.3 shows the results. The optimal cut value for the likelihood ratio was determined simultaneously with the particle identification cut values during the sensitivity study.

During Monte Carlo studies, it was found that 8% of signal events had a vertex that failed reconstruction and, therefore, had no meaningful Δz value. Rather than discarding these

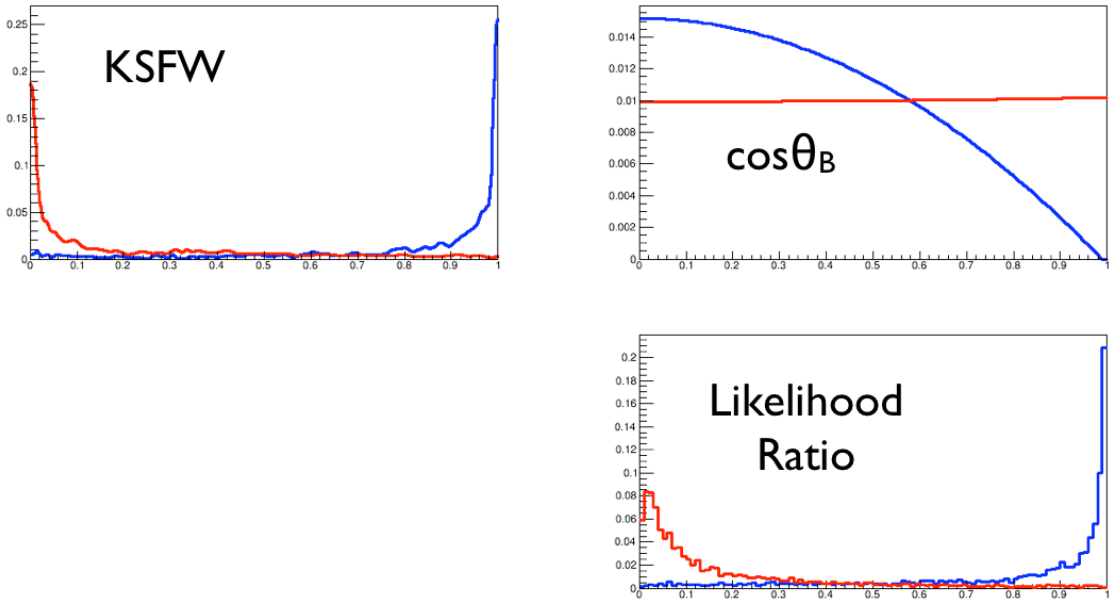


Figure 6.4: Continuum Suppression Variables for Events without a valid Δz

otherwise good signal events, they were analyzed as a separate category with their own continuum suppression likelihood ratio containing only the KSFW likelihood ratio and $cos\theta_B$.

6.4 Sensitivity Study

After the cuts on the number of charged tracks and on the pion/kaon identification, the next step was to perform a sensitivity study. The purpose of the sensitivity study was twofold; it simultaneously determined the best attainable limit on branching fraction, and it determined the optimized cuts on lepton particle identification and the continuum suppression likelihood ratio.

At this point, the events with and without a valid Δz were split for separate analysis. For each background and signal mode, a model was created in ROOT of M_{bc} and ΔE spanning the fitting region (5.2 GeV/c^2 $\leq$ M_{bc} $\leq$ 5.3 GeV/c^2 and $|\Delta E| \leq 0.5$ GeV). Using the expected number of background events, a toy Monte Carlo was generated. To reduce the effects of small numbers, the number of background events generated was multiplied by one hundred and treated as a Poisson distribution. Null signal was used. Then, the set of toy Monte Carlo data was fit to the signal and background PDFs and the signal yield extracted. Once 10,000 toy MC fits were performed the 90% confidence level upper limit ($UL90$) was found and used to calculate the limit on branching fraction (BF_{UL})

$$BF_{UL} = \frac{UL90}{100\epsilon_{sig}N_{events}} \tag{6.1}$$

where the BF_{UL} is the best possible upper limit for this study, $UL90$ is the point at which 90% of the positive signal yield is contained, ϵ_{sig} is the signal efficiency, and N_{events} is the total number of $B\bar{B}$ events in the Belle data sample. One hundred is included in the denominator to account for the multiplier in the number of background events generated. The full procedure is then repeated for different cut values of lepton ID and continuum suppression likelihood ratio. A separate sensitivity study was performed for each signal mode and for events with and without a valid Δz.

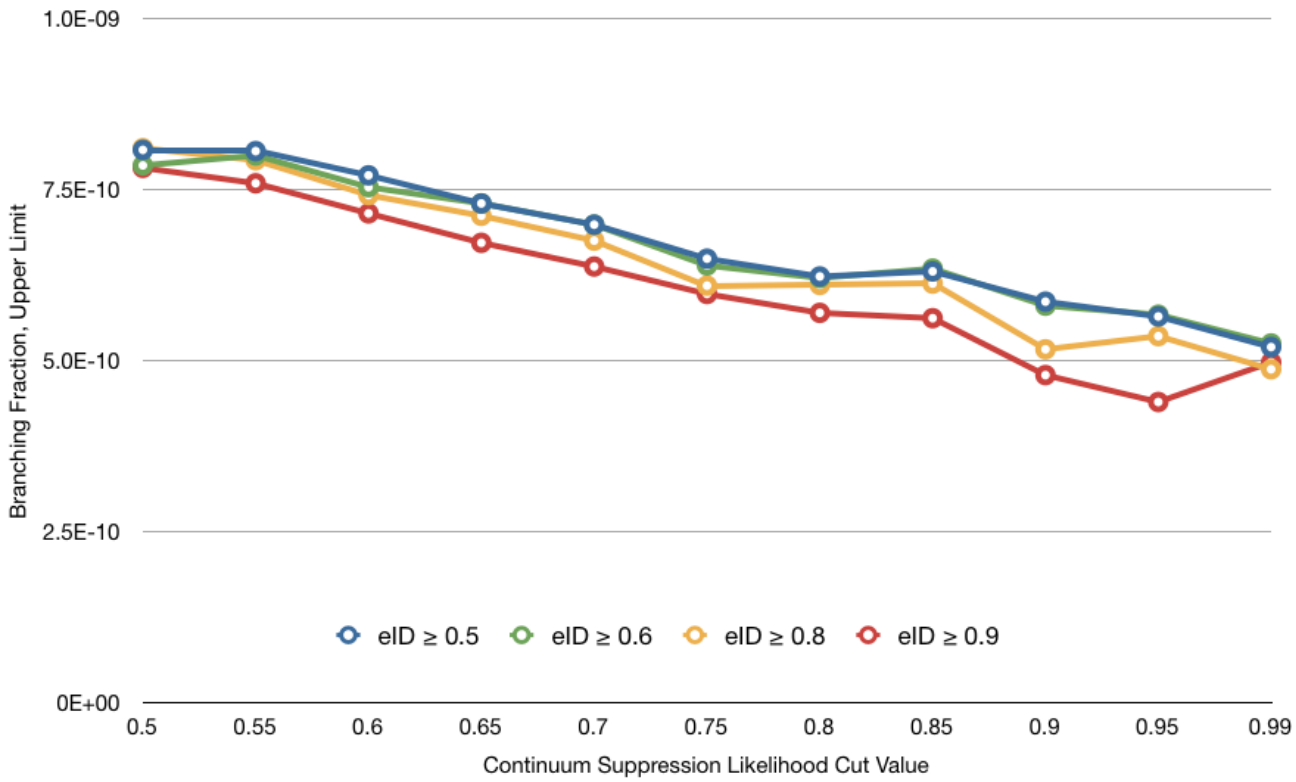


Figure 6.5: Branching Fraction Upper Limits for $B \to ee$

The results for $B \to ee$ are shown in Figure 6.5. In this chart, the horizontal axis represents the cut value of the continuum suppression likelihood ratio. Each colored line represents a different cut value for the electron identification. Both signal tracks must pass the cut value. The resulting upper limit of branching fraction is shown along the vertical axis. The lowest value of the branching fraction upper limit is located at the optimized combination of cut values for continuum suppression likelihood ratio and electron identification. The most advantageous cut values for $B \to ee$ mode are an electron identification cut at 0.9 and continuum suppression likelihood ratio cut of 0.95. By implementing these cut values, the

resulting branching fraction upper limit is 4.4×10^{-10} at 90% C.L.

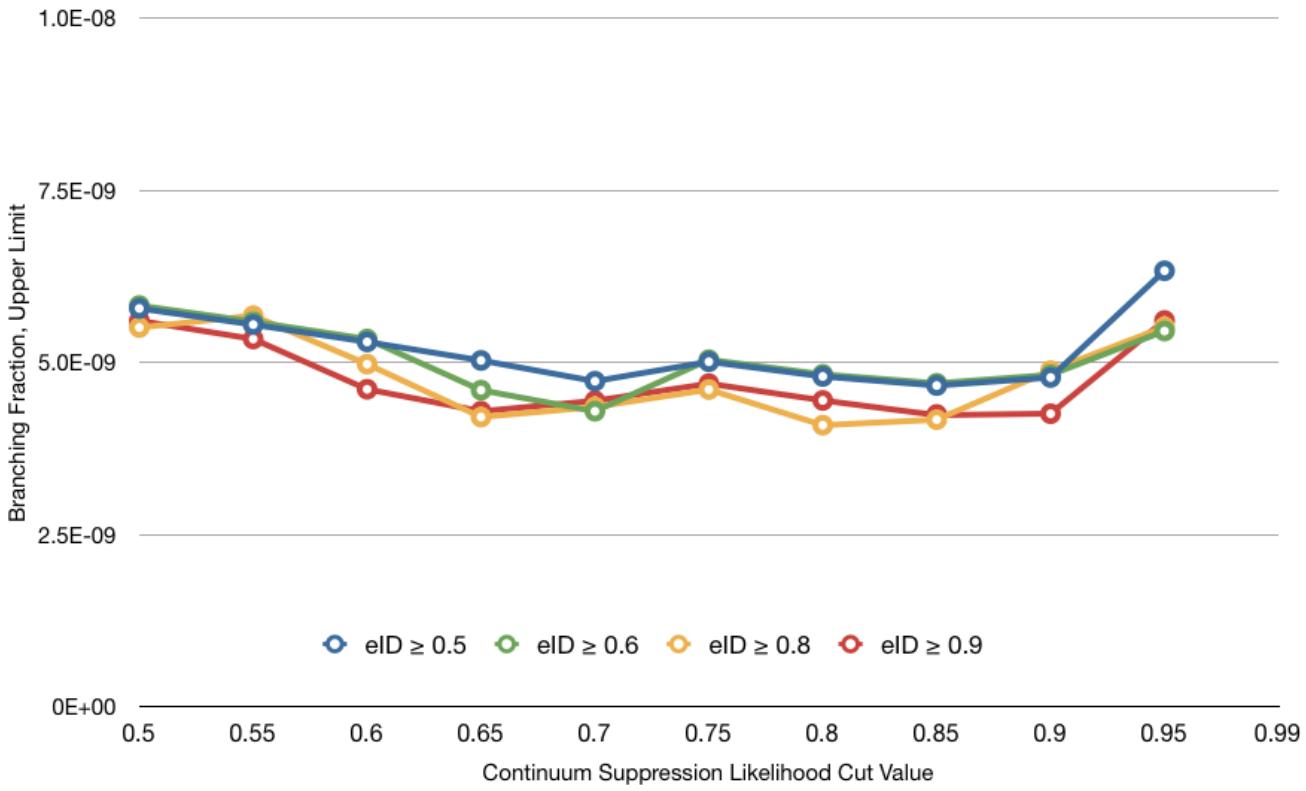


Figure 6.6: Branching Fraction Upper Limits for $B \to ee$ events without Δz

Figure 6.6 shows the results of the sensitivity study for $B \to ee$ mode events without a valid Δz value. This graph shows a large, relatively flat area. Based on the resulting branching fraction limits as well as for the sake of consistency, an electron identification cut was placed at $eID \geq 0.9$. Along the continuum suppression likelihood ratio axis, $LR \geq 0.85$ was chosen. The resulting upper limit of branching fraction is 4.2×10^{-9}.

