國立臺灣大學理學院物理學系

碩士論文

Department of Physics
College of Science
National Taiwan University

Master Thesis

尋找 B^0 介子衰變至 $l^+l^-l^+l^-$ 之分析

Search for $B^0 \to l^+l^-l^+l^-$ Decay at Belle

程章銘
Chang-Ming Cheng

指導教授：張寶棣教授
Advisor: Prof. Pao-Ti Chang

中華民國 107 年 1 月
January, 2018

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

國立臺灣大學碩士學位論文
口試委員會審定書

尋找B^0介子衰變至$l^+l^-l^+l^-$之分析
Search for $B^0 \to l^+l^-l^+l^-$ Decay at Belle

本論文係程章銘君（R03222050）在國立臺灣大學物理學系、所完成之碩士學位論文，於民國 107 年 1 月 29 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

張寶棣 （簽名）

（指導教授）

徐靜戈

doi:10.6342/NTU201800725

誌謝

在這做研究的一路上我受到了許多人的幫助與鼓勵，因此才能完成這篇論文。

首先我要感謝我的指導教授張寶棣老師，有您的指導才能讓我深入認識高能物理實驗以及得到寶貴的研究經驗，讓我能了解做研究的方法與實驗精神。此外也感謝王名儒老師和徐靜戈老師願意當我的口試委員為我的研究給些評論和建議。

我還要感謝高能實驗室的同學、學長和學弟們，在我做研究時遇到困難都不吝嗇地給予許多幫助和建議，多虧你們才能讓我渡過許多難關。當然還要感謝我的朋友們，有你們的鼓勵和打氣才能讓我的研究堅持下去。

最後要感謝我的家人，正因為有你們的支持與栽培才能讓我平安順利地完成論文，謝謝你們。

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

摘要

本論文利用蒙地卡羅方法模擬日本高能加速器 B 介子工廠 (KEKB) 中之 Belle 偵測器所收集的 771 百萬個 B 介子對與背景來尋找 B^0 介子衰變至四個輕子的事件，其中包含三種衰變模式 $B^0 \to \mu^+\mu^-\mu^+\mu^-$、$B^0 \to e^+e^-e^+e^-$ 和 $B^0 \to e^+e^-\mu^+\mu^-$。我們使用 NeuroBayes 演算法來分離訊號與背景以提升訊噪比，接著再以 B^0 介子衰變至一個 J/ψ 介子、一個 K 介子和一個 π 介子，J/ψ 介子再衰變至一個輕子對的事件作為對照組來確認分析步驟是否正確並得到校正因子，然後利用校正因子作為 $B^0 \to l^+l^-l^+l^-$ 之中模擬與數據的校正。最後我們估計在 90% 的信賴區間之下的衰變機率上限為：$\mathcal{B}(B^0 \to \mu^+\mu^-\mu^+\mu^-) < 1.28 \times 10^{-7}$，$\mathcal{B}(B^0 \to e^+e^-e^+e^-) < 8.34 \times 10^{-8}$ 和 $\mathcal{B}(B^0 \to e^+e^-\mu^+\mu^-) < 5.09 \times 10^{-8}$。

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

Abstract

This thesis uses Monte Carlo method to simulate 771 million $B\bar{B}$ pairs and background collected in Belle detector at KEKB to search for $B^0 \to l^+l^-l^+l^-$, it includes $B^0 \to \mu^+\mu^-\mu^+\mu^-$, $B^0 \to e^+e^-e^+e^-$ and $B^0 \to e^+e^-\mu^+\mu^-$. We use NeuroBayes algorithm to separate signal from background so that it enhances the signal-to-noise ratio, then we do $B^0 \to J/\psi(J/\psi \to l^+l^-)K^+\pi^-$ as the control sample to check whether the analysis is correct or not and get the calibration factor, and then we use calibration factor to be the correction between simulation and data in $B^0 \to l^+l^-l^+l^-$. Finally, we estimate the upper limit of branching fraction at 90% confidence interval: $\mathcal{B}(B^0 \to \mu^+\mu^-\mu^+\mu^-) < 1.28 \times 10^{-7}$, $\mathcal{B}(B^0 \to e^+e^-e^+e^-) < 8.34 \times 10^{-8}$ and $\mathcal{B}(B^0 \to e^+e^-\mu^+\mu^-) < 5.09 \times 10^{-8}$.

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

Contents

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

List of Figures

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

List of Tables

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

Chapter 1

Introduction

1.1 Particle Physics

The purpose of particle physics is to search for the most basic composition in matter. In 1897, J.J. Thomson found the electron in cathode ray experiment, in 1919, Ernest Rutherford found the proton by using α particle collided the Nitrogen. In 1920s, people considered the matter consisting of the protons and electrons. However, since 1930s, many experiments such as nuclear physics and cosmic ray, scientists found many kinds of elementary particles. To search for the elementary particles became a complicated problem, so the particle physics was born in 20th century.

Nowadays, particle physics includes experiment and theory, the experimental particle physicists do the research by using accelerator which can produce high energy particles collision to research production and decay phenomenon of high energy particles. High energy detector in accelerator can detect high energy particles which can help physicists to search for the properties and interactions. Therefore, the accelerator is important for experimental particle physicists to do the research.

The theoretical particle physicists use physical theory and mathematics to explain the experimental results and develop the new physical laws. The modern theory in particle physics such as Standard Model is not perfect. For example, the Standard Model can't explain and include dark matter perfectly, it only explains the observable matter in the universe. The final goal of theoretical particle physicists is to search for the theory of

doi:10.6342/NTU201800725

everything.

1.2 Standard Model

The Standard Model is currently a dominant theory of particle physics. The theory is established by relativistic quantum field theory in 1970s. The particle physics experimental results and theoretical predictions can be explained by Standard Model. In Standard Model, there are four kind classes of elementary particles: quark, lepton, gauge boson and Higgs boson. Besides, there are three fundamental interactions: electromagnetic interaction, strong interaction and weak interaction (Fig. 1.1).

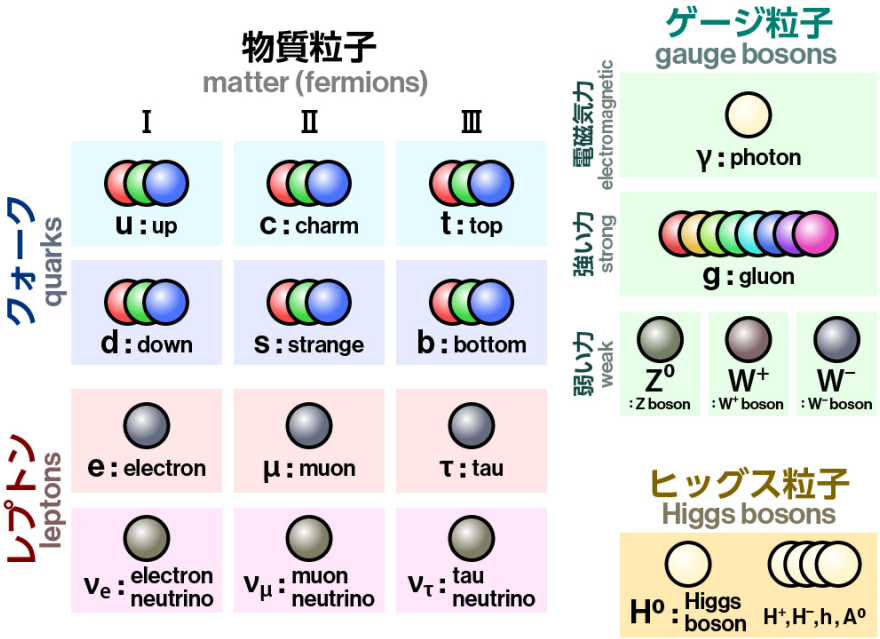


Figure 1.1: Standard Model [1].

There are three generations of quarks and leptons. The quarks,including up (u), down (d), charm (c), strange (s), top (t) and bottom (b) carry fractional unit charges, +2/3 or -1/3, and the leptons, including electron (e), electron neutrino (ν_e), muon (μ), muon neutrino (ν_μ), tau (τ) and tau neutrino (ν_τ) carry integral charge,-1 or 0. Both quarks and leptons carry spin 1/2. The fundamental interactions due to exchange gauge bosons between elementary particles. The electromagnetic interaction acts on quarks, charged leptons and

doi:10.6342/NTU201800725

W^+, W^- due to the photon (γ). The strong interaction acts on quarks and gluon due to the gluon (g). The weak interaction acts on all fermions due to the Z^0 and $W^\pm$ bosons.

The Higgs bosons (H) is postulated by Higgs mechanism, which explains the origination of matter mass. In July 2012, the European Organization for Nuclear Research (CERN) found new boson which liked Higgs boson. In March 2013, the CERN confirmed the new boson which was Higgs boson, the prediction of Higgs mechanism is successful. The interactions between elementary particles is shown in Fig. 1.2.