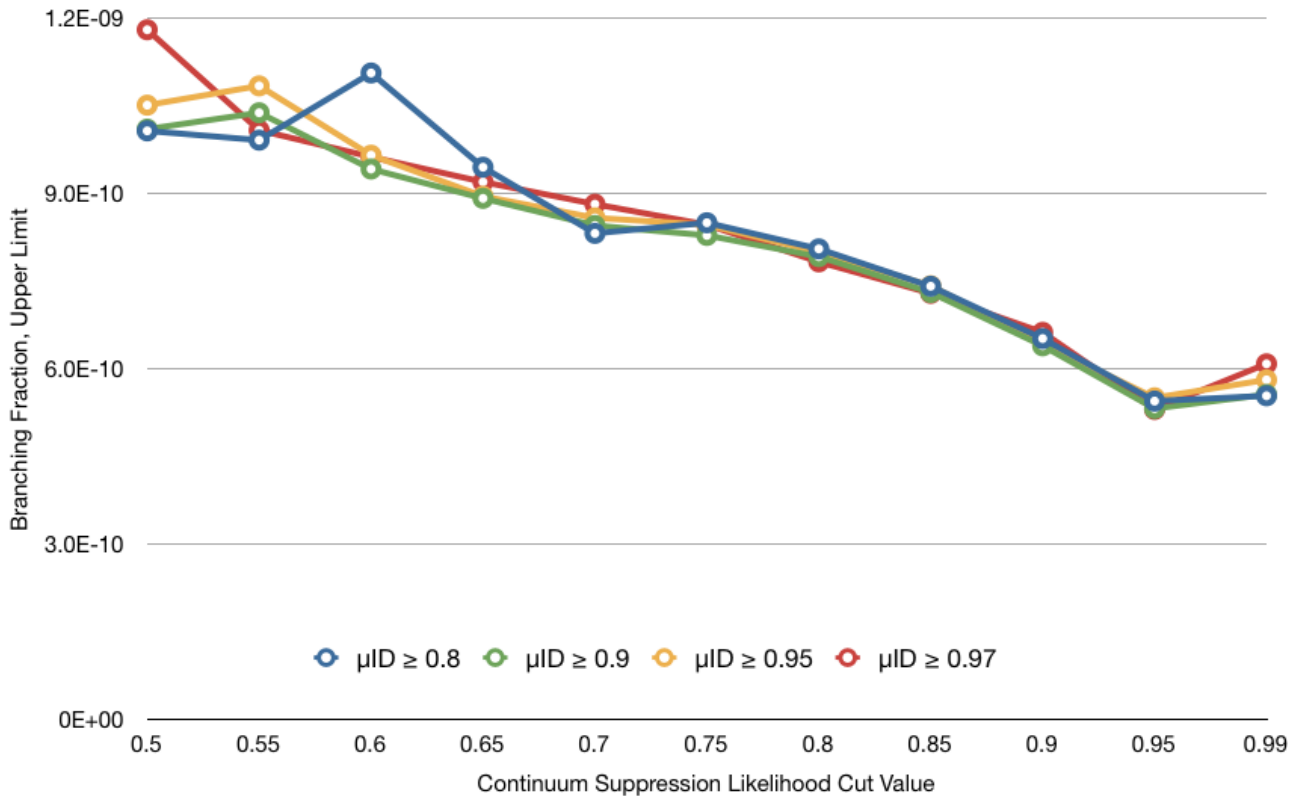


Figure 6.7: Branching Fraction Upper Limits for $B \to \mu\mu$

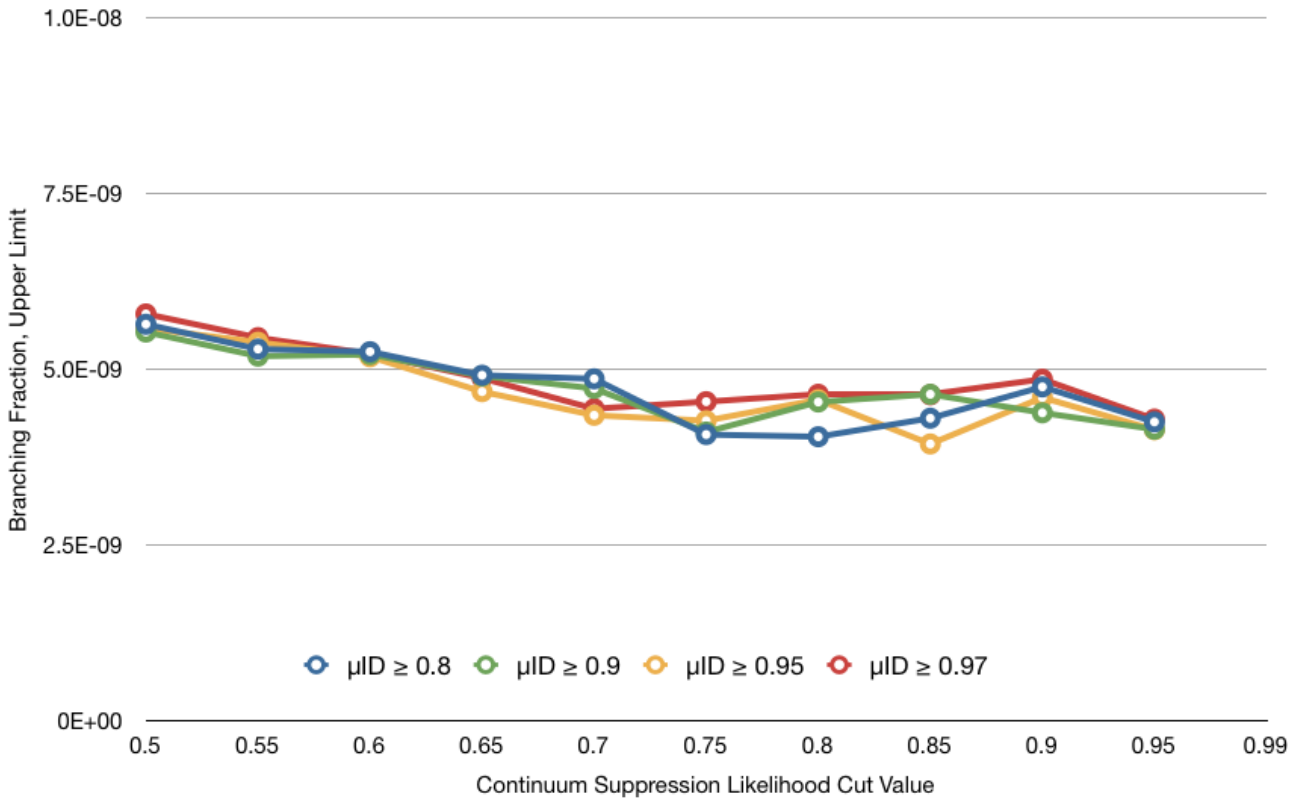
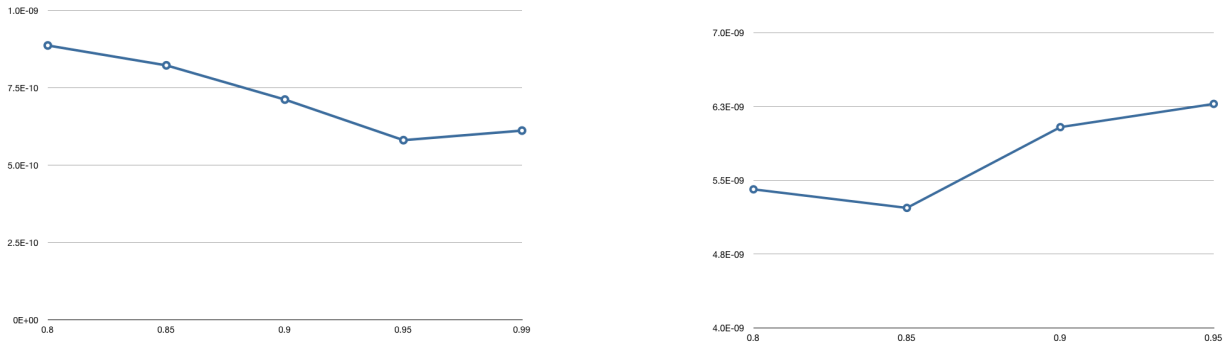


Figure 6.8: Branching Fraction Upper Limits for $B \to \mu\mu$ events without Δz

The results of $B \to \mu\mu$ are shown in Figure 6.7. This mode was analyzed in the same manner as $B \to ee$ but with the $B\bar{B}$ background modes of $B \to \pi\pi$ and $B \to K\pi$ fit separately from the other background modes because of the similarity of the masses of the muons and pions. The resulting upper limits of branching fraction show that cuts placed on muon identification at $\mu ID \geq 0.9$ and continuum suppression likelihood ratio value at $LR \geq 0.95$ provide the best limit, consistent with the results of $B \to ee$ mode. These cuts produce a branching fraction upper limit of 5.3×10^{-10}.



(a) Events with Δz (b) Events without Δz

Figure 6.9: Branching Fraction Upper Limits for $B \to e\mu$ events

Figure 6.8 displays the results of the sensitivity study for $B \to \mu\mu$ mode for events without a valid Δz value. Like the $B \to ee$ without Δz mode, there exists a large flat area with several equivalent choices for the cut values. The cuts were placed at $\mu ID \geq 0.9$ and $LR \geq 0.85$ which produces an upper limit on branching fraction of 4.2×10^{-9}.

With the cuts on lepton identification established and shown to be consistent across the analysis modes, $B \to e\mu$ mode was analyzed in a truncated fashion. The results are shown in Figure 6.9. This mode was analyzed only with the lepton identification cut (for both electrons and muons) placed at $\ell ID \geq 0.9$. For events with a valid Δz value, placing a cut at $LR \geq 0.95$ produces a branching fraction limit of 5.8×10^{-10}. For events without a valid Δz value, placing a cut at $LR \geq 0.85$ produces a branching fraction limit of 5.2×10^{-9}. The results of the sensitivity study are summarized in Table 6.3.

Table 6.3: Branching Fraction Upper Limits

Mode		Lepton ID Cut	LR Cut	Signal Efficiency	BF_{UL}
$B \to e^+e^-$	with Δz	$eID \geq 0.9$	$LR \geq 0.95$	21.2 %	4.4×10^{-10}
	without Δz	$eID \geq 0.9$	$LR \geq 0.85$	27.6 %	4.2×10^{-9}
$B \to \mu^+\mu^-$	with Δz	$\mu ID \geq 0.9$	$LR \geq 0.95$	25.3 %	5.3×10^{-10}
	without Δz	$\mu ID \geq 0.9$	$LR \geq 0.85$	34.0 %	4.2×10^{-9}
$B \to e^\pm\mu^\mp$	with Δz	$\ell ID \geq 0.9$	$LR \geq 0.95$	23.3 %	5.8×10^{-10}
	without Δz	$\ell ID \geq 0.9$	$LR \geq 0.85$	30.9 %	5.2×10^{-9}

6.5 Conclusions

This chapter details the optimization and cuts on data. A cut on the number of charged tracks was increased from three to five to reduce the tau production background. Using a figure of merit, a cut on the pion identification number was optimized to be $\mathcal{L}_{\pi/K} \geq 0.5$ which greatly reduced the effects of the $B\bar{B}$ backgrounds involving kaons. Finally, a sensitivity study was performed to determine the upper limits on the branching fraction, listed in Table 6.3, and optimize the cut values on lepton identification and continuum suppression likelihood ratio. The lepton ID cut was placed at $\ell ID \geq 0.9$. The continuum suppression likelihood ratio cut was placed at $LR \geq 0.95$ for events with a valid Δz value and at $LR \geq 0.85$ for events without a valid Δz value. A full summary of the cuts is shown in Table 6.4.

Table 6.5 summarizes the final efficiencies for this study. Since the efficiency of each category of Monte Carlo data varies based on the signal mode with which it is being analyzed, each

Table 6.4: Summary of Data Cuts

Variable	Cut Value
Number of Charged Tracks	$N_{track} \geq 5$
Pion ID	$\mathcal{L}_{\pi/K} \geq 0.5$
Lepton ID	$\ell ID \geq 0.9$ where $\ell = e$ and μ
Continuum Supression Likelihood Ratio	$LR \geq 0.95$ events with Δz $LR \geq 0.85$ events without Δz

column represents a different analysis mode. Each row represents the MC data category being tested. After the application of all cuts, the signal efficiency varies from 21.6% to 25.6%. The efficiency of the background modes varies but ranges from 10^{-6} to 10^{-8} for non-$B\bar{B}$ modes and is on the order of 10^{-3} for $B\bar{B}$ background modes. Table 6.6 shows the expected number of background events for each of the analysis modes. The charm background shows the largest contribution with 31 to 72 events expected across the fitting region. The uds background contributes 5 to 17 events. Tau-pair showed the largest contribution to $B \to e^{\pm}\mu^{\mp}$ mode where 64 events are expected. $B \to \pi\pi$ showed less that one event across the analysis modes. $B \to K\pi$ has an efficiency of $< 7.40 \times 10^{-4}$ % for all analysis modes which corresponds to less than 0.11 events. Note that these background events are distributed across the fitting region and that only a small fraction are expected to populate in the much smaller "signal" region.

Table 6.5: Final Efficiencies

Mode	$B \to e^+e^-$	$B \to \mu^+\mu^-$	$B \to e^\pm\mu^\mp$
Signal	21.6 %	25.6 %	23.9 %
Charm	3.44 $\times 10^{-6}$ %	7.00 $\times 10^{-6}$ %	8.00 $\times 10^{-6}$ %
UDS	1.18 $\times 10^{-6}$ %	9.74 $\times 10^{-7}$ %	3.48 $\times 10^{-7}$ %
$ee \to \tau\tau$	1.79 $\times 10^{-6}$ %	5.87 $\times 10^{-6}$ %	1.43 $\times 10^{-5}$ %
$B \to \pi\pi$	$< 1.50 \times 10^{-3}$ %	1.33 $\times 10^{-2}$ %	4.44 $\times 10^{-3}$ %
$B \to K\pi$	$< 7.40 \times 10^{-4}$ %	$< 7.40 \times 10^{-4}$ %	$< 7.40 \times 10^{-4}$ %

Table 6.6: Expected Number of Background Events

Mode	$B \to e^+e^-$	$B \to \mu^+\mu^-$	$B \to e^\pm\mu^\mp$
Charm	31	63	72
UDS	17	14	5
$ee \to \tau\tau$	11	36	64
$B \to \pi\pi$	< 0.13	0.79	0.26
$B \to K\pi$	< 0.11	< 0.11	< 0.11

Chapter 7

Linearity Study

To validate the methods and fitting procedure used in this analysis, a linearity study was conducted. The goal of the linearity study is to demonstrate that the fitting method can accurately detect the number of signal events found with no bias or to determine the expected bias. Many attempts were made to discover the most effective method of fitting in ROOT. Selected methods, procedures, and results are described below.

7.1 Free Fit Method

The linearity study was set up similarly to the procedure used in the sensitivity study. Using the PDFs from the fitting region (5.2 GeV/c^2 $\leq$ M_{bc} $\leq$ 5.3 GeV/c^2 and -0.5 GeV $\leq$ ΔE $\leq$ 0.5 GeV), a toy Monte Carlo data set was generated representing the number

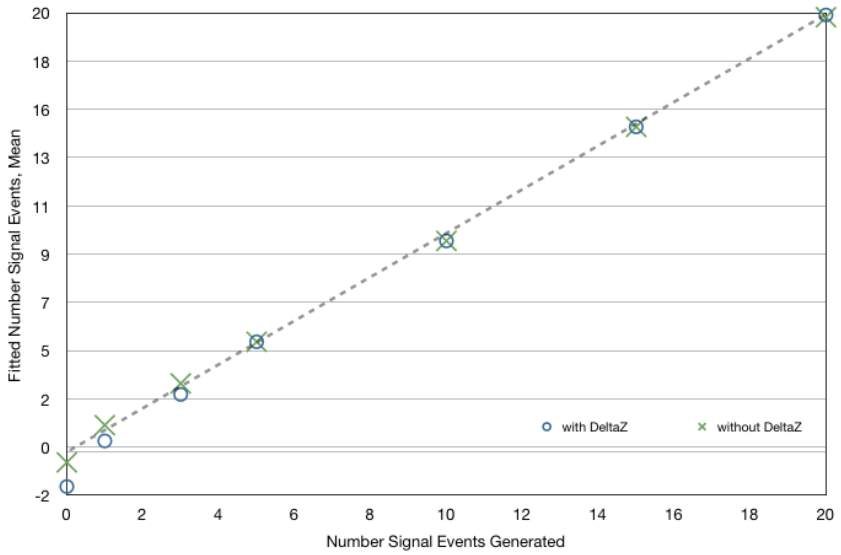


Figure 7.1: Linearity study using fitting method: results for $B \to e^+e^-$ mode

of background events expected to pass the cuts optimized in the previous chapter. The number of signal events included in the toy MC data set was varied from zero to twenty. The generated data set was then fit to PDFs of both the signal and background. On early attempts, the number of signal events found was freely fit in the range $-50 \leq$ signal yield ≤ 50. To more accurately represent the possible outcomes of real data, the number of background events generated was treated as a Poisson distribution. Unlike the sensitivity study, no multiplier in the number of events was used. For each value of generated signal events, ten-thousand toy Monte Carlo data sets were created and fit. The extracted signal yield was then compared to the number of signal events generated. Ideally, the number of signal events found by the fit equals the number of signal events generated.

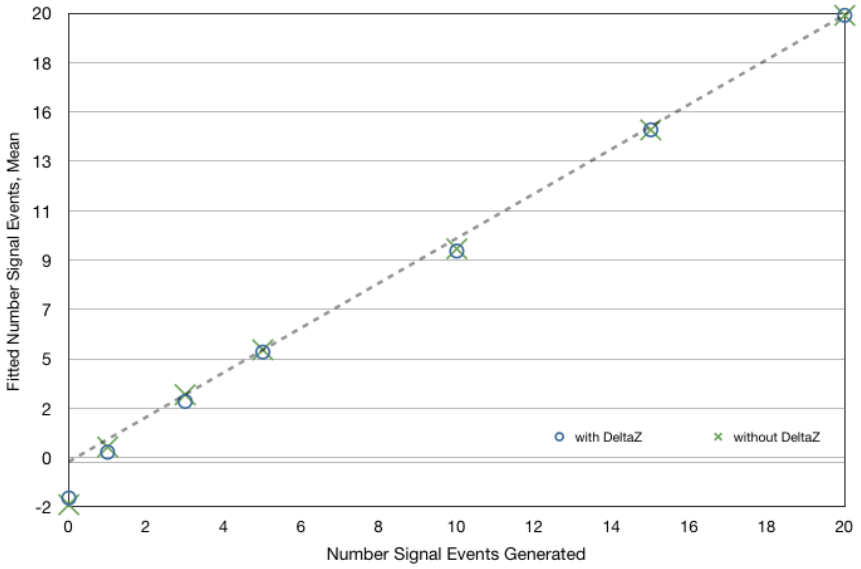


Figure 7.2: Linearity study using fitting method: results for $B \to \mu^+\mu^-$ mode

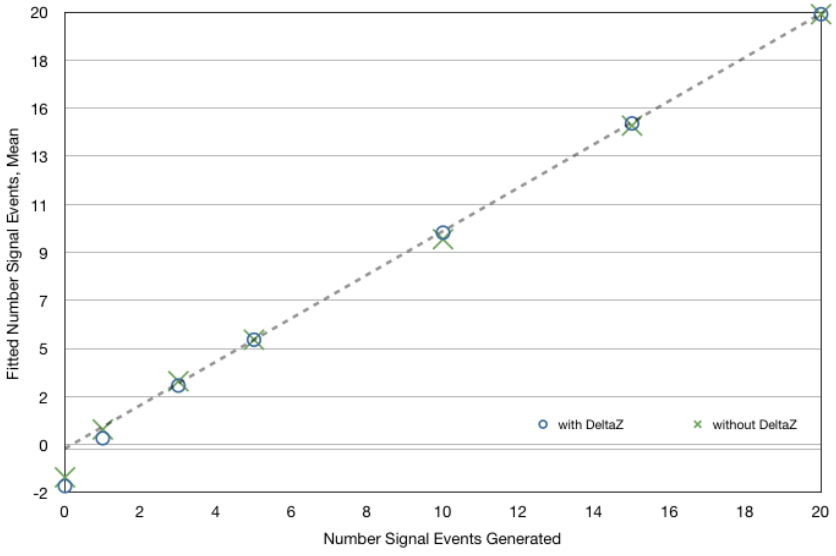


Figure 7.3: Linearity study using fitting method: results for $B \to e^{\pm}\mu^{\mp}$ mode

For tests with five or more signal events generated, the linearity study showed excellent results. For low numbers of signal events generated, most notably zero or one signal event generated, the linearity study showed a negative bias as seen in Figures 7.1, 7.2, and 7.3 for the three signal modes. Upon further investigation, it was found that, while the fits peak at the correct value of signal yield, these sets of toy MC data produced a large tail on the negative side as seen in Figure 7.4. Figure 7.4 shows the resulting signal yield for $B \to \mu^+\mu^-$ when zero signal events were generated. This disproportionately large number of negative results appears to stem from the ROOT fitting method getting stuck in a local minimum. One such fit is included as an example in Figure 7.5. It shows the fit for one toy MC on $B \to \mu^+\mu^-$ where the best-fit result was a signal yield of -18 signal events. As illustrated, the signal found dips well below the shape of the generated toy MC data where there happens to be a zero-empty bin.

Several different methods were explored in attempts to correct this bias. One method used a fixed value of signal yield and fit the number of background events to the data set. The number of signal events stepped through a range of values and the negative log likelihood (NLL) was examined for each fit. The fit that corresponded to the best NLL was considered the best fit and saved. The procedure was then repeated for many trials and with different numbers of signal events generated.

A second method tested set multiple smaller ranges of signal yield. The ranges then stepped

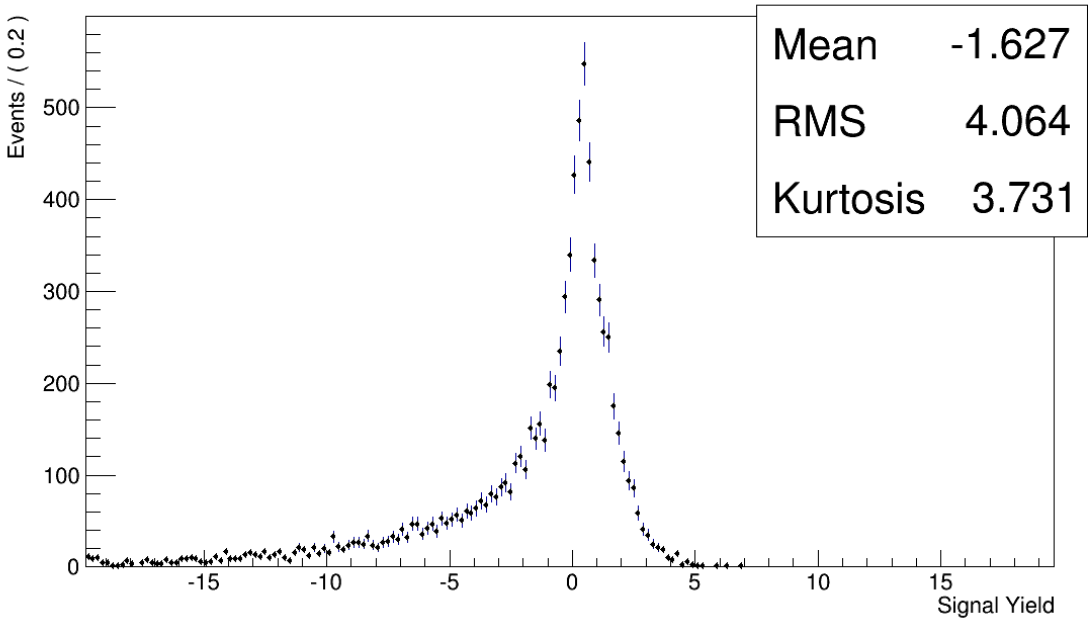
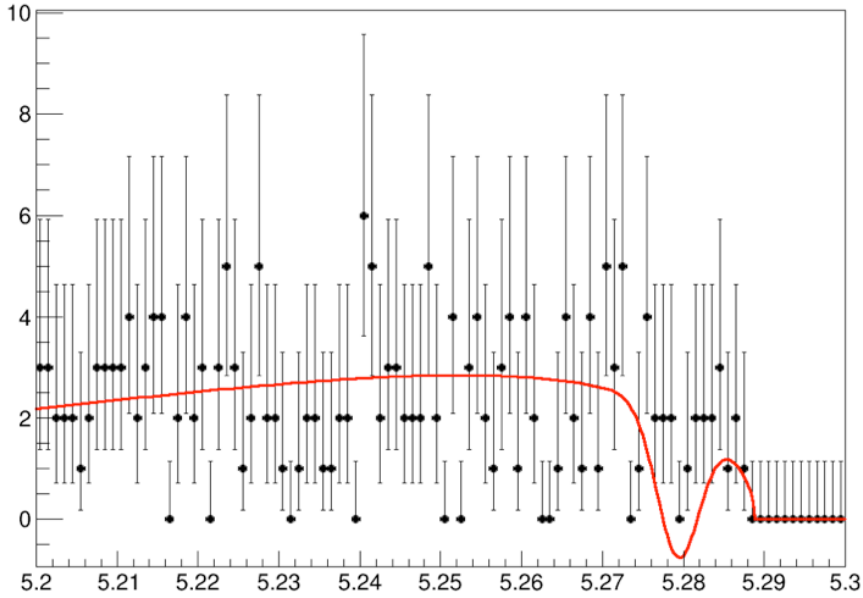


Figure 7.4: Signal yield for $B \to \mu^+\mu^-$ with null signal generated

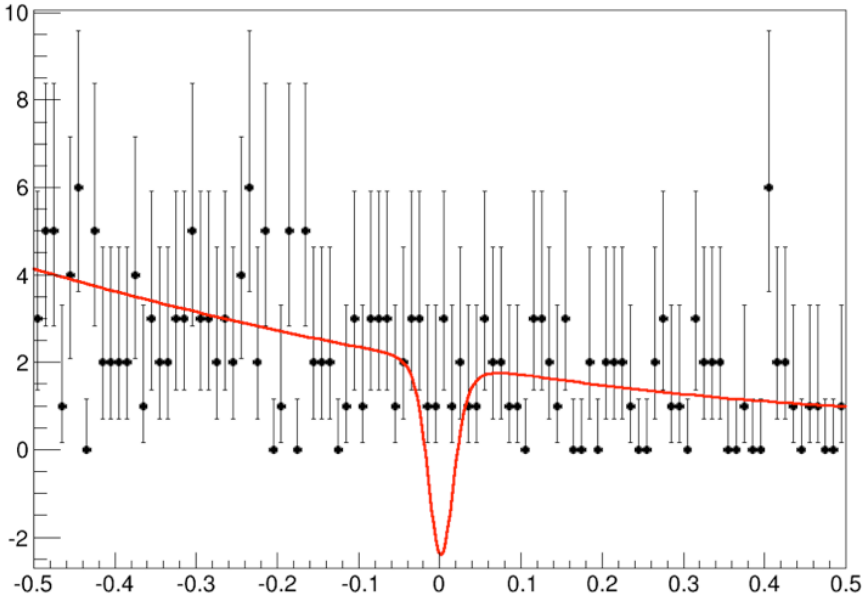
through a wide set of possible values. Both the signal and background were fit. The fit which produced the best value of NLL was saved. Both of these methods proved inferior to a linearity study in which the initial guess value was varied.