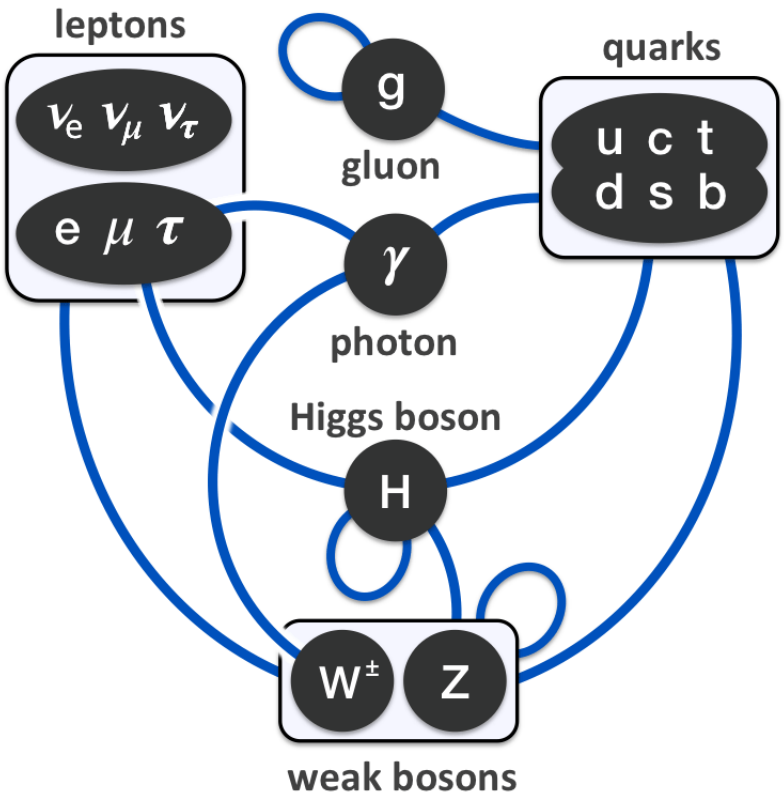


Figure 1.2: Interactions between elementary particles [2].

1.2.1 Cabibbo-Kobayashi-Maskawa Matrix

In Standard Model, the d, s and b quarks can change their flavor via weak interaction. In 1963, Nicola Cabibbo established Cabibbo angle to describe the second generation quarks. In 1973, Makoto Kobayashi and Toshihide Maskawa extended Cabibbo angle to Cabibbo-Kobayashi-Maskawa (CKM) matrix, it can describe the third generation quarks.

doi:10.6342/NTU201800725

The CKM matrix can explain CP violation. In mathematics, quarks changing their flavor can be represented as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.1}$$

where the 3×3 unitary matrix is called CKM matrix.

All the entries in CKS matrix can be parameterized, and we can get the magnitude of each entry from experimental results [3]

$$V_{CKM} = \begin{pmatrix} 0.97428 \pm 0.00015 & 0.2253 \pm 0.0007 & 0.00347^{+0.00016}_{-0.00012} \\ 0.2252 \pm 0.0007 & 0.97345^{+0.00015}_{-0.00016} & 0.0410^{+0.0011}_{-0.0007} \\ 0.00862^{+0.00026}_{-0.00020} & 0.0403^{+0.0011}_{-0.0007} & 0.999152^{+0.000030}_{-0.000045} \end{pmatrix}. \tag{1.2}$$

1.2.2 CP Violation

In physics, the symmetry means a physical quantity that it is invariant in value under a operation. In particle physics, physicists have found many new conservation, such as parity (P), lepton number (L), baryon number (B) and charge conjugation symmetry (C). Some of quantities can't obey its conservation in special conditions. For example, in weak interaction, it will happen parity violation.

CP symmetry is the product of charge conjugation symmetry and parity symmetry. CP symmetry is also not a conservation absolutely. In 1964, J.W. Cronin and V.R. Fitch found the K_L^0 meson decay has probability of 3/1000 so that it happened CP violation. CP symmetry and time reversal symmetry (T) is called CPT symmetry, CPT symmetry is an absolute conservation law in particle physics. Therefore, the CP violation implies time reversal violation.

In cosmology, why is the amount of matter much larger than the antimatter? The antimatter consists of antiparticles, the charge and quantum number of antiparticle are opposite to its corresponding particle. CP violation can provide a reasonable solution. If the universe obeys CP symmetry absolutely, the amount of matter would same as antimatter,

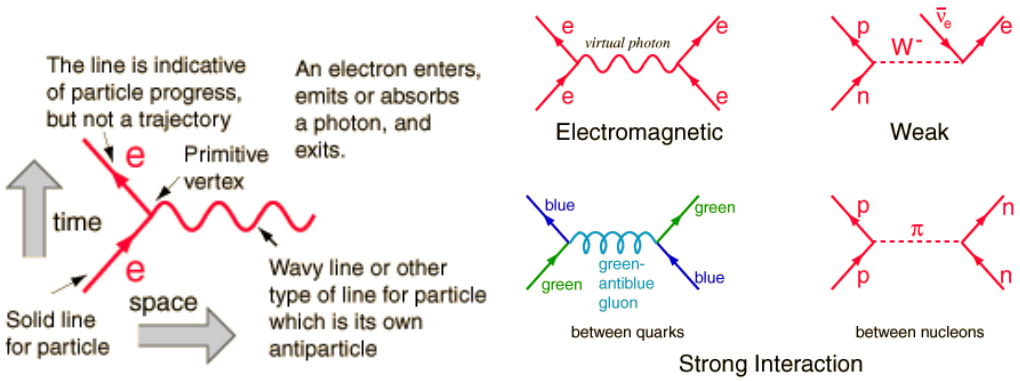
doi:10.6342/NTU201800725

it happens pair annihilation so that the universe doesn't have any matter, but it is not true.

According to experimental results, CP violation is classified two kinds: indirect CP violation and direct CP violation. Currently, the SLAC National Accelerator Laboratory and High Energy Accelerator Research Organization (KEK) have observed direct CP violation in B meson decay [4].

1.2.3 Feynman Diagram

Feynman diagram can represent the behavior of elementary particles and interactions graphically. It was invented by Richard Feynman in the middle of the 20th century. There are some rules in Feynman diagram. The time axis is upward and space axis to the right. (In particle physics often reverse that orientation.) Fermions are represented by solid lines, photons are represented by wavy lines, bosons are represented by dash lines and gluons are represented by helical lines, an example rules are shown in Fig. 1.3(a). Two or more than lines converge a point which is called a vertex, interactions can represent graphically between two vertices (Fig. 1.3 (b)) [5].



(a) Basic rules for Feynman diagram. (b) Vertex rules for Feynman diagram.

Figure 1.3: Feynman diagram.

1.3 B Meson Physics

In 1973, Makoto Kobayashi and Toshihide Maskawa introduced a bottom quark (b), which was the third generation quark in order to explain CP violation. In 1977, Fermilab E288

doi:10.6342/NTU201800725

experiment team discovered the bottom quark from Upsilon meson, the notation is Υ(1S), which consists of bottom quark and its antiparticle [6]. There are many resonant states of Upsilon system. For example, Υ(1S), Υ(2S), Υ(3S) and Υ(4S), they can be created by high energy accelerator. Fig. 1.4 shows the spectrum of hadron production, the figure originated from Upsilon Spectroscopy: Transitions in the Bottomonium System by D. Besson and T. Skwarnicki [7].

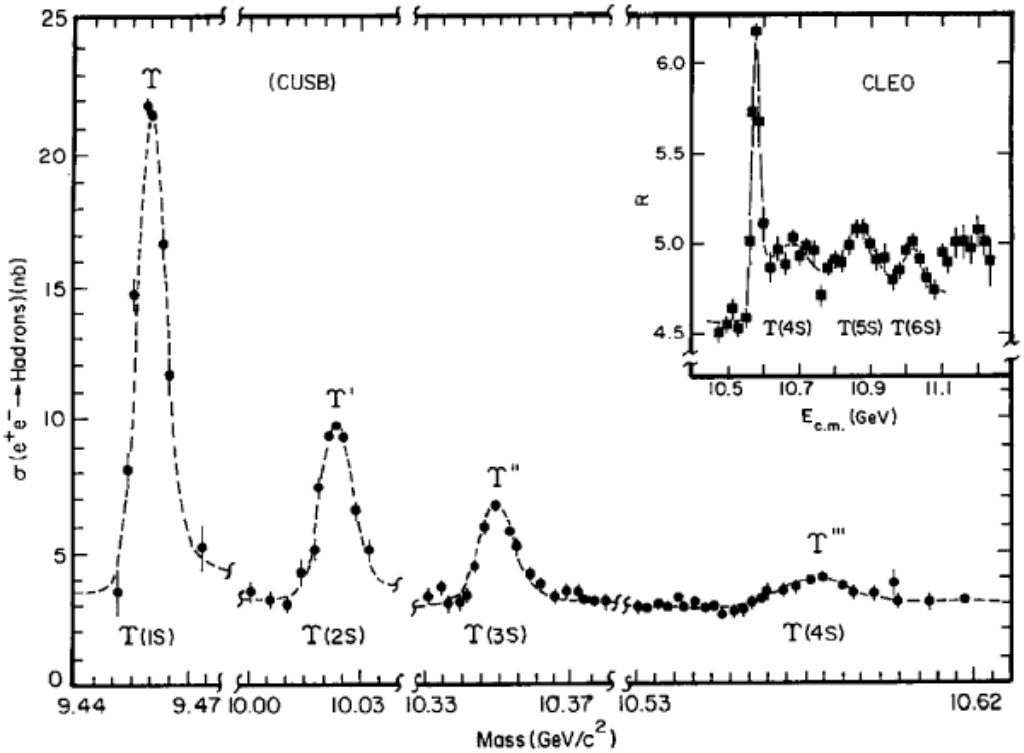


Figure 1.4: The spectrum shows the cross section for inclusive production of hadrons as a function of center-of-mass energy. In plot, there are many peaks correspond with each resonant states of Upsilon system.

The B meson consists of a bottom antiquark and other a quark. There are four kinds of B mesons, a charged B meson (B^+) consists of a bottom antiquark and a up quark (u), a neutral B meson (B^0) consists of a bottom antiquark and a down quark (d), a strange B meson (B_s^0) consists of a bottom antiquark and a strange quark (s), a charmed B meson (B_c^+) consists of a bottom antiquark and a charm quark (c). Table 1.1 shows the properties of B mesons and its antiparticles.

In order to research CP violation of B meson, the KEK and SLAC built a e^+e^- colliders, which can produce asymmetry energy of an e^+e^- pair. The $B\overline{B}$ pair can be created

doi:10.6342/NTU201800725

Particle	Antiparticle	Quark content	Isospin	Rest mass (Mev/c^2)	lifetime (ps)
B^+	B^-	$u\bar{b}$	$\frac{1}{2}$	5279.29±0.15	1.638±0.004
B^0	$\overline{B}^0$	$d\bar{b}$	$\frac{1}{2}$	5279.61±0.16	1.520±0.004
B_s^0	$\overline{B}_s^0$	$s\bar{b}$	0	5366.79±0.23	1.510±0.005
B_c^+	B_c^-	$c\bar{b}$	0	6275.1±1.0	0.507±0.009

Table 1.1: Properties of B mesons.

through the reaction $e^+e^- \to \Upsilon(4S) \to B\overline{B}$. The mass of $\Upsilon(4S)$ is 10.5794 ± 0.0012 GeV/c^2 and branching fraction of $\Upsilon(4S)$ decay to $B\overline{B}$ is larger than 96% at 95% confidence interval [8]. The decay process has threshold energy. In center-of-mass frame, the energy of colliding beams including electrons and positrons at least equal $\Upsilon(4S)$ resonance. Fig. 1.5 shows the Feynman diagram of this reaction.

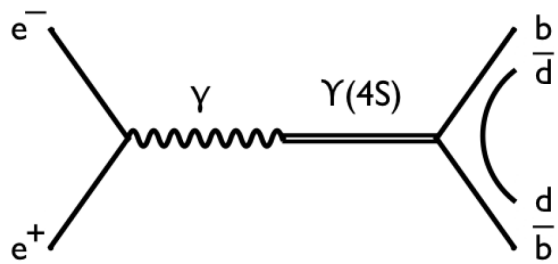


Figure 1.5: Feynman diagram for $e^+e^- \to \Upsilon(4S) \to B\overline{B}$.

1.4 Motivation

The rare decay $B^0_{(s)} \to l^+l^-l^+l^-$ is predict by Standard Model [9] and minimal supersymmetric model (MSSM) [10]. A B meson decay to four leptons such as $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ has resonant SM channel and nonresonant SM channel. For the resonant mode, the process is $B^0_s \to J/\psi\phi(1020)$ and then both the J/ψ and $\phi(1020)$ decay to two muons. We can calculate the branching fraction for $B^0_s \to J/\psi(\to \mu^+\mu^-)\phi(1020)(\to \mu^+\mu^-)$ by product of $\mathcal{B}(B^0_s \to J/\psi\phi)$, $\mathcal{B}(J/\psi \to \mu^+\mu^-)$ and $\mathcal{B}(\phi \to \mu^+\mu^-)$ [8], the result value is $(2.3\pm0.9)\times10^{-8}$. The nonresonant SM mode is $B^0_{(s)} \to \mu^+\mu^-\gamma(\to \mu^+\mu^-)$, $B^0_{(s)}$ decay to one muon pair via virtual photon and the other muon pair via box diagram or electroweak penguin, the branching fraction for $B^0_{(s)} \to \mu^+\mu^-\gamma(\to \mu^+\mu^-)$ is less than the order of

doi:10.6342/NTU201800725

10^{-10} [11]. Fig. 1.6 shows the Feynman Diagram for resonant and nonresonant SM mode for $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$.

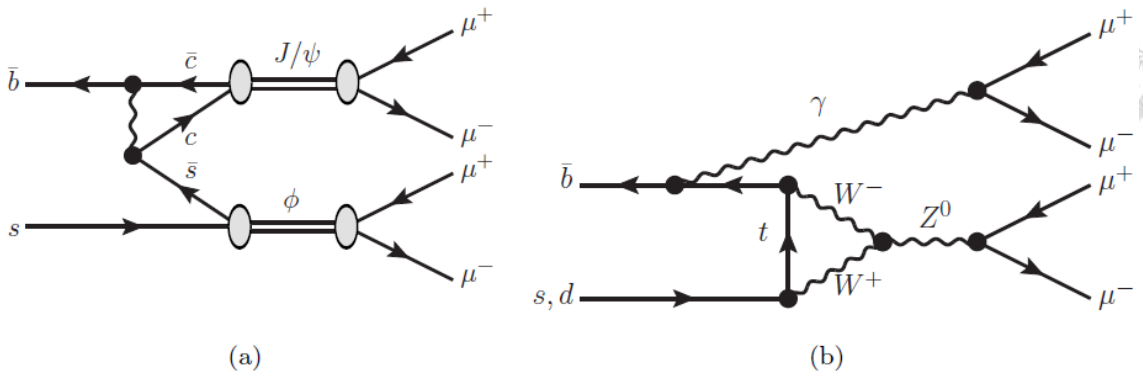


Figure 1.6: Feynman diagram for $B^0 \to \mu^+\mu^-\mu^+\mu^-$ and $B^0_s \to \mu^+\mu^-\mu^+\mu^-$ decay, (a) is resonant SM channel and (b) is nonresonant SM channel [12].

This thesis focus on $B^0 \to l^+l^-l^+l^-$ excluding $B^0_s \to l^+l^-l^+l^-$, it will cover $B^0 \to \mu^+\mu^-\mu^+\mu^-$, $B^0 \to e^+e^-e^+e^-$ and $B^0 \to e^+e^-\mu^+\mu^-$. The branching fraction of these modes are expected to be less than 10^{-10} in SM model [11]. Thus, enhancing the branching fraction can discover new physics beyond the SM model. For example, minimal supersymmetric model (MSSM).

In MSSM, $B^0_{(s)}$ decay to two muon pairs via scalar S and pseudoscalar P particles respectively. S and P are supersymmetric fermions in MSSM, the Feynman Diagram is shown in Fig. 1.7.

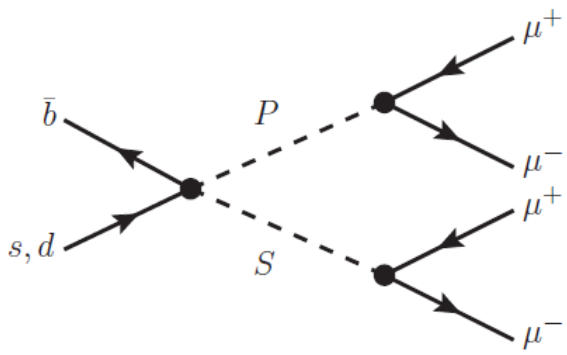


Figure 1.7: Feynman diagram for MSSM channel of $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ [12].

The LHCb collaboration measured that the upper limit of branching fraction is $\mathcal{B}(B^0 \to \mu^+\mu^-\mu^+\mu^-) < 6.6(5.3) \times 10^{-9}$ at 95% (90%) confidence interval in SM model [12]. In

doi:10.6342/NTU201800725

this thesis, we use Monte Carlo method and Belle data to search the $B^0 \to l^+l^-l^+l^-$ in following chapters.

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

Chapter 2

Belle Experiment

Nowadays, high energy physics experiments often are international collaboration, Belle experiment, which locates at High Energy Accelerator Research Organization (KEK) in Tsukuba, Japan is one of them. There are many research areas in KEK, the purpose of Belle experiment is searching CP violation in B mesons. This chapter will introduce KEKB accelerator and Belle detectors. An aerial photo of KEK is shown below.

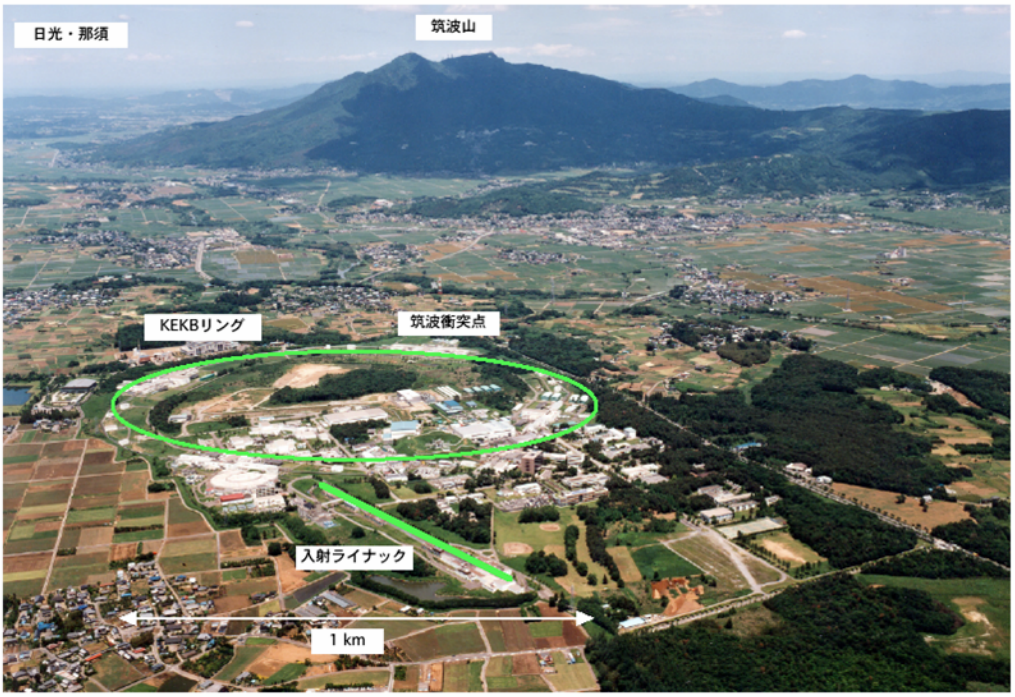


Figure 2.1: An aerial photo of KEK.

doi:10.6342/NTU201800725

2.1 KEKB Accelerator

KEK B-factory, is called KEKB for short, is an asymmetric energy of positron-electron collider at KEK. There are two operation rings in KEKB: a low-energy ring (LER) and a high-energy ring (HER). HER can accelerate electrons to reach 8.0 GeV and LER can accelerate positrons to reach 3.5 GeV. The total length of rings is 3016 m, there are four straight sections in the circular tunnel: Tsukuba area, Oho area, Fuji area and Nikko area, the Belle experiment locates at Tsukuba area. In this area, electron beam collides positron beams that we call interaction point (IP). Two beams collide each other at IP with ± 11 mrad (Fig. 2.2). Beam current in LER reaches 1.8 A and HER reaches 1.3 A [13], it will generate 10.58 GeV energy in centre-of-mass frame, it's equal to the mass of $\Upsilon(4S)$.