7.2 Initial Guess Method

When defining a variable in RooFit [97], the specified parameters include the minimum value, the maximum value, and the initial guess [98]. A linearity study was conducted where the toy MC data set was fit several times with different initial guesses. After the signal and



(a) M_{bc}



(b) ΔE

Figure 7.5: A sample fit for $B \to \mu^+\mu-$ with null signal generated

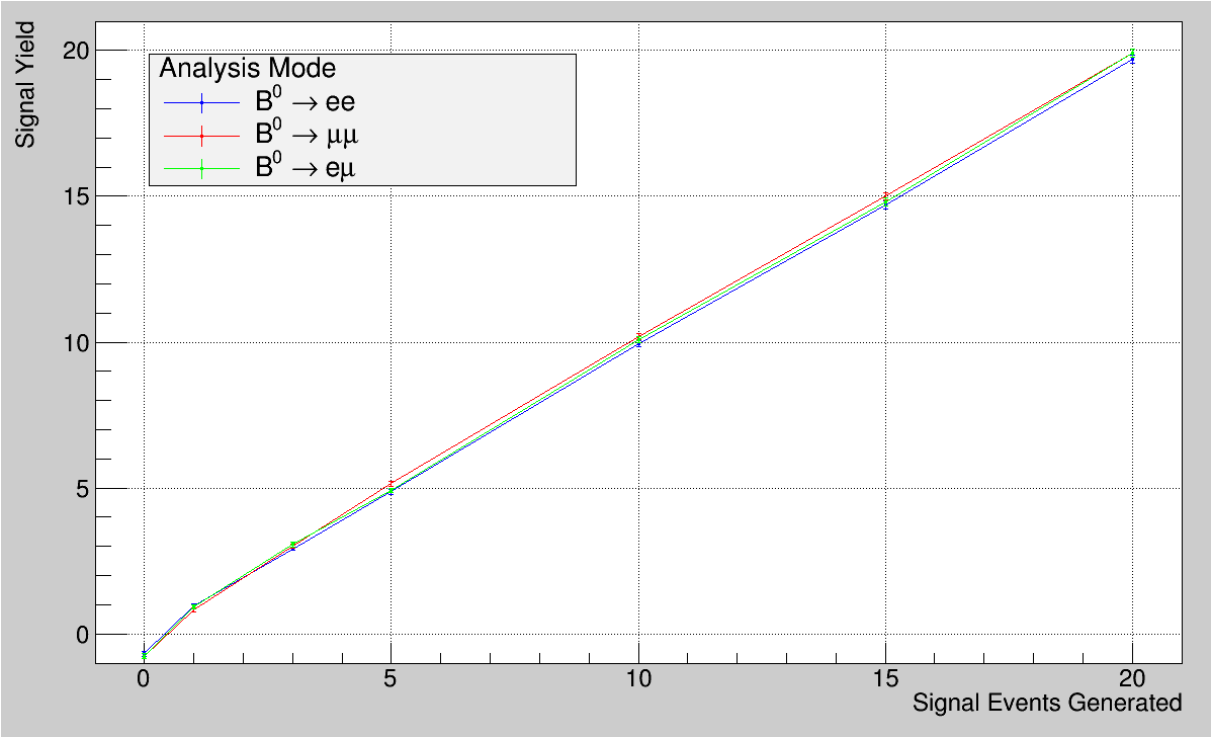


Figure 7.6: Results for linearity study using initial guess method

background were fit to the data set, the fit which produced the best NLL was saved. The procedure was repeated for up to twenty signal events generated.

Figure 7.6 shows the results of the linearity study. Here, each line represents a different analysis mode. The results show good agreement for all trials with one or more signal events included in the data set. For trials with null signal, the resulting signal yield still falls below the expected outcome but does show marked improvement when compared to other fitting methods. The signal yield and errors are included in Table 7.1.

Table 7.1: Results for linearity study using initial guess method.

Signal events generated	Signal Yield $B \to e^+e^-$	Signal Yield $B \to \mu^+\mu^-$	Signal Yield $B \to e^\pm\mu^\mp$
0	-0.68 ± 0.08	-0.76 ± 0.08	-0.76 ± 0.08
1	0.96 ± 0.07	0.82 ± 0.07	0.93 ± 0.06
3	2.92 ± 0.07	3.02 ± 0.07	3.05 ± 0.07
5	4.87 ± 0.08	5.15 ± 0.08	4.91 ± 0.08
10	10.0 ± 0.10	10.2 ± 0.11	10.1 ± 0.11
15	14.7 ± 0.13	15.0 ± 0.12	14.8 ± 0.13
20	19.7 ± 0.15	19.9 ± 0.15	19.9 ± 0.15

7.3 Conclusions

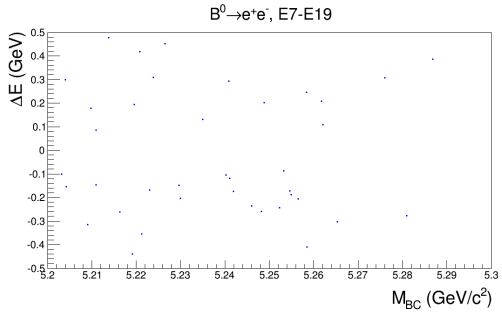
Early methods of fitting produced a bias in the linearity study. Several attempts were made to correct this. The most successful fitting procedure uses multiple initial guesses and chooses the fit with the best negative log likelihood. Though this method shows significant improvement, it does still produce a bias for null signal. For trials with signal events present, the chosen method yields no bias.

Chapter 8

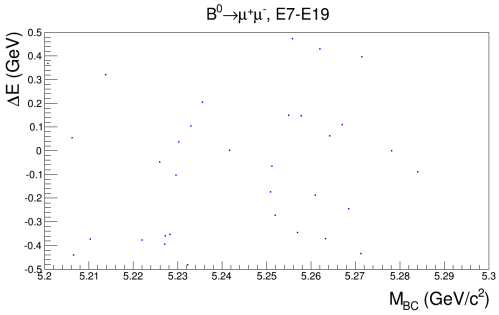
Results

To prove that the analysis methods are effective, they were compared to the results of Ming-Chuan Chang who did the first Belle analysis of these decay modes in 2003. Chang reported upper limits on the branching fractions of $\mathcal{B}(B^0 \to e^+e^-) < 1.9 \times 10^{-7}$, $\mathcal{B}(B^0 \to \mu^+\mu^-) < 1.6 \times 10^{-7}$, and $\mathcal{B}(B^0 \to e^\pm\mu^\mp) < 1.7 \times 10^{-7}$ [99, 28]. Her study corresponded to a $78fb^{-1}$ data set or 85 million $B\bar{B}$ pairs. For a comparison and validation of the present analysis, the procedures established in this study were applied to the same $78fb^{-1}$ data set.

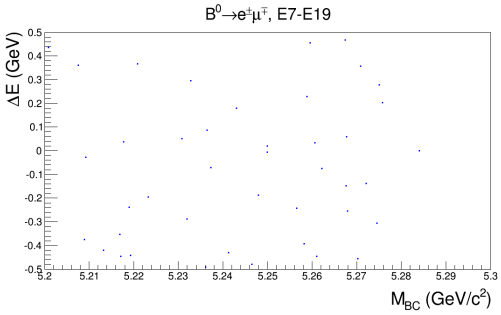
Of the 85 million $B\bar{B}$ pairs, 39 events passed the cuts for $B^0 \to e^+e^-$, 33 events passed for $B^0 \to \mu^+\mu^-$, and 42 events passed for $B^0 \to e^\pm\mu^\mp$. The results are shown in Figure 8.1. After the implementation of the optimized cuts, the fit procedure was applied using the previously established PDFs for M_{BC} and ΔE. At the 90% confidence level, the signal yield was 1.5



(a) $B^0 \to e^+e^-$



(b) $B^0 \to \mu^+\mu^-$



(c) $B^0 \to e^\pm\mu^\mp$

Figure 8.1: $B^0 \to \ell^+\ell^-$ events passing cuts for 85 million $B\bar{B}$ pairs data set.

for $B^0 \to e^+e^-$, 2.1 for $B^0 \to \mu^+\mu^-$, and 1.6 for $B^0 \to e^\pm\mu^\mp$. For comparison, a simple count of the number of events within the much smaller "signal box" ($5.27 < M_{bc} < 5.29$ and $|\Delta E| < 0.05$) contains 0, 1, and 1 events, respectively.

Table 8.1: Signal Yield for $B^0 \to \ell^+\ell^-$ using a 85 million $B\bar{B}$ pairs data set.

Mode	No. Events Passing	Signal Efficiency ϵ	Signal Yield
$B^0 \to e^+e^-$	39	21.6 %	1.5
$B^0 \to \mu^+\mu^-$	33	25.6 %	2.1
$B^0 \to e^\pm\mu^\mp$	42	23.9 %	1.6

.

8.1 Estimation of Systematic Uncertainties

To carefully evaluate the errors in this study, established methods of calculation were chosen. Some calculations are standard procedures performed at the Belle experiment. Other uncertainties were independently calculated but established using proven methods. This chapter details all systematic uncertainties on the measured branching fractions.

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

The number of $B\bar{B}$ pairs is measured directly by the Belle detector. It is determined by comparing the difference in cross section when the recording data at 80 MeV below the $\Upsilon(4S)$ resonance and data at the $\Upsilon(4S)$ resonance.

$$N(B\bar{B}) = N(on) - acN(off) \tag{8.1}$$

The ratio c, of efficiency for recording background $q\bar{q}$ events in an on and off-resonance sample is determined from Monte Carlo. The scale factor a, for on and off-resonance $q\bar{q}$ background is calculated from Bhabha pairs and muon pairs from data. The systematic uncertainties have been globally determined for all Belle analyses; here, the systematic error is 1.3% [100].

8.1.2 Track Reconstruction

The tracking systematics are uncertainties that affect the efficiency of reconstruction. Tracking efficiency was measured by studying the decay chain $D^* \to \pi D^0$, $D^0 \to \pi\pi K_s^0$, and $K_s^0 \to \pi^+\pi^-$. By calculating the ratio of tracking efficiency in data and MC, the systematic uncertainty associated with tracking has been evaluated. For particles with momentum higher than 200 MeV, the uncertainty is 0.35% per track and is additive. Since $B^0 \to \ell^+\ell^-$ decays fall into this momentum region and the decays involve two tracks, the total uncertainty for tracking systematics is 0.70% [101].

8.1.3 Particle Identification

In order to create a consistent and manageable approach to the calculation of PID uncertainties within the Belle experiment, the PID Joint Group developed a table of uncertainties. For each charged track, the tracking efficiency and associated uncertainty are evaluated as a function of the PID cut value and the track momentum and polar angle. Kaons and pions were evaluated for their misidentification rate. The efficiency and error were assessed for electrons and muons. For all three decay modes, kaon identification contributed 0.73% to the systematic uncertainty. The ℓID contributes 4.4% to $B^0 \to e^+e^-$, 4.4% to $B^0 \to \mu^+\mu^-$, and 4.6% to $B^0 \to e^\pm\mu^\mp$.

8.1.4 Likelihood Ratio

The systematic uncertainty stemming from the valued used for the likelihood ratio cut was estimated by comparing the Monte Carlo data with a control sample from previous studies [99]. The efficiency with and without applying the likelihood ratio cuts were calculated for both the MC and the data. The total efficiency from this ratio of ratios is

$$\epsilon_{total} = \frac{data}{MC} = \frac{0.30 \pm 0.011}{0.311 \pm 0.052} = 0.96 \pm 0.06. \tag{8.2}$$

This systematic uncertainty is statistically limited.

8.1.5 Fit Model

During the extraction of the signal yield, the final fit used PDFs for M_{BC} and ΔE. For each fit parameter contained within the PDFs, there exists an uncertainty in its value. To evaluate the contribution of each such uncertainty to the final signal yield's uncertainty, each parameter was both increased and decreased by one standard deviation from its nominal value. Then, the signal extraction was performed again and the deviation from the nominal fit's signal yield was determined. The fit proved to be very stable showing little variation due this procedure. The uncertainty associated with the fit model is 0.2%.

Table 8.2: Contributions to the total systematic uncertainty. All contributions are given in units of percent.

	$B^0 \to e^+e^-$	$B^0 \to \mu^+\mu^-$	$B^0 \to e^\pm\mu^\mp$
$N_{B\bar{B}}$	1.3	1.3	1.3
Track Reconstruction	0.70	0.70	0.70
ℓID	4.4	4.4	4.6
$K\pi ID$	0.73	0.73	0.73
Fitting Parameters	0.2	0.2	0.2
LR	5.8	5.8	5.8
Total	7.4	7.4	7.6

.

8.1.6 Total Systematic Uncertainties

Table 8.2 summarizes the statistical uncertainties contained within this analysis. The total systematic uncertainties amount to 7.4% for $B^0 \to e^+e^-$ and $B^0 \to \mu^+\mu^-$ modes and 7.6% for $B^0 \to e^\pm\mu^\mp$.

8.2 Final Calculation of Branching Fraction

The branching fractions were calculated using the following equation:

$$BF = \frac{Signal\ Yield}{\epsilon_{sig} N_B (1 - \sigma_{sys})} \tag{8.3}$$

Here, the signal yield is taken at the 90% confidence level, ϵ_{sig} is the signal efficiency, N_B represents the number of B mesons, and σ_{sys} is the uncertainty arising from systematic errors. The resulting branching fractions are presented in Table 8.3.

Table 8.3: Comparison of branching fraction results for $B^0 \to \ell^+\ell^-$ using a 85 million $B\bar{B}$ pairs data set. The previous analysis was performed in 2003 by M.-C. Chang [28].

Mode	ϵ_{sig}	BF (10^{-7})	ϵ_{sig}	BF (10^{-7})
	Previous Analysis		This Analysis	
$B^0 \to e^+e^-$	14.3 %	< 1.9	21.6 %	< 0.88
$B^0 \to \mu^+\mu^-$	16.9 %	< 1.6	25.6 %	< 1.0
$B^0 \to e^\pm\mu^\mp$	15.8 %	< 1.7	23.9 %	< 0.85

.

For the 85 million $B\bar{B}$ data set, the resulting upper limits on branching fractions were 8.8×10^{-8} for $B^0 \to e^+e^-$, 1.0×10^{-7} for $B^0 \to \mu^+\mu^-$, and 8.5×10^{-8} for $B^0 \to e^\pm\mu^\mp$ which are slight improvements over the previous study.

8.3 Conclusions

The exceeding rarity of the $B^0 \to \ell^+\ell^-$ decays provide unique challenges and opportunities to researchers. The minute branching fraction requires that the data be handled meticulously and that all plausible sources of background be evaluated and suppressed. This study used Monte Carlo data to optimize event selection while maintaining a higher signal efficiency than previous studies. When applied to the same data set as the previous studies, the analysis procedures developed here show not only consistency, but also some improvement in the final measure of the limit on branching fraction. The methods used are now ready to be applied to the entire Belle data set.