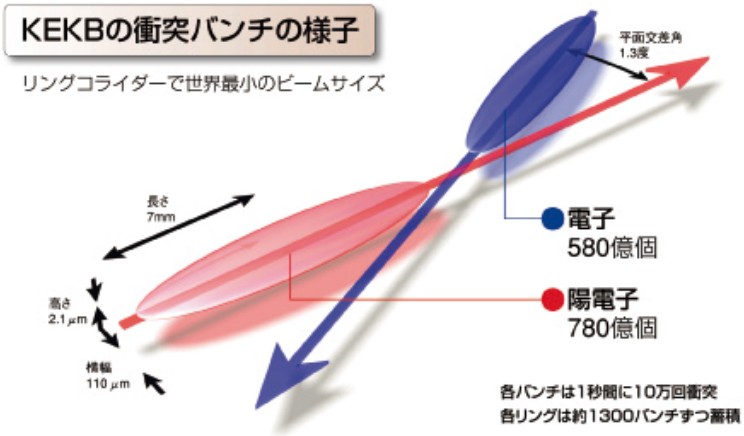


Figure 2.2: Two bunches collision at interaction point.

In particle physics experiments, luminosity means collision frequency per cross section, KEKB is designed to operate with maximum luminosity at $1.0 \times 10^{34} cm^{-2} s^{-1}$, it has $10^8 B\bar{B}$ pairs per year. Now, the maximum luminosity in KEKB has reached 2.1083×10^{34}, total integrated luminosity has reached 1052 fb^{-1}, which were recorded by Belle detector [13]. More parameters of KEKB accelerator are shown in Table 2.1.

In December 1998, the HER started operating, in June 1999, Belle detector measured the first particle interaction. Now, the KEKB accelerator and Belle detector are updating to Belle II and SuperKEKB, the luminosity will reach 10^{40} and help us to search more new physics. The configuration of KEKB accelerator as shown in Fig. 2.3.

doi:10.6342/NTU201800725

Ring	Notation	LER	HER	Unit
Energy	E	3.5	8.0	GeV
Circumference	C	3016.26		m
Luminosity	L	1×10^{34}		$cm^{-2}s^{-1}$
Crossing angle	θ_x	± 11		mrad
Tune shifts	ξ_x/ξ_y	0.039/0.052		
Beta function at IP	β_x^*/β_y^*	0.33/0.01		m
Beam current	I	2.6	1.1	A
Natural bunch length	σ_z	0.4		cm
Energy spread	σ_ε	7.1×10^{-4}	6.7×10^{-4}	
Bunch spacing	s_b	0.59		m
Particles/bunch	N	3.3×10^{10}	1.4×10^{10}	
Emittance	$\varepsilon_x/\varepsilon_y$	$1.8 \times 10^{-8}/3.6 \times 10^{-10}$		
Synchrotron	ν_s	$0.01 \sim 0.02$		
Betatron tune	ν_x/ν_y	45.52/45.08	47.52/43.08	
Momentum compaction factor	α_p	$1 \times 10^{-4} \sim 2 \times 10^{-4}$		
Energy loss/turn	U_0	$0.81^\dagger/1.5^{\dagger\dagger}$	3.5	MeV
RF voltage	V_c	$5 \sim 10$	$10 \sim 20$	MV
RF frequency	f_{RF}	508.887		MHz
Harmonic number	h	5120		
Longitudinal damping time	τ_ε	$43^\dagger/23^{\dagger\dagger}$	23	ms
Total beam power	P_b	$2.7^\dagger/4.5^{\dagger\dagger}$	4.0	MW
Radiation power	P_{SR}	$2.1^\dagger/4.0^{\dagger\dagger}$	3.8	MW
HOM power	P_{HOM}	0.57	0.15	MW
Bending radius	ρ	16.3	104.5	m
Length of bending magnet	l_B	0.915	5.86	m

Table 2.1: Parameters of KEKB accelerator, †: without wigglers, ††: with wigglers.

doi:10.6342/NTU201800725

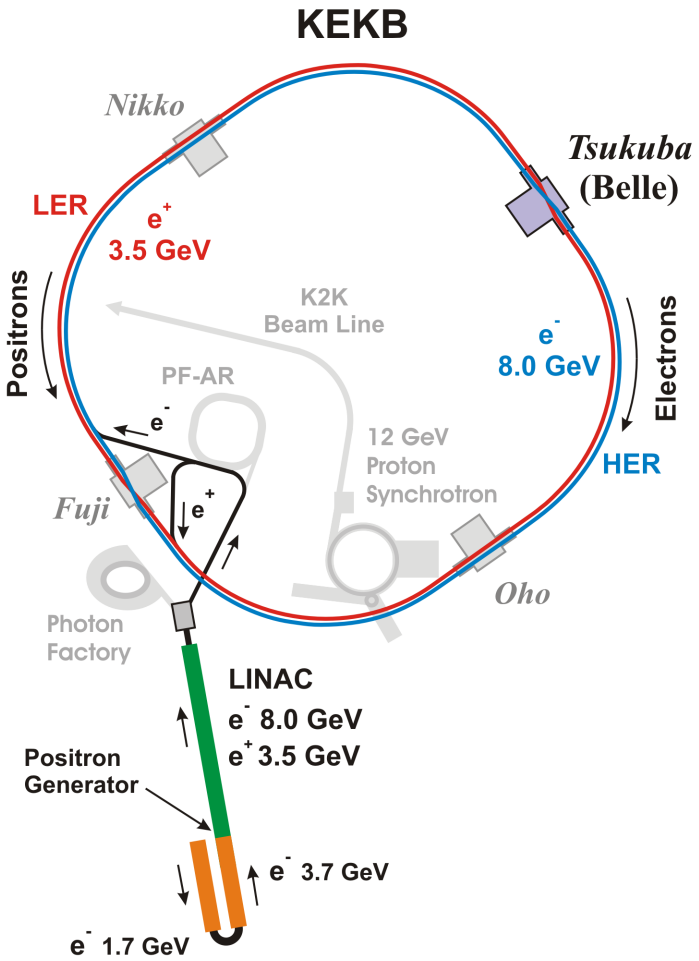


Figure 2.3: Configuration of KEKB accelerator. The beams are accelerated in linear accelerator and then into circular tunnel, the IP locates at Tsukuba area [14].

2.2 Belle Detector

In Belle experiment, we often set up a specific coordinate system to describe directions:

- z: direction of the HER axis,
- y: direct to sky,
- x: perpendicular to y-z plain.

We also use cylindrical (r, ϕ, z) and spherical coordinates (r, θ, ϕ) to express the directions. θ is polar angle and ϕ is azimuthal angle, they are calculated from z axis and x axis respectively. Besides, $r = \sqrt{x^2 + y^2}$.

doi:10.6342/NTU201800725

Belle detector consists of a 1.5 T superconductor solenoid and an iron structure surrounding the electron and positron beams at IP. There are many components of Belle detector [15], the silicon vertex detector (SVD) can measure B meson decay vertices and SVD locates just outside of a cylindrical beryllium (Be) beam pipe. Charged particle tracking information is provided by the central drift chamber (CDC). Besides, CDC, aerogel Cherenkov counters (ACC) and time-of-flight counters (TOF) provide dE/dx measurements so that it can get the particle identification (PID), ACC and TOF locate outside of CDC radially. The electromagnetic calorimeter (ECL) consisting of an array of CsI(Tl) crystals locates inside the solenoid coil. K_L and muon detector (KLM) can identify the K_L^0 mesons and muons, KLM locates outside of the solenoid coil. The coverage of θ region is extending from $17°$ to $150°$ for the detector. In order to measure the uncovered small angle in forward and backward directions, we design the extreme forward calorimeter (EFC), which has a pair of BGO crystal arrays located on the surfaces of the QCS cryostats in the forward and backward directions. The θ coverage of each detector is shown in Table 2.2, Table 2.3 shows more details of Belle detector and Fig. 2.4, Fig. 2.5 show sectional drawing of Belle detector [15].

Detector	Total	forward	barrel	backward
CDC	$17° - 150°$		$17° - 150°$	
ACC	$17° - 127°$	$17° - 34°$	$34° - 127°$	
ECL	$12° - 155°$	$12° - 31.4°$	$32.2° - 128.7°$	$130.7° - 155.1°$
KLM (inner)	$25° - 145°$	$25° - 51°$	$51° - 117°$	$117° - 145°$
KLM (outer)	$17° - 145°$	$17° - 51°$	$51° - 117°$	$117° - 155°$

Table 2.2: θ coverage of CDC, ACC, ECL and KLM [16].

2.2.1 Beam Pipe

Measuring CP violation is a dominant goal for Belle experiment. It is important to get the precise measurement of the decay vertex (decay point). In order to obtain a good solution on z-vertex position, we require the thickness of beam pipe as thin as possible. Moreover, because the vertex resolution improves inversely with the distance to the first detection layer, the vertex detector has to be placed as close to the IP as possible [15].

doi:10.6342/NTU201800725

Detector	Type	Configuration	Readout	Performance
Beam pipe for DS-I	Beryllium double wall	Cylindrical, r = 20mm, 0.5/2.5/0.5(mm) = Be/He/Be w/ He gas cooled		
Beam pipe for DS-II	Beryllium double wall	Cylindrical, r = 15mm, 0.5/2.5/0.5(mm) = Be/PF200/Be		
EFC	BGO	Photodiode readout Segmentation : 32 in ϕ; 5 in θ	160×2	Rms energy resolution: 7.3% at 8 GeV 5.8% at 2.5 GeV
SVD1	Double-sided Si strip	3-layers: 8/10/14 ladders Strip pitch: 25(p)/50(n)μm	ϕ: 40.96k z: 40.96k	$\sigma(z_{CP}) \sim 78.0\mu$m for $B \to \phi K^0 2_s$
SVD2	Double-sided Si strip	4-layers: 6/12/18/18 ladders Strip pitch: 75(p)/50(n)μm (layer1-3) 73(p)/65(n)μm (layer4)	ϕ: 55.29k z: 55.296k	$\sigma(z_{CP}) \sim 78.9\mu$m for $B \to \phi K_s^0$
CDC	Small cell drift chamber	Anode: 50 layers Cathode: 3 layers r = 8.3 - 86.3 cm $-77 \leq z \leq 160$ cm	Anode: 8.4k Cathod: 1.8k	$\sigma_{r\phi} = 130\mu$m $\sigma_z = 200 \sim 1400\mu$m $\sigma_{Pt}/Pt = 0.3\%\sqrt{p_t^2+1}$ $\sigma_{dE/dx} = 0.6\%$
ACC	Silica aerogel	960 barrel/228 end-cap FM-PMT readout		$N_{p.e.} \geq 6$ K/π seperation: $1.2 < p < 3.5$GeV/c
TOF	Scintillator	128 ϕ segmentation r = 120 cm, 3-cm long	128×2	σ_t = 100 ps K/π seperation:
TSC		64 ϕ segmentation	64	up to 1.2 GeV/c
ECL	CsI (Towered-structure)	Barrel: r = 125 - 162 cm End-cap: z = -102 cm and +196cm	6624 1152(F) 960(B)	$\sigma_E/E = 1.3\%/\sqrt{E}$ $\sigma_{pos} = 0.5$ cm/$\sqrt{E}$ (E in GeV)
KLM	Resistive plate counters	14 layers (5 cm Fe + 4cm gap) 2 RPCs in each gap	θ: 16k ϕ: 16k	$\Delta\phi = \Delta\theta = 30$mr for K_L $\sim$ 1% hadron fake
Magnet	Supercon.	Inner radius = 170 cm		B=1.5T

Table 2.3: Performance parameters for the Belle detector. There are two configurations of inner detectors used to collect two data sets, DS-I and DS-II, corresponding to a 3-layer SVD1 and a 4-layer SVD2 with a smaller beam pipe respectively.

doi:10.6342/NTU201800725

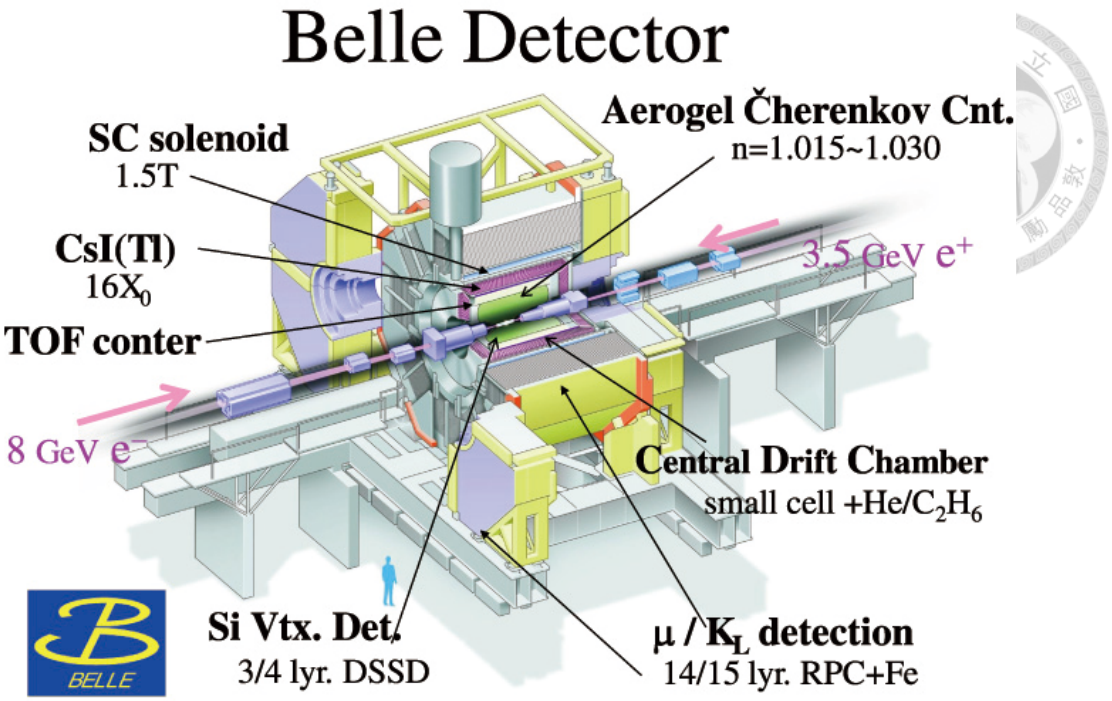


Figure 2.4: 3D sectional drawing of Belle detector.

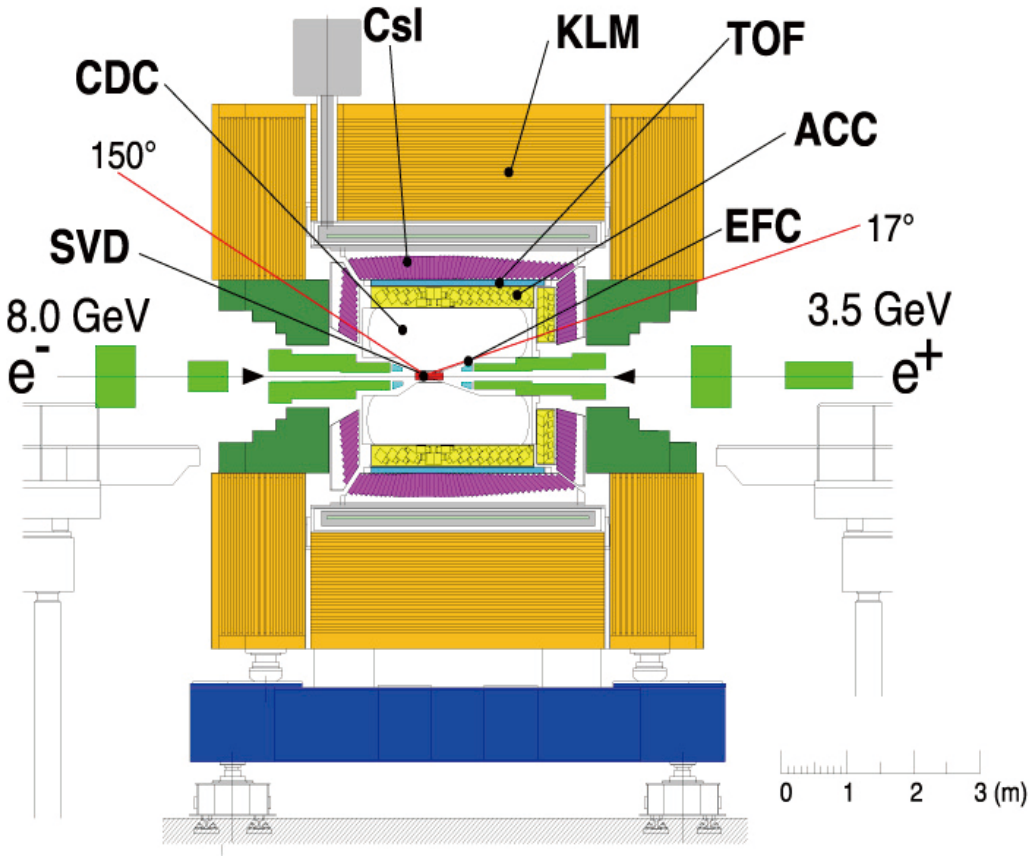


Figure 2.5: 2D sectional drawing of Belle detector.

doi:10.6342/NTU201800725

The beam pipe has an outer wall and an inner wall, the thickness of each beryllium wall is 0.5 mm. Between inner and outer walls the Helium (He) gas is filled in this gap in order to cool down the beam pipe. Fig. 2.6 shows the cross-section of the beam pipe near the IP [15].