Appendix A

EvtGen Code

The decay.dec input files used to generate the $B^0 \to \ell^+\ell^-$ Monte Carlo data in EvtGen are very simple. It starts with the $\Upsilon(4S)$ resonance which decays to a $B^0\bar{B}^0$ pair. One B meson is forced to decay into a pair of leptons. The other B meson is allowed to decay generically. These input files (plain text) are included below in Figure A.1 through Figure A.4.

```
### B0 -> e+ e- and B0bar -> generic or charged conjugate

Alias MyB0 B0

Decay Upsilon(4S)
0.5  MyB0 anti-B0     VSS;
0.5  MyB0 B0            VSS;
Enddecay

Decay MyB0
1.0  e+    e-    PHOTOS SLL;
Enddecay

End
```

Figure A.1: Decay.dec for mode $B \to ee$

```
### B0 -> mu+ mu- and B0bar -> generic or charged conjugate

Alias MyB0 B0

Decay Upsilon(4S)
0.5  MyB0 anti-B0     VSS;
0.5  MyB0 B0            VSS;
Enddecay

Decay MyB0
1.0  mu+   mu-   PHOTOS SLL;
Enddecay

End
```

Figure A.2: Decay.dec for mode $B \to \mu\mu$

```
### B0 -> e+ mu- and B0bar -> generic or charged conjugate

Alias MyB0 B0

Decay Upsilon(4S)
0.5  MyB0 anti-B0     VSS;
0.5  MyB0 B0            VSS;
Enddecay

Decay MyB0
1.0  e+    mu-    PHOTOS SLL;
Enddecay

End
```

Figure A.3: Decay.dec for mode $B \to e\mu$

```
### B0 -> mu+ e- and B0bar -> generic or charged conjugate

Alias MyB0 B0

Decay Upsilon(4S)
0.5  MyB0 anti-B0     VSS;
0.5  MyB0 B0            VSS;
Enddecay

Decay MyB0
1.0  mu+   e-    PHOTOS SLL;
Enddecay

End
```

Figure A.4: Decay.dec for mode $B \to \mu e$

Appendix B

HadronB Event Selection

The HadronB skim was designed to select 100% of $B\bar{B}$ events and the maximum number of $e^+e^- \to q\bar{q}$ continuum events possible. Ideally, all non-hadronic events should be rejected. To achieve this result, cuts are applied to track properties as well as event variables taken from the mdst tables.

The hadronic event selection variables are calculated using reconstructed tracks and ECL clusters. The reconstructed tracks used have $p_T > 100$ MeV/c, and impact parameters from the nominal IP of $dr < 2.0$ cm and $|dz| < 4.0$ cm. ECL clusters with energy $E > 100$ MeV and those not matched to CDC tracks define photon candidates within the CDC acceptance ($17^\circ \leq \theta \leq 150^\circ$). The tracks and photon candidates are boosted into the $\Upsilon(4S)$ rest frame with pion and zero mass assumptions respectively. In the following, the square of the total

center of mass energy is denoted by s, where $\sqrt{s} = 10.58$ GeV.

The HadronB criteria are:

- The number of charged tracks,

$$N_{TRACK} \geq 3; \tag{B.1}$$

- The total visible energy defined as the sum of track and photon energies

$$E_{VISIBLE} \geq 0.2\sqrt{s}; \tag{B.2}$$

- The sum of track and photon momentum projected onto the z-axis

$$\left|\sum p_z c\right| \leq 0.5\sqrt{s}; \tag{B.3}$$

- The primary event vertex, determined from all tracks, must be positioned within a cylindrical surface defined by

$$r < 1.5 \text{ cm and } |z| < 3.5 \text{ cm} \tag{B.4}$$

and centered at the IP.

- The sum of the cluster energies in the ECL within the CDC acceptance

$$0.1\sqrt{s} \leq E_{sum}^{CDC} \leq 0.8\sqrt{s}; \tag{B.5}$$

- The number of ECL clusters in the barrel region of the detector ($-0.7 \leq \cos\theta \leq 0.9$)

$$N_{CLUSTER} > 1; \tag{B.6}$$

- The energy sum of clusters within the ECL, and the Heavy Jet Mass, HJM

$$E_{sum}^{ECL} > 0.18\sqrt{s} \text{ or } HJM > 1.8 \text{ GeV}/c^2; \tag{B.7}$$

  To calculate the HJM, an event is split into two hemispheres with the dividing plane placed perpendicular to the event thrust axis. The invariant mass is calculated for each hemisphere by combining the track and photon four-momenta. The larger of the two mass values becomes the Heavy Jet Mass.

- If $HJM \leq 1.8$ GeV/c^2, then

$$HJM > 0.25E_{VISIBLE}; \tag{B.8}$$

Table B.1: Cross-sections and efficiencies for various processes subject to the HadronB selection criteria.

Process $e^+e^- \rightarrow$	$B\bar{B}$	$q\bar{q}$	$\tau^+\tau^-$	$e^+e^-(\gamma)$	$\gamma\gamma$
$\epsilon(\%)$	99.1	79.5	4.9	0.002	0.4
σ(nb)	1.09	2.62	0.05	0.001	0.04

- The average cluster energy calculated from clusters within the ECL

$$\frac{E_{sum}^{ECL}}{N_{CLUSTER}^{ECL}} < 1 \text{ GeV}; \tag{B.9}$$

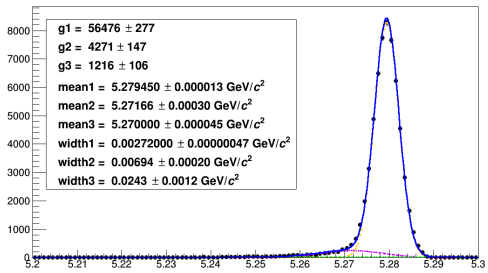
The combination of the cuts produces a skim set that retains nearly all $B\bar{B}$ events and a large portion of $e^+e^- \rightarrow q\bar{q}$ continuum events. Meanwhile most occurrences of $\tau^+\tau^-$ production events, QED events, and two-photon events are rejected. The resulting efficiencies and cross-sections for individual processes after the application of HadronB criteria are listed in Table B.1 [100].

Appendix C

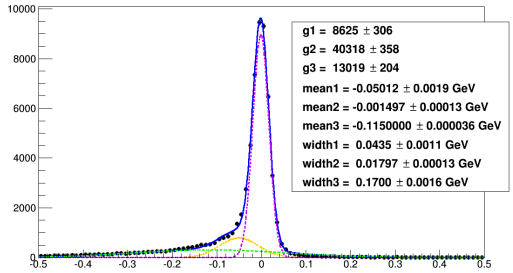
Fits and Parameters

This appendix contains the fit parameters for signal and background modes involved in $B \to \ell^+\ell^-$ analysis. Figures C.1, C.2, and C.3 are the signal modes: $B \to e^+e^-$, $B \to \mu^+\mu^-$, and $B \to e^\pm\mu^\mp$. For all three signal modes, M_{bc}, ΔE, and Δz were modeled using a summation of three Gaussian distributions. By using multiple Gaussians, the PDFs were able to accommodate skewness, such as in the $B \to e^+e^-$ M_{bc} PDF, and distributions with wider bases, such as in Δz fits. The $\cos\theta_B$ distribution was fit using a second-order polynomial function. Since the absolute value of $\cos\theta_B$ is used in analysis, the fit has a linear component set to zero. KSFW likelihood ratio and lepton ID distributions were treated with second-order histogram functions.

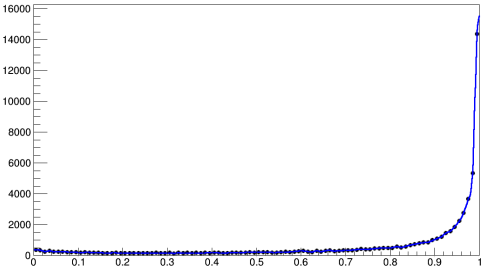
Figures C.4 through C.6 contain the fits and parameters for continuum Monte Carlo and



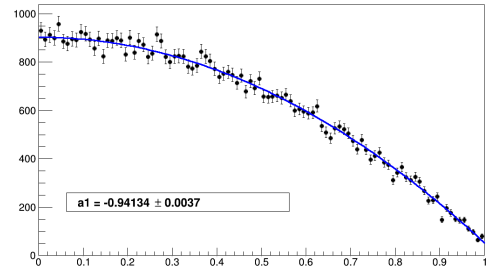
(a) M_{bc}



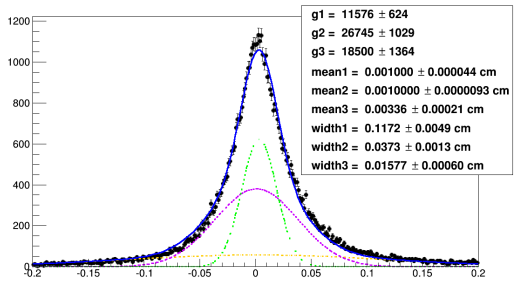
(b) ΔE



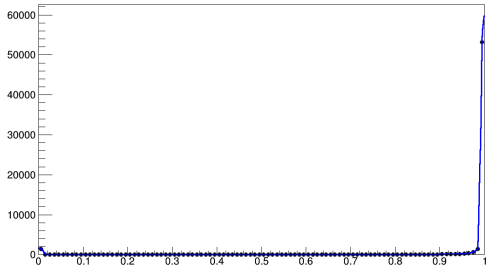
(c) KSFW Likelihood Ratio



(d) $\cos\theta_B$

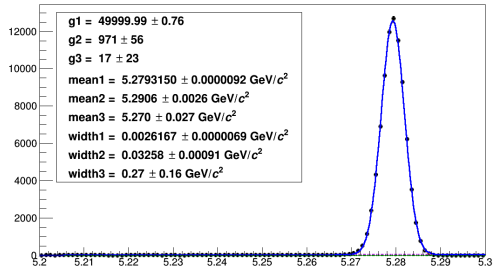


(e) Δz

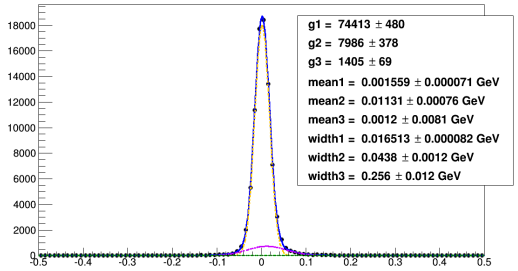


(f) Lepton PID

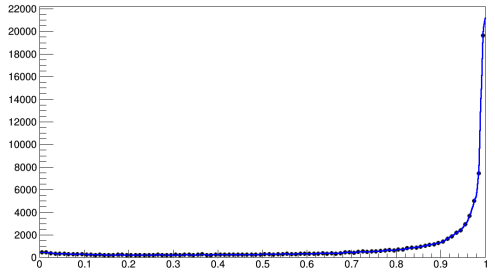
Figure C.1: Fits with parameters for $B \to e^+e^-$ MC



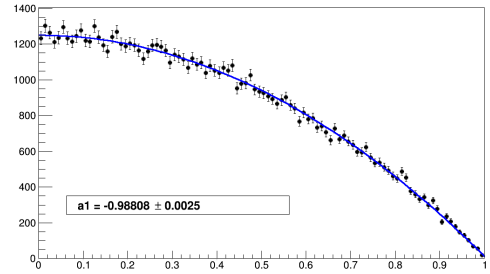
(a) M_{bc}



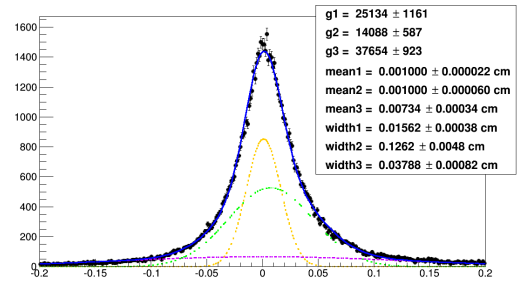
(b) ΔE



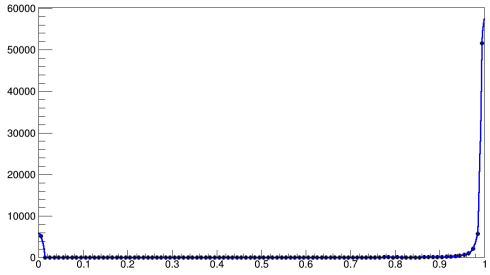
(c) KSFW Likelihood Ratio



(d) $\cos\theta_R$

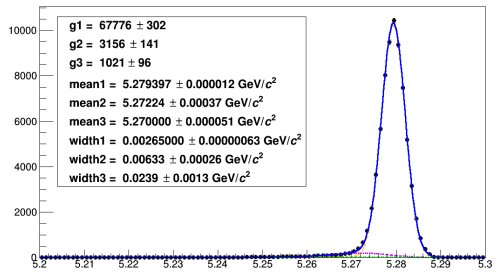


(e) Δz

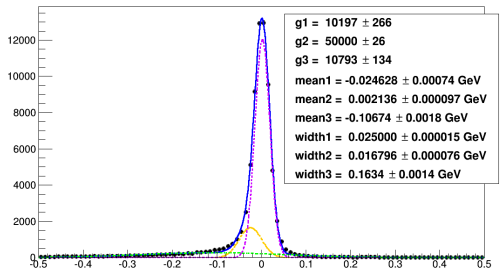


(f) Lepton PID

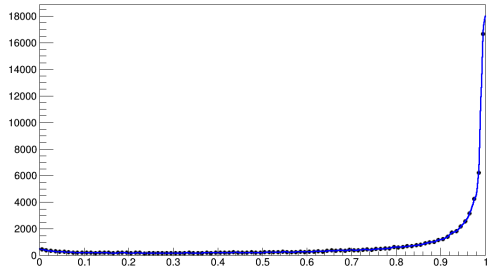
Figure C.2: Fits with parameters for $B \to \mu^+\mu^-$ MC