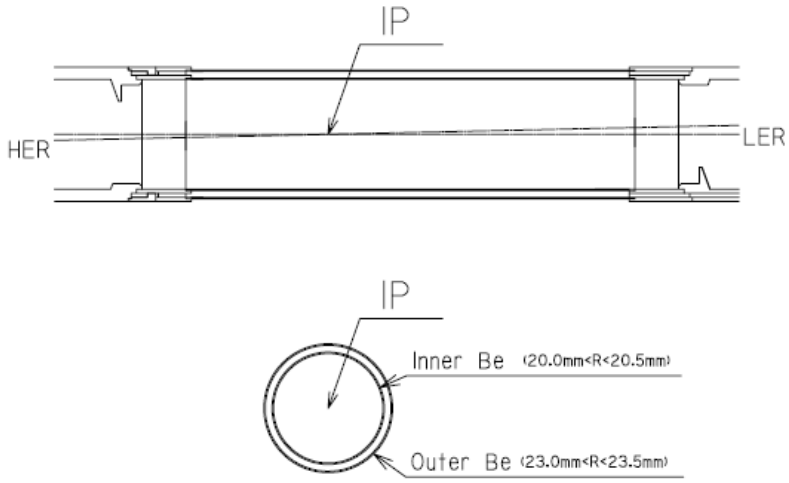


Figure 2.6: The cross-section of the beryllium beam pipe at the IP.

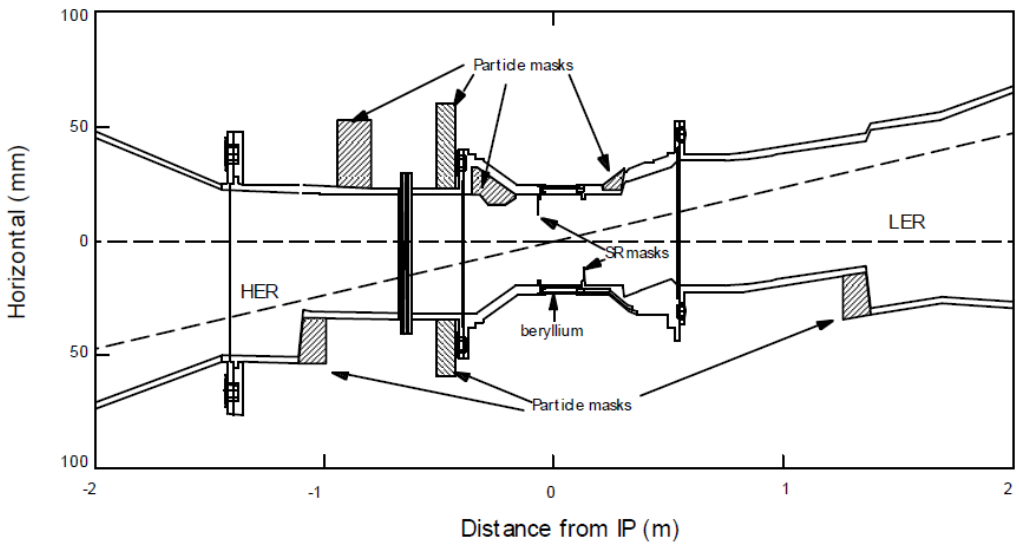


Figure 2.7: The arrangement of the beam pipe and horizontal masks.

The synchrotron radiation backgrounds are eliminated because of the separation-bend magnets near IP. Besides, when synchrotron radiation from QCS and QC1 pass through the walls, radiation won't hit them due to the well-designed apertures of beam pipe. Fig. 2.7 shows the configuration of beam pipe [15].

doi:10.6342/NTU201800725

2.2.2 Silicon Vertex Detector (SVD)

In order to observe time-dependent CP violation precisely in Belle experiment, it requires precision of $\sim 100\ \mu m$ when it measured the difference in z-vertex positions for B meson pairs. The SVD has old SVD and new SVD, old SVD is called SVD1 and new SVD is called SVD2. SVD1 consists of three layers of double-sided silicon strip detectors (DSSD), each layer has 8, 10 and 14 ladders. The radius of three layers are 30 mm, 45.5 mm, 60.5mm and the θ coverage is $23^\circ < \theta < 139^\circ$. It corresponds 86% of full solid angle. Each ladder has some DSSDs and there are 1280 sense strips and 640 readout pads at opposite sides for each DSSD. Fig. 2.8 shows the outline of SVD [15].

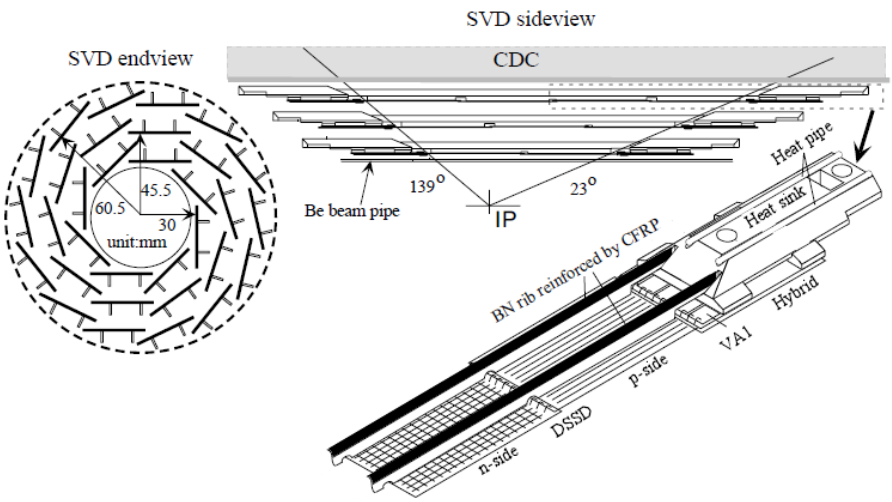


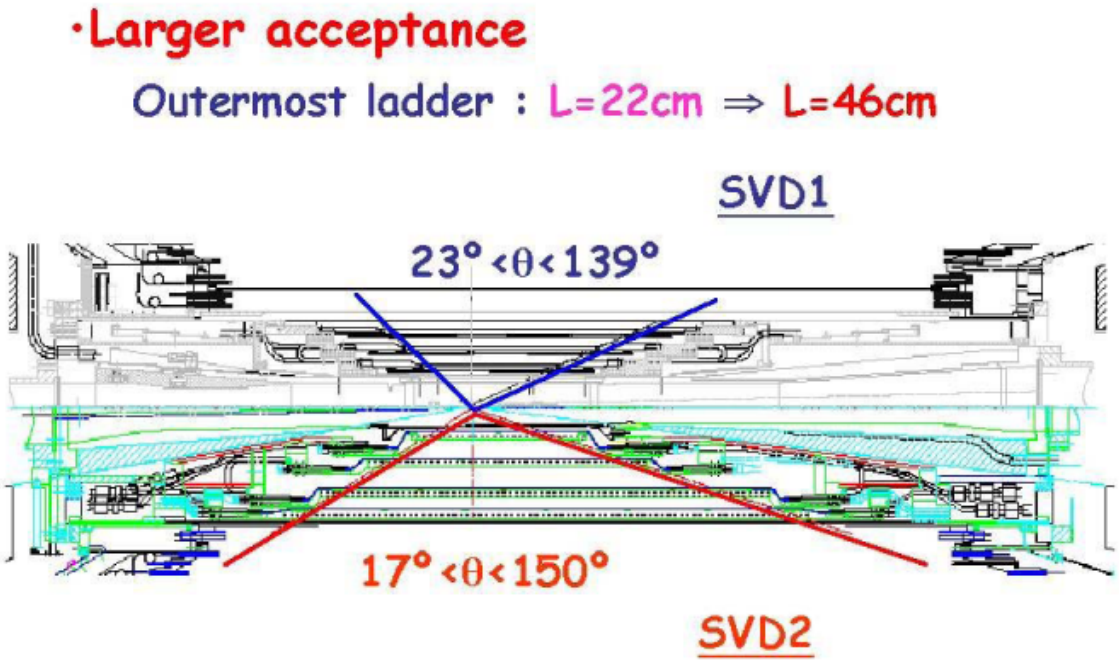
Figure 2.8: Side view and end view of SVD.

In 2003 the SVD1 was upgraded to SVD2 [17]. SVD2 consists of fours layers of DSSDs, each layer has 6, 12, 18 and 18 ladders. The θ coverage is $17^\circ < \theta < 150^\circ$, it is larger than SVD1. The performance of SVD2 such as detecting efficiency, vertex resolution and radiation tolerance were improved. The comparison between SVD1 and SVD2 are shown in Fig. 2.9 [15].

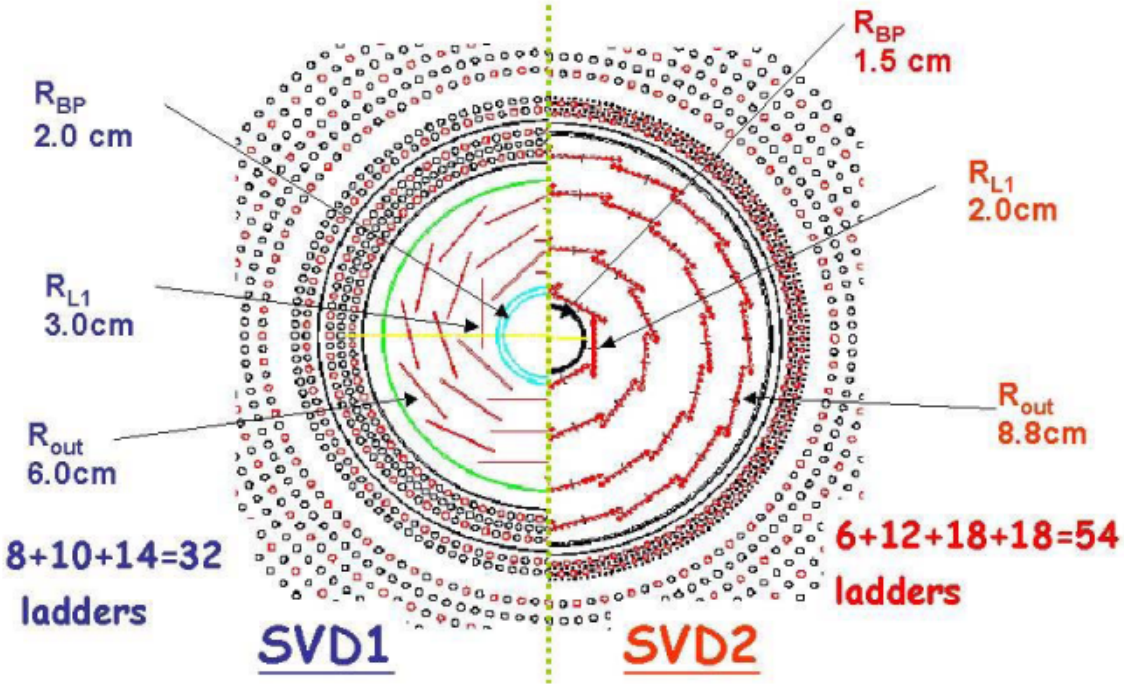
2.2.3 Extreme Forward Calorimeter (EFC)

The EFC is equipped at extreme forward side and extreme backward side surrounding the beam pipe. EFC can improve experimental sensitivity for some physics decay such as

doi:10.6342/NTU201800725



(a) Side view comparison of SVD1 and SVD2.