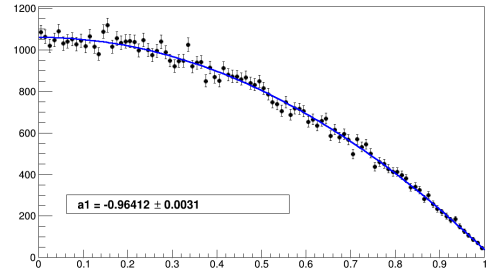
(a) M_{bc}



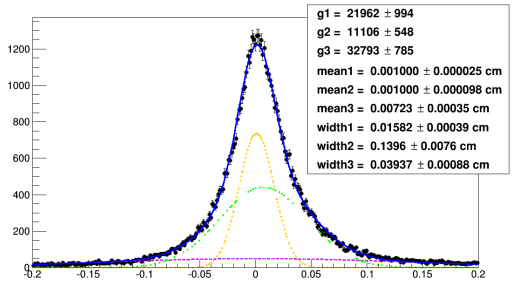
(b) ΔE



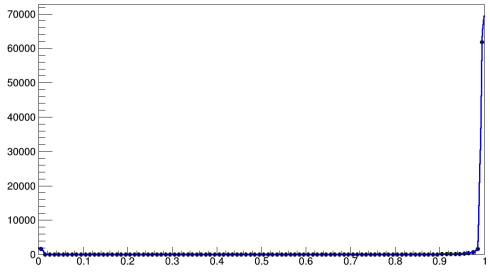
(c) KSFW Likelihood Ratio



(d) $\cos\theta_R$

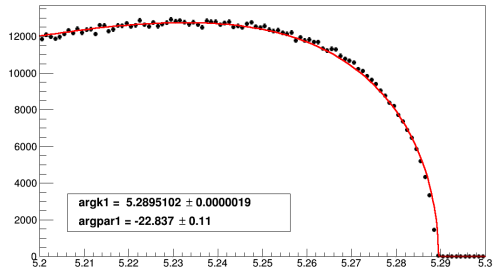


(e) Δz

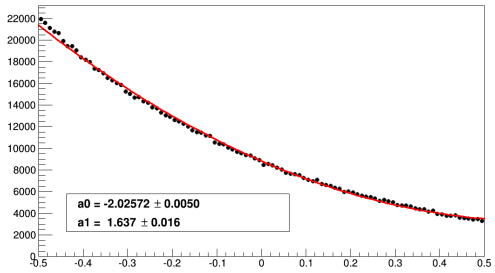


(f) Lepton PID

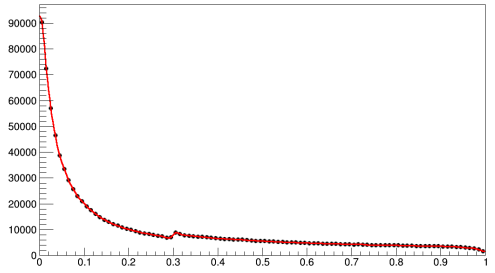
Figure C.3: Fits with parameters for $B \to e^{\pm}\mu^{\mp}$ MC



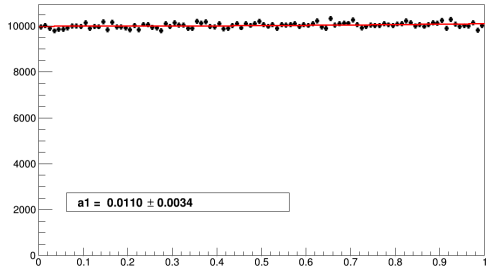
(a) M_{bc}



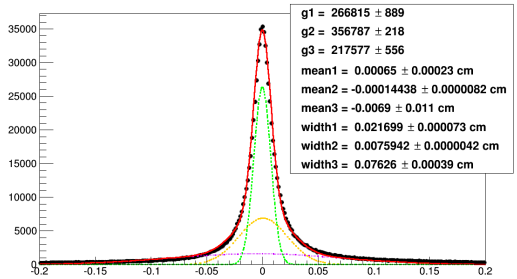
(b) ΔE



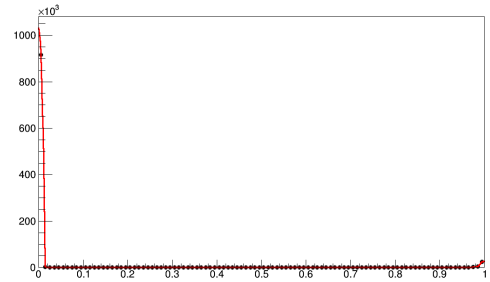
(c) KSFW Likelihood Ratio



(d) $\cos\theta_B$



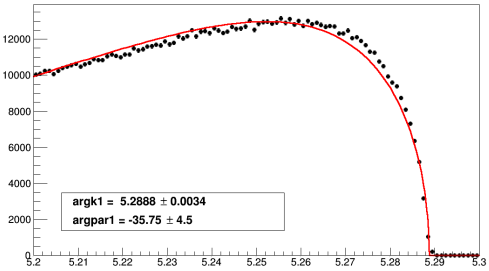
(e) Δz



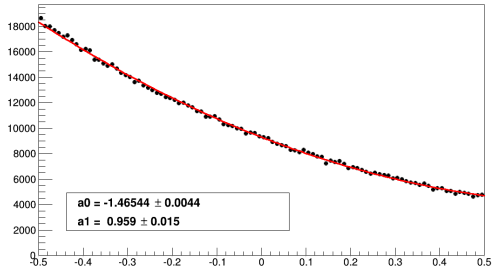
(f) Lepton PID

Figure C.4: Fits with parameters for $e^+e^- \to c\bar{c}$ MC

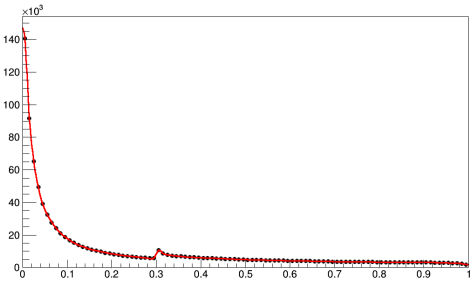
tau-pair production Monte Carlo. Like the signal modes, Δz was fit using the sum of three common-mean Gaussian distributions. KSFW likelihood ratio and lepton ID were fit with second-order histogram functions. However, M_{bc} distributions were fit using an ARGUS function. Like $\cos\theta_B$, ΔE was fit using a second-order polynomial function but the linear component was allowed to vary.



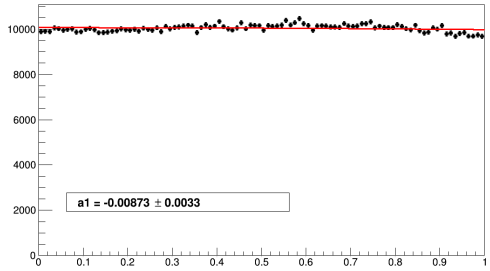
(a) M_{bc}



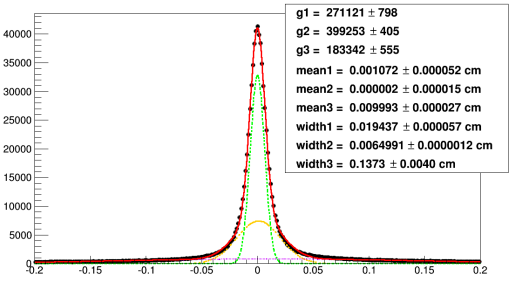
(b) ΔE



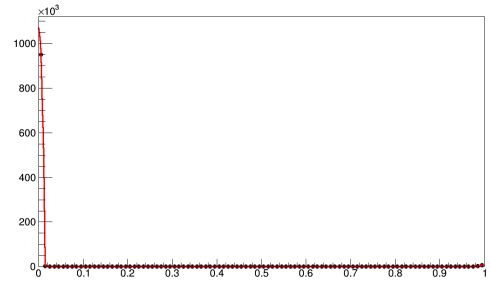
(c) KSFW Likelihood Ratio



(d) $\cos\theta_R$

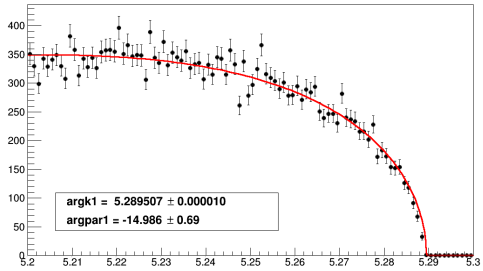


(e) Δz

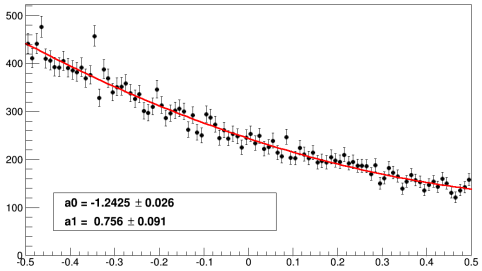


(f) Lepton PID

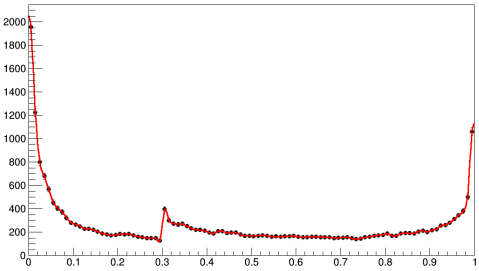
Figure C.5: Fits with parameters for uds MC



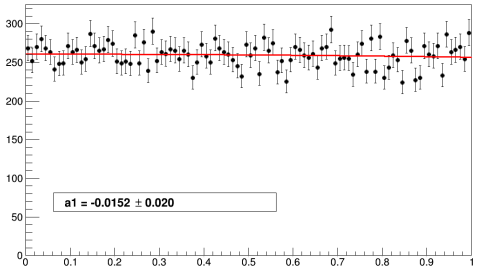
(a) M_{bc}



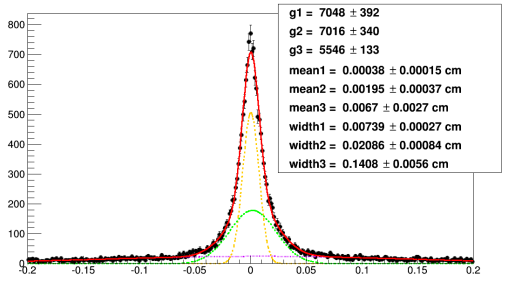
(b) ΔE



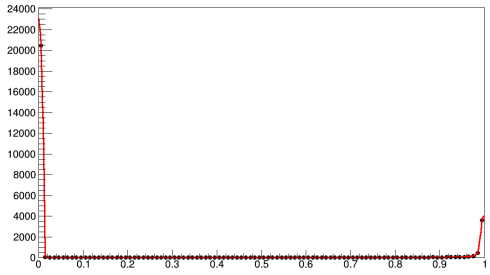
(c) KSFW Likelihood Ratio



(d) $\cos\theta_R$



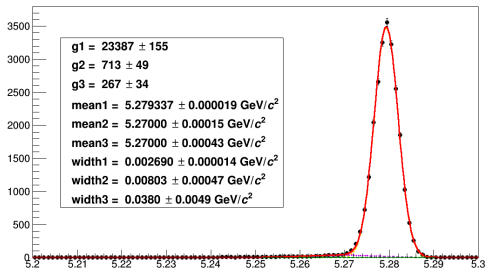
(e) Δz



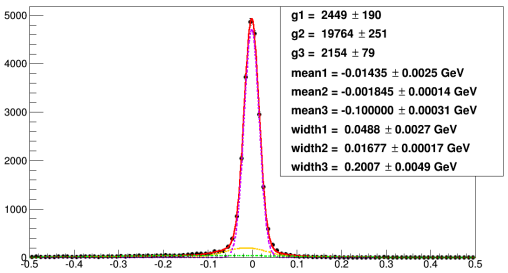
(f) Lepton PID

Figure C.6: Fits with parameters for $e^+e^- \to \tau^+\tau^-$ MC

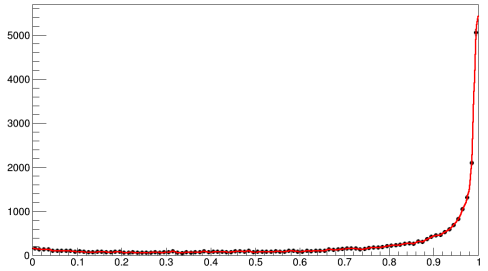
Figures C.7 through C.9 show the fits for the relevant $B\bar{B}$ background modes: $B^0 \to \pi^+\pi^-$, $B^0 \to K^-\pi^+$, and $B^0 \to K^+K^-$. Because the PDFs for these modes are similar to the signal modes, they were fit using the same functions as signal distributions: a summation of Gaussian distributions for M_{BC}, ΔE, and Δz; a polynomial function for $\cos\theta_B$; and histogram functions for KSFW likelihood ratio and lepton ID distributions. Histogram functions are non-parametric PDFs based on an input histogram where the value equals the normalized bin content [102].