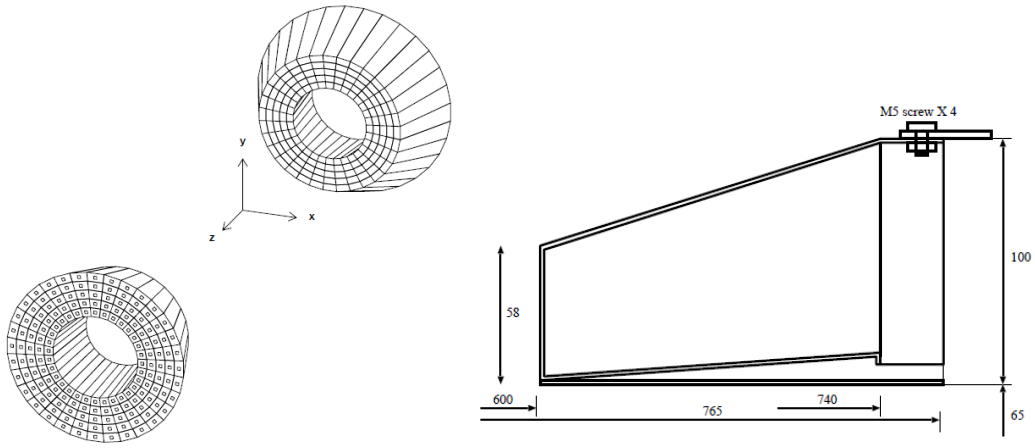


(b) End view comparison of SVD1 and SVD2.

Figure 2.9: Sectional drawing of SVD1 and SVD2.

doi:10.6342/NTU201800725

$B \to \tau\nu$ and extend the coverage of polar angle. In addition, EFC is also a beam mask to reduce backgrounds for CDC, a beam monitor for KEKB control and a luminosity monitor. For forward region, the θ coverage is $6.4° < \theta < 11.5°$ and backward region is $163.3° < \theta < 171.2°$. Because EFC locates near IP and there has very high radiation level, EFC need to tolerate high radiation. Therefore, BGO (Bismuth Germanate, $Bi_4Ge_3O_{12}$) crystal is a good material for making a EFC. Moreover, BGO has good e/γ energy resolution of $(0.3-1)\%/\sqrt{E(GeV)}$. The side view of EFC and BGO crystals arrangement are shown below [15].



(a) The BGO crystals of the forward and backward EFC detectors.

(b) Side view of the mounting of forward EFC.

Figure 2.10: Outline of the EFC.

2.2.4 Central Drift Chamber (CDC)

The CDC is used for determining the momentum of charged particles precisely and reconstructing charged particle tracks when the charged particle passed the coverage of CDC, the θ coverage is $17° < \theta < 150°$. We can get the transverse momentum (p_t) by observing curvatures in the transverse plane. Besides, from track information, we will get the momentum of charged particle in z direction (p_z). Moreover, the dE/dx measurements provide PID and useful information for trigger. It is required for a momentum resolution of $\sigma_{p_t}/p_t \sim 0.5\%\sqrt{1+p_t^2}$ (p_t in GeV/c) for all charged particles with $p_t \geq 100MeV/c$ in the coverage of CDC [15].

Fig. 2.11 shows the structure of CDC. The cylindrical shape that inner radius is 103.5

doi:10.6342/NTU201800725

mm and outer is 874 mm. In addition, it has 50 layers, 3 cathode strip layers and 8400 drift cells. The more information of CDC layers are listed in Table 2.4. The cathode strip on the cylinder walls can read out the signals from the drift cells in the inner layers. The cell arrangement is shown in Fig. 2.12 [15].

Superlayer type and no.	No. of layers	Signal channels per layer	Radius (mm)	Stereo angle (mrad) and strip pitch (mm)
Cathode	1	$64(z) \times 8(\phi)$	83.0	(8.2)
Axial 1	2	64	88.0-98.0	0.
Cathode	1	$80(z) \times 8(\phi)$	103.0	(8.2)
Cathode	1	$80(z) \times 8(\phi)$	103.5	(8.2)
Axial 1	4	64	108.5-159.5	0.
Stereo 2	3	80	178.5-209.5	71.46-73.75
Axial 3	6	96	224.5-304.0	0.
Stereo 4	3	128	322.5-353.5	-42.28- -45.80
Axial 5	5	144	368.5-431.5	0.
Stereo 6	4	160	450.5-497.5	45.11-49.36
Axial 7	5	192	512.5-575.5	0.
Stereo 8	4	208	594.5-641.5	-52.68- -57.01
Axial 9	5	240	656.5-719.5	0.
Stereo 10	4	256	738.5-785.5	62.10-67.09
Axial 11	5	288	800.5-863.0	0.

Table 2.4: Configurations of the CDC sense wires and cathode strips.

The Coulomb scattering can affect momentum resolution. Thus, we use low atomic number gas in CDC. It has a 50% helium - 50% ethane mixture gas in CDC, it has a good dE/dx resolution due to the large portion of ethane. The dE/dx is used for identifying the kinds of charged particles. For example, the momentum below 0.5 GeV/c, the CDC can separate kaons (K) and pions (π). The two dimensional plot for dE/dx versus momentum shows the example in Fig. 2.13 [15].

2.2.5 Aerogel Cherenkov Counters (ACC)

Except the CDC, the ACC is also a PID system for separating the K and π. Though the K and π are identified by dE/dx measurement from CDC and time-of-flight measurement from TOF, the ACC extends the momentum coverage over the CDC and TOF [15].

In the barrel region along the ϕ direction, the ACC has 960 counter modules segmented

doi:10.6342/NTU201800725

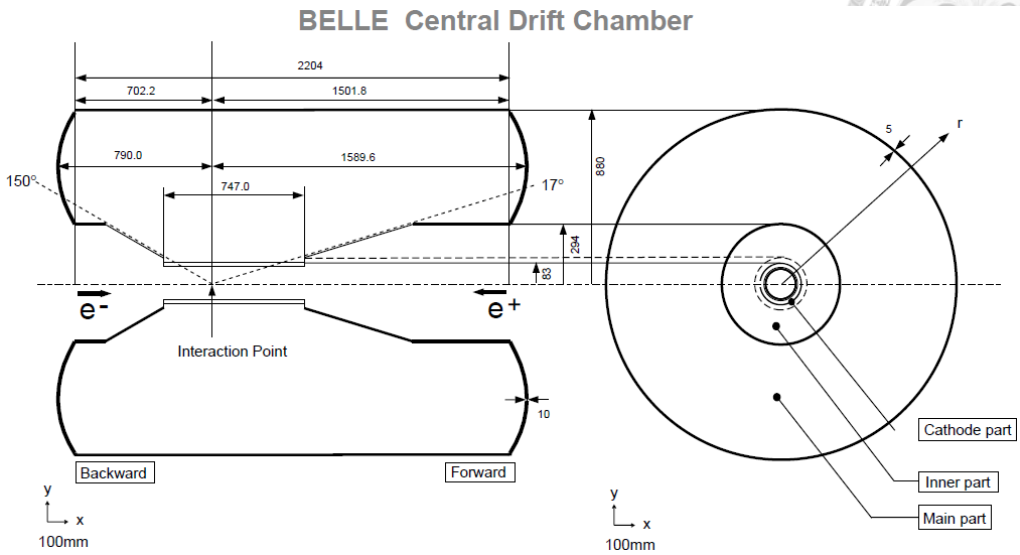
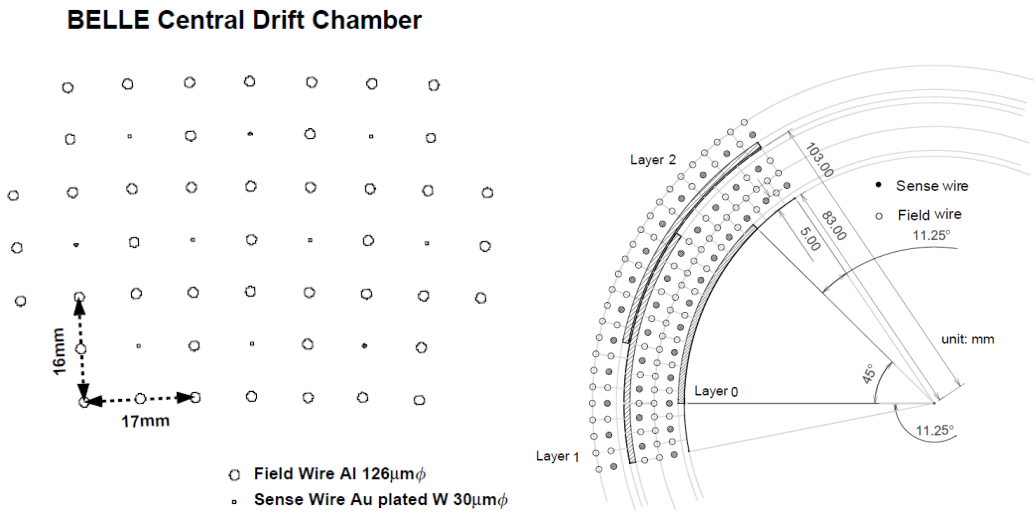


Figure 2.11: Overview of the CDC structure. The lengths in the figure are in units of mm.