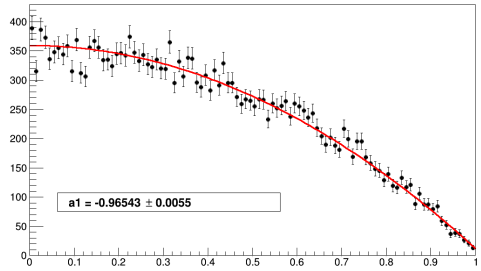
(a) M_{bc}



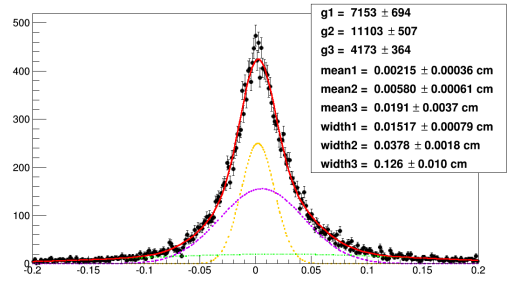
(b) ΔE



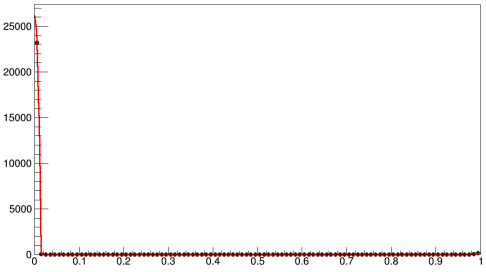
(c) KSFW Likelihood Ratio



(d) $\cos\theta_R$

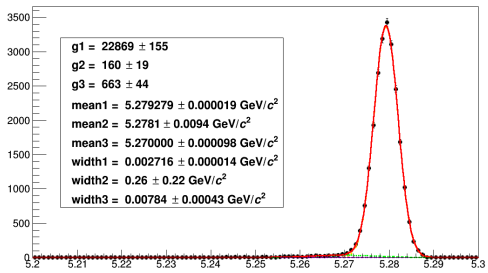


(e) Δz

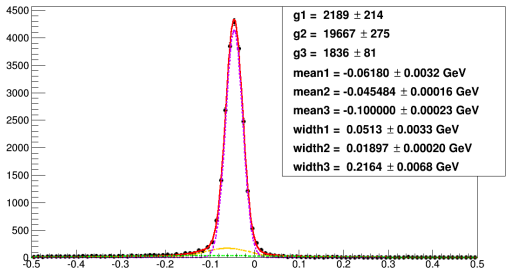


(f) Lepton PID

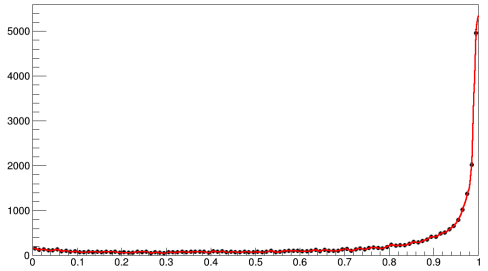
Figure C.7: Fits with parameters for $B\bar{B} \to \pi^+\pi^-$ MC



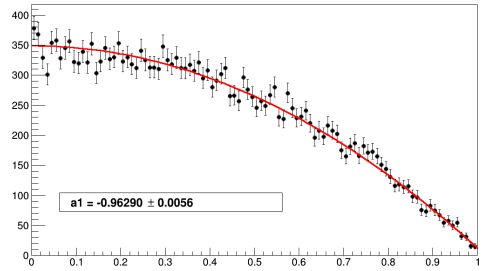
(a) M_{bc}



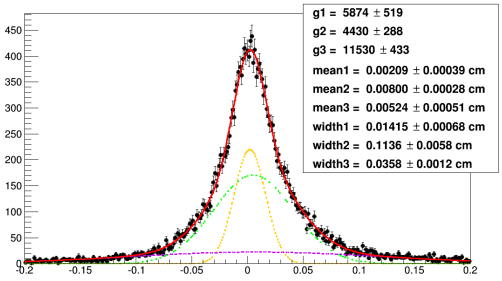
(b) ΔE



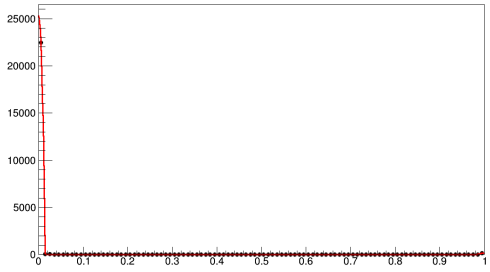
(c) KSFW Likelihood Ratio



(d) $\cos\theta_R$

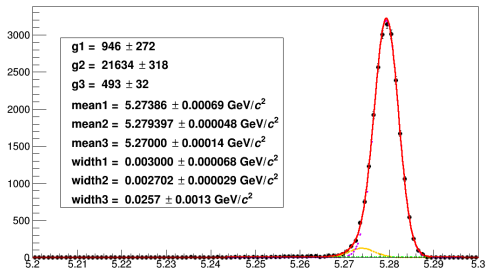


(e) Δz

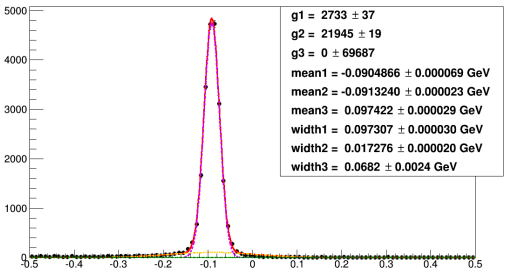


(f) Lepton PID

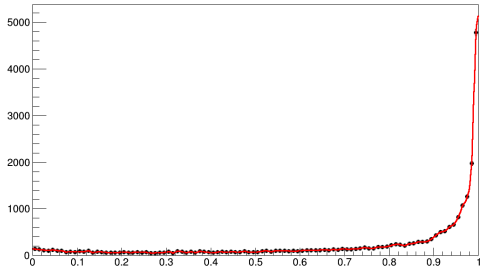
Figure C.8: Fits with parameters for $B\bar{B} \to K^-\pi^+$ MC



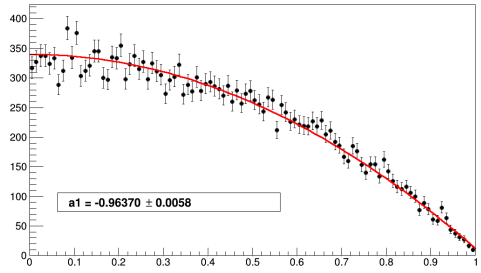
(a) M_{bc}



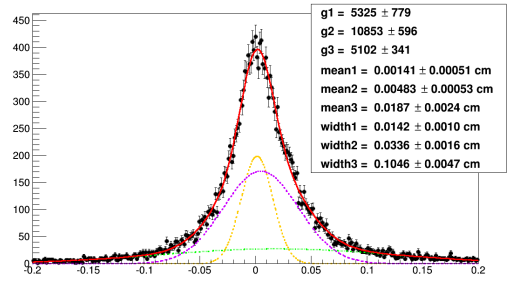
(b) ΔE



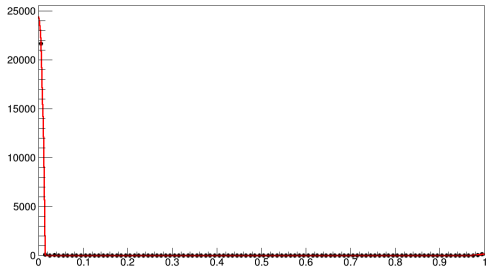
(c) KSFW Likelihood Ratio



(d) $\cos\theta_R$



(e) Δz



(f) Lepton PID

Figure C.9: Fits with parameters for $B\bar{B} \to K^+K^-$ MC

Bibliography

[1] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[2] The CMS Collaboration. Evidence for the direct decay of the 125 GeV Higgs boson to fermions. *Nat Phys*, 10(8):557–560, 08 2014.

[3] K. S. Babu and Christopher Kolda. Higgs-Mediated $B^0 \to \mu^+\mu^-$ in Minimal Supersymmetry. *Phys. Rev. Lett.*, 84:228–231, Jan 2000.

[4] Piotr H. Chankowski and Łucja Sławianowska. $B^0_{d,s} \to \mu^-\mu^+$ decay in the MSSM. *Phys. Rev. D*, 63:054012, Feb 2001.

[5] Jogesh C. Pati and Abdus Salam. Lepton number as the fourth "color". *Phys. Rev. D*, 10:275–289, Jul 1974.

[6] Antonio Pich. Effective field theory: Course. In *Probing the standard model of particle interactions. Proceedings, Summer School in Theoretical Physics, NATO Advanced Study Institute, 68th session, Les Houches, France, July 28-September 5, 1997. Pt. 1, 2*, pages 949–1049, 1998.

[7] J. Harrison. Searches for lepton flavour violation and baryon number violation at LHCb. *Nuclear Physics B - Proceedings Supplements*, 253–255:20 – 23, 2014. The Twelfth International Workshop on Tau-Lepton Physics (TAU2012).

[8] Monika Blanke, Andrzej J. Buras, Bjoern Duling, Anton Poschenrieder, and Cecilia Tarantino. Charged Lepton Flavour Violation and $(g-2)_\mu$ in the Littlest Higgs Model with T-Parity: A Clear Distinction from Supersymmetry. *JHEP*, 05:013, 2007.

[9] Rodolfo A. Diaz, R. Martinez, and Carlos E. Sandoval. Flavor changing neutral currents from lepton and B decays in the two Higgs doublet model. *Eur. Phys. J.*, C41:305–310, 2005.

[10] A. Ilakovac. Lepton flavor violation in the standard model extended by heavy singlet Dirac neutrinos. *Phys. Rev. D*, 62:036010, Jul 2000.

[11] A. Vicente. Charged lepton flavor violation beyond minimal supersymmetry. *Nuclear Physics B - Proceedings Supplements*, 248–250:20 – 25, 2014. 1st Conference on Charged Lepton Flavor Violation1st Conference on Charged Lepton Flavor Violation.

[12] Mikołaj Misiak and Jörg Urban. QCD corrections to FCNC decays mediated by Z-penguins and W-boxes. *Physics Letters B*, 451(1–2):161 – 169, 1999.

[13] Gerhard Buchalla and Andrzej J. Buras. The rare decays $K \to \pi\nu\bar{\nu}, B \to X\nu\bar{\nu}$ and $B \to l^+l^-$: an update. *Nuclear Physics B*, 548(1–3):309 – 327, 1999.

[14] C. Bobeth, T. Ewerth, F. Krüger, and J. Urban. Analysis of neutral Higgs-boson contributions to the decays $B_s \to l^+l^-$ and $\overline{B} \to \mathrm{Kl}^+l^-$. *Phys. Rev. D*, 64:074014, Sep 2001.

[15] S. Rai Choudhury and Naveen Gaur. Dileptonic decay of B_s meson in SUSY models with large tanβ. *Physics Letters B*, 451(1–2):86 – 92, 1999.

[16] Gerhard Buchalla and Andrzej J. Buras. QCD corrections to rare K^- and B-decays for arbitrary top quark mass. *Nuclear Physics B*, 400(1–3):225 – 239, 1993.

[17] Gerhard Buchalla, Andrzej J. Buras, and Markus E. Lautenbacher. Weak decays beyond leading logarithms. *Rev. Mod. Phys.*, 68:1125–1144, 1996.

[18] Andrzej J. Buras. Minimal flavour violation and beyond: Towards a flavour code for short distance dynamics. *Acta Phys.Polon.*, B41:2487–2561, 2010.

[19] Andrzej J. Buras. Relations between Δ M(s, d) and B(s, d) $\to \mu\bar{\mu}$ in models with minimal flavor violation. *Phys.Lett.*, B566:115–119, 2003.

[20] Andrzej J. Buras, Jennifer Girrbach, Diego Guadagnoli, and Gino Isidori. On the Standard Model prediction for BR(Bs,d to mu+ mu-). *Eur. Phys. J.*, C72:2172, 2012.

[21] K. Nakamura and Particle Data Group. Review of Particle Physics. *Journal of Physics G: Nuclear and Particle Physics*, 37(7A):075021, 2010.

[22] The LEP Electroweak Working Group. http://lepewwg.web.cern.ch, via Particle Data Group.

[23] R. J. Dowdall, C. T. H. Davies, R. R. Horgan, C. J. Monahan, and J. Shigemitsu. B-Meson Decay Constants from Improved Lattice Nonrelativistic QCD with Physical u, d, s, and c Quarks. *Phys. Rev. Lett.*, 110(22):222003, 2013.

[24] J. Beringer *et al.* (Particle Data Group). *PR D*, 86(010001): URL: http://pdg.lbl.gov, 2012.

[25] Marco Ciuchini, G. D'Agostini, E. Franco, V. Lubicz, G. Martinelli, F. Parodi, P. Roudeau, and A. Stocchi. 2000 CKM triangle analysis: A Critical review with updated experimental inputs and theoretical parameters. *JHEP*, 07:013, 2001.

[26] J. Charles, Andreas Hocker, H. Lacker, S. Laplace, F. R. Le Diberder, J. Malcles, J. Ocariz, M. Pivk, and L. Roos. CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories. *Eur. Phys. J.*, C41:1–131, 2005.

[27] B. Aubert. *et al.* (BABAR Collaboration). Search for decays of B^0 mesons into e^+e^-, $\mu^+\mu^-$ and $e^\pm\mu^\mp$ final states. *Phys. Rev. D*, 77:032007, Feb 2008.

[28] M.-C. Chang. *et al.* (Belle Collaboration). Search for $B^0 \to \ell^+\ell^-$ at the Belle detector. *Phys. Rev. D*, 68:111101, Dec 2003.

[29] T. Aaltonen. *et al.* (CDF Collaboration). Search for the decays $B^0_{(s)} \to e^+\mu^-$ and $B^0_{(s)} \to e^+e^-$ in CDF Run II. *Phys. Rev. Lett.*, 102:201801, May 2009.

[30] T. Bergfeld. *et al.* (CLEO Collaboration). Search for decays of B^0 mesons into pairs of leptons: $B^0 \to e^+e^-$, $B^0 \to \mu^+\mu^-$, and $B^0 \to e^\pm\mu^\mp$. *Phys. Rev. D*, 62:091102, Oct 2000.

[31] S. Chatrchyan. *et al.* (CMS Collaboration). Measurement of the $B_s^0 \to \mu^+\mu^-$ Branching Fraction and Search for $B^0 \to \mu^+\mu^-$ with the CMS Experiment. *Phys. Rev. Lett.*, 111:101804, Sep 2013.

[32] R. Aaij. *et al.* (LHCb Collaboration). Strong constraints on the rare decays $B_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$. *Phys.Rev.Lett.*, 108:231801, 2012.

[33] R. Aaij. *et al.* (LHCb Collaboration). Search for the Lepton-Flavor-Violating Decays $B_s^0 \to e^\pm\mu^\mp$ and $B^0 \to e^\pm\mu^\mp$. *Phys. Rev. Lett.*, 111:141801, Sep 2013.

[34] M.C. Chang. Search for B→ll at Belle. *Belle Note 639*, 2007.

[35] K.A. Olive *et al.* (Particle Data Group). *Chin. Phys. C*, 38:09001, (2014) and 2015 update.

[36] CMS Collaboration and LHCb Collaboration. Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data. *Nature*, 522(7554):68–72, 06 2015.

[37] Makoto Kobayashi and Toshihide Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[38] Ed. A.J. Bevan, B. Golob, Th. Mannel, S. Prell, and B.D. Yabsley. The Physics of the B Factories. *Eur. Phys. J.*, C74:3026, 2014.

[39] B. Barish *et al.* (CLEO Collaboration). Measurement of the B Semileptonic Branching Fraction with Lepton Tags. *Phys. Rev. Lett.*, 76:1570–1574, Mar 1996.

[40] J. Beringer *et al.* (Particle Data Group). *Phys. Rev. D*, 86(1528):1528, Jul 2012.

[41] J. P. Alexander *et al.* (CLEO Collaboration). Measurement of the Relative Branching Fraction of $\Upsilon(4S)$ to Charged and Neutral B-meson pairs. *Phys. Rev. Lett.*, 86:2737–2741, Mar 2001.

[42] Bohringer, T. *et al.* (CLEO Collaboration). Observation of Υ, Υ', and Υ'' at the Cornell Electron Storage Ring. *Phys. Rev. Lett.*, 44:1111–1114, Apr 1980.

[43] Finocchiaro, G. *et al.* (CLEO Collaboration). Observation of the Υ''' at the Cornell Electron Storage Ring. *Phys. Rev. Lett.*, 45:222–225, Jul 1980.

[44] I. Abe *et al.* (Belle Collaboration). The KEKB injector linac. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 499(1):167 – 190, 2003. KEK-B: The KEK B-factory.

[45] T. Abe *et al.* (Belle Collaboration). Compensation of the Crossing Angle with Crab Cavities at KEKB. In *22nd Particle Accelerator Conference (PAC 07)*, page 27, 2007.

[46] S. Kurokawa and E. Kikutani. Overview of the KEKB accelerators. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 499(1):1 – 7, 2003. KEK-B: The KEK B-factory.

[47] K. Abe *et al.* (Belle Collaboration). Observation of Large *CP* Violation in the Neutral *B* Meson System. *Phys. Rev. Lett.*, 87:7, Aug 2001.

[48] Tetsuo Abe et al. Achievements of KEKB. *PTEP*, 2013:03A001, 2013.

[49] KEKB commissioning members. History of Records. $http://www-acc.kek.jp/kekb/commissioning/record/History-of-record.html$, Last checked 11/11/2015.

[50] A. Abashian *et al.* (Belle Collaboration). The Belle detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 479(1):117 – 232, 2002. Detectors for Asymmetric B-factories.

[51] T. Kawasaki. The Belle Silicon Vertex Detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 494(1–3):94 – 101, 2002. Proceedings of the 8th International Conference on Instrumentatio n for Colliding Beam Physics.

[52] Masashi Hazumi. Performance of the Belle silicon vertex detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 473(1–2):1 – 6, 2001. Proceedings of the 9th International Workshop on Vertex Detectors.

[53] V. Chabaud *et al.* (DELPHI Collaboration). The DELPHI silicon strip microvertex detector with double sided readout. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 368(2):314 – 332, 1996.

[54] O. Toker, S. Masciocchi, E. Nygård, A. Rudge, and P. Weilhammer. VIKING, a CMOS low noise monolithic 128 channel frontend for Si-strip detector readout. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 340(3):572 – 579, 1994.

[55] R. Abe *et al.* (Belle Collaboration). Status of the Belle SVD detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 478(1–2):296 – 298, 2002. Proceedings of the ninth Int.Conf. on Instrumentation.

[56] G. Alimonti *et al.* (Belle Collaboration). The Belle Silicon Vertex Detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 453(1–2):71 – 77, 2000. Proc. 7th Int. Conf on Instrumentation for colliding Beam Physics.

[57] K. Abe *et al.* (Belle Collaboration). Letter of Intent for KEK Super B Factory. Part II: Detector. Technical report, Tech. rep. KEK Report 04-4, 2004.

[58] M. Yokoyama, H. Aihara, M. Hazumi, H. Ishino, J. Kaneko, Y. Li, D. Marlow, S. Mikkelsen, E. Nygard, H. Tajima, J. Talebi, G. Verner, and H. Yamamoto. Ra-

diation hardness of VA1 with submicron process technology. *Nuclear Science, IEEE Transactions*, 48(3):440–443, Jun 2001.

[59] Z. Natkaniec *et al.* (Belle Collaboration). Status of the Belle silicon vertex detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 560(1):1 – 4, 2006. Proceedings of the 13th International Workshop on Vertex DetectorsVERTEX 200413th International Workshop on Vertex Detectors.

[60] H. Aihara *et al.* (Belle Collaboration). Belle SVD2 vertex detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 568(1):269 – 273, 2006. New Developments in Radiation Detectors. Proceedings, 10th European Symposium on Semiconductor Detectors, Wildbad Kreuth, Germany, June 12-16, 2005.

[61] H. Aihara *et al.* (Belle Collaboration). Status and upgrade plans of the Belle silicon vertex detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 582(3):709 – 713, 2007. {VERTEX} 2006Proceedings of the 15th International Workshop on Vertex Detectors.

[62] A. Kibayashi. Status of the Belle Silicon Vertex Detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and As-*

sociated Equipment, 569(1):5 – 7, 2006. Proceedings of the 14th International Workshop on Vertex DetectorsVERTEX 200514th International Workshop on Vertex Detectors.

[63] Y. Ushiroda. Belle silicon vertex detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 511(1–2):6 – 10, 2003. Proceedings of the 11th International Workshop on Vertex Detectors.

[64] H. Hirano *et al.* (Belle Collaboration). A high-resolution cylindrical drift chamber for the KEK B-factory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 455(2):294 – 304, 2000.

[65] E Nakano. Belle PID. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 494(1–3):402 – 408, 2002. Proceedings of the 8th International Conference on Instrumentatio n for Colliding Beam Physics.

[66] T. Iijima *et al.* (Belle Collaboration). Aerogel Cherenkov counter for the BELLE detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 453(1–2):321 – 325, 2000. Proc. 7th Int. Conf on Instrumentation for colliding Beam Physics.

[67] M.H.R. Khan and *et al.* (Belle Collaboration). The monitoring system for the aerogel Cherenkov counter of the BELLE detector. *Nuclear Instruments and Methods*

in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 413(2–3):201 – 212, 1998.

[68] T. Iijima et al. Study on fine mesh PMTs for detection of aerogel Cherenkov light. *Nucl. Instrum. Meth.*, A387:64–68, 1997.

[69] H. Kichimi *et al.* (Belle Collaboration). The BELLE TOF system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 453(1–2):315 – 320, 2000. Proc. 7th Int. Conf on Instrumentation for colliding Beam Physics.

[70] H. Sagawa. The BELLE CsI calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 453(1–2):259 – 261, 2000. Proc. 7th Int. Conf on Instrumentation for colliding Beam Physics.

[71] V. Zhilich. Online Luminosity Measurements with CSI Calorimeter at the Belle Detector. *Belle Note 465*, 2001.

[72] Kenkichi Miyabayashi. Belle electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 494(1–3):298 – 302, 2002. Proceedings of the 8th International Conference on Instrumentatio n for Colliding Beam Physics.

[73] H. Ikeda *et al.* (Belle Collaboration). A detailed test of the CsI(Tℓ) calorimeter for BELLE with photon beams of energy between 20 MeV and 5.4 GeV. *Nuclear In-*

struments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 441(3):401 – 426, 2000.

[74] M. Yamaga *et al.* (Belle Collaboration). RPC systems for BELLE detector at KEKB. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 456(1–2):109 – 112, 2000. Proceedings of the 5th Int. Workshop on Resistive Plate Chambers and Related Detectors.

[75] Y. Yusa. Belle muon identification. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 598(1):183 – 186, 2009. Instrumentation for Collding Beam PhysicsProceedings of the 10th International Conference on Instrumentation for Colliding Beam Physics.

[76] Y. Ushiroda, A. Mohapatra, H. Sakamoto, Y. Sakai, M. Nakao, Q. An, and Y. F. Wang. Development of the central trigger system for the BELLE detector at the KEK B factory. *Nucl. Instrum. Meth.*, A438:460–471, 1999.

[77] K. Hanagaki, M. Hazumi, and H. Kakuno. Level 4 Software Trigger at BELLE. *Belle Note 299*, 2000.

[78] Belle Collaboration internal web page. Number of B events in Hadron B(J). http://belle.kek.jp/secured/nbb/nbb.html.

[79] S. H. Lee. *et al.* (Belle Collaboration). Evidence for $B^0 \to \pi^0\pi^0$. *Phys. Rev. Lett.*, 91:261801, Dec 2003.

[80] Yuuji Unno. Study of CP violation in charmless hadronic two-body decays of B mesons. 2004.

[81] S. Nishida. Charged and Neutral Particles. In *Software Festa 2007*. KEK, Sep. 6, 2007.

[82] L. Hinz. Lepton ID efficiency correction and systematic error. *Belle Note 954*, 2006.

[83] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto. Electron identification in Belle. *Nucl.Instrum.Meth.*, A485:490–503, 2002.

[84] S. Nishida. How to use PID Packages. In *Belle Analysis School*. KEK, Feb. 11, 2010.

[85] A. Abashian *et al.* (Belle Collaboration). Muon identification in the Belle experiment at KEKB. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 491(1–2):69 – 82, 2002.

[86] S Nishida. Study of Kaon and Pion Identification Using Inclusive D^* Sample. *Belle Note 779*, 2005.

[87] Belle internal web. Cross section of relevant processes at $\sqrt{s}$ = 10.58 GeV. http://belle.kek.jp/secured/wiki/doku.php?id=physics:taup:cross, last accessed 20 May 2015.

[88] Anders Ryd, David Lange, Natalia Kuznetsova, Sophie Versille, Marcello Rotondo, et al. EvtGen: A Monte Carlo Generator for B-Physics. 2005.

[89] David J. Lange. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors*

and Associated Equipment, 462(1–2):152 – 155, 2001. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.

[90] H. Albrecht. *et al.* (ARGUS Collaboration). Search for hadronic b→u decays. *Physics Letters B*, 241(2):278 – 282, 1990.

[91] Rene Brun and Fons Rademakers. ROOT — An object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(1–2):81 – 86, 1997. New Computing Techniques in Physics Research V.

[92] Mikihiko Nakao and Nobu Katayama. cabs3 user's manual. *Belle Note 991*, 2007.

[93] Nobu Katayama, Ryosuke Itoh, Atsushi Manabe, and Takashi Sasaki. Belle computing model. *Computer Physics Communications*, 110(1–3):22 – 25, 1998.

[94] I. Antcheva *et al.* ROOT — A C++ framework for petabyte data storage, statistical analysis and visualization. *Computer Physics Communications*, 180(12):2499 – 2512, 2009. 40 YEARS OF CPC: A celebratory issue focused on quality software for high performance, grid and novel computing architectures.

[95] ROOT Reference Guide 6.05/03. TTree Class Reference. *https://root.cern.ch/doc/master/classTTree.html*, Reference Guide Generated on Tue Nov 24 2015.

[96] R. Brun, Ph. Canal, M. Frank, A. Kreshuk, S. Linev, P. Russo, and F. Rademakers. Developments in ROOT I/O and trees. *J. Phys. Conf. Ser.*, 119:042006, 2008.

[97] Wouter Verkerke and David P. Kirkby. The RooFit toolkit for data modeling. *eConf*, C0303241:MOLT007, 2003. [186(2003)].

[98] W Verkerke and D Kirkby. RooFit Users Manual v2. 07.

[99] Ming-Chuan Chang. *Search for two-body leptonic and baryonic B decays at Belle*. PhD thesis, Taiwan, Natl. Taiwan U., 2004.

[100] Brendan Casey. HadronB. *Belle Note 390*, 2001.

[101] Bipul Bhuyan. High P_T Tracking Efficiency Using Partially Reconstructed D^* Decays. *Belle Note 1165*, 2010.

[102] Adrian Bevan and Fergus Wilson. A ROOFit User Guide. *http* : *//pprc.qmul.ac.uk/ bevan/afit/afit.pdf*.