(a) Cell structure of CDC.

(b) Cell structure and the cathode sector configuration.

Figure 2.12: Cell arrangement for CDC.

doi:10.6342/NTU201800725

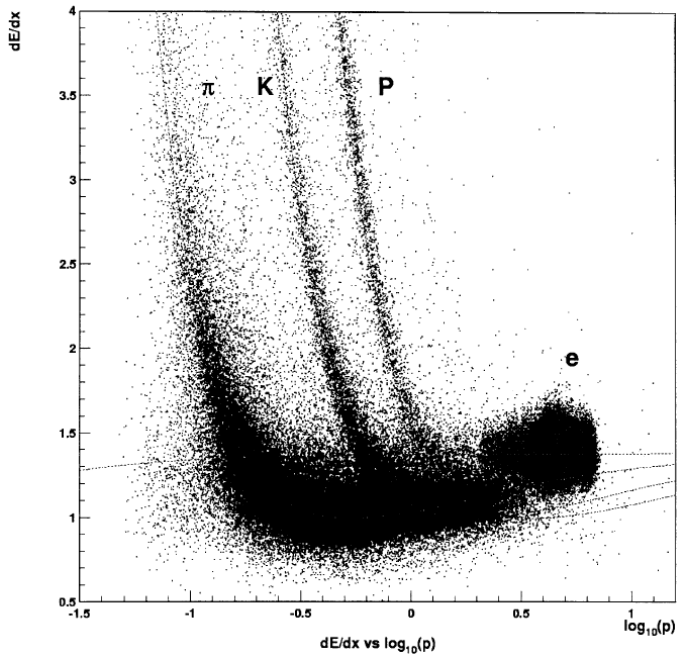


Figure 2.13: Truncated mean of dE/dx vs. momentum observed in collision data.

into 60 cells. Besides, there are 228 counter modules segmented in the forward end-cap region of ACC. The ACC covers the θ coverage of $17° < \theta < 127°$. Fig. 2.14 shows arrangement of ACC. A typical single ACC module is shown in Fig. 2.15. The refractive index of aerogels is 1.01 to 1.03. Moreover, the fine mesh-type photomultiplier tubes (FM-PMTs), which is attached to the aerogels can be used for observing Cherenkov radiation [15].

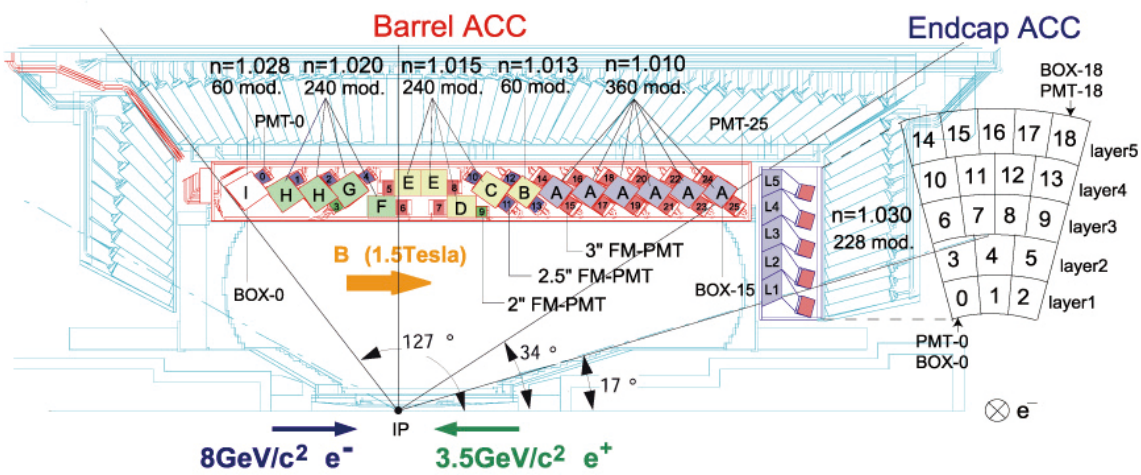
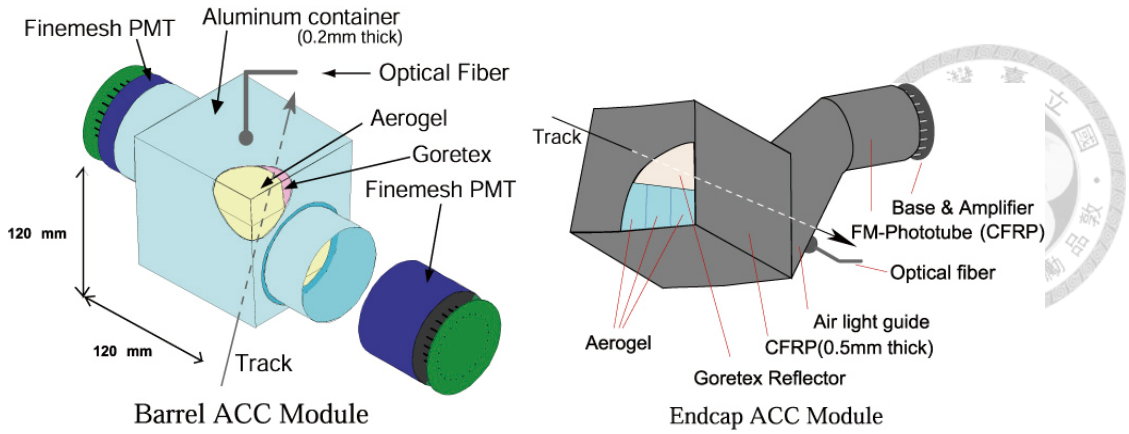


Figure 2.14: The arrangement of ACC at the central part of the Belle detector.

doi:10.6342/NTU201800725



(a) Barrel ACC Module. (b) End-cap ACC Module.

Figure 2.15: Schematic drawing of a typical ACC counter module.

2.2.6 Time-of-Flight Counters (TOF)

The TOF system is a powerful system for PID. For the particle momentum below 1.2 GeV/c, TOF has a time resolution with 100 ps and a 1.2 m flight path, there are 90% of particles from the $\Upsilon(4S)$ decays to be covered by TOF. This system also provides fast timing signals for trigger system. If the fast trigger rate keeps below 70 kHz in beam background, TOF should be increased by thin trigger scintillation counters (TSC) [15].

The TOF system, which has a θ coverage of $34^\circ < \theta < 120^\circ$ consists of 128 TOF counters and 64 TSC counters. The minimum p_t of charged particles reaching the TOF counters is about 0.28 GeV/c. Fig. 2.16 shows the TOF/TSC module and parameters of TOF and TSC are listed in Table. 2.5 [15].

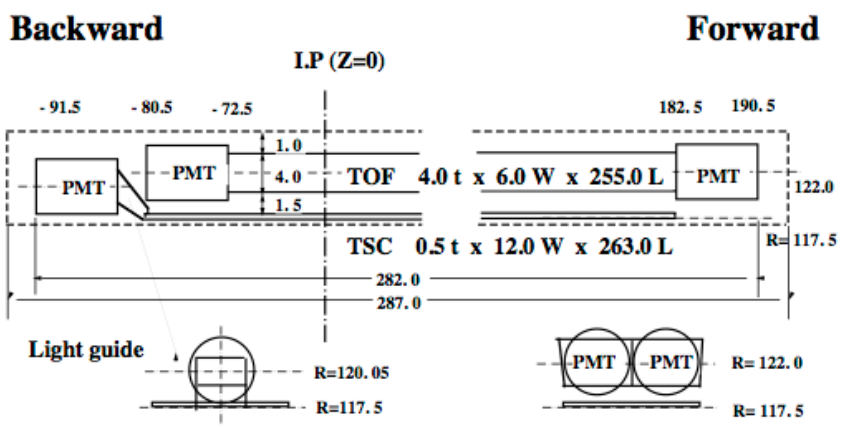


Figure 2.16: Configuration of a TOF/TSC module.

doi:10.6342/NTU201800725

Counter	Thickness	z coverage (cm)	r (cm)	ϕ segm.	No. of PMTs
TOF	4.0	-72.5 to +182.5	122.0	128	2
TSC	0.5	-80.5 to +182.5	117.5	64	1

Table 2.5: Parameters of the TOF and TSC counters.

The TOF can get the mass distribution for each track in hadron events, it is calculated by this equation

$$M^2 = (\frac{1}{\beta^2} - 1)P^2 = ((\frac{cT_{obs}^{twc}}{L_{path}})^2 - 1)P^2, \tag{2.1}$$

where T_{obs}^{twc} is time walk correction to get a precise observed time, P is momentum and L_{path} is path length of the charged particle collected from the CDC track. Fig. 2.17 shows the result of mass distribution by the TOF system [15].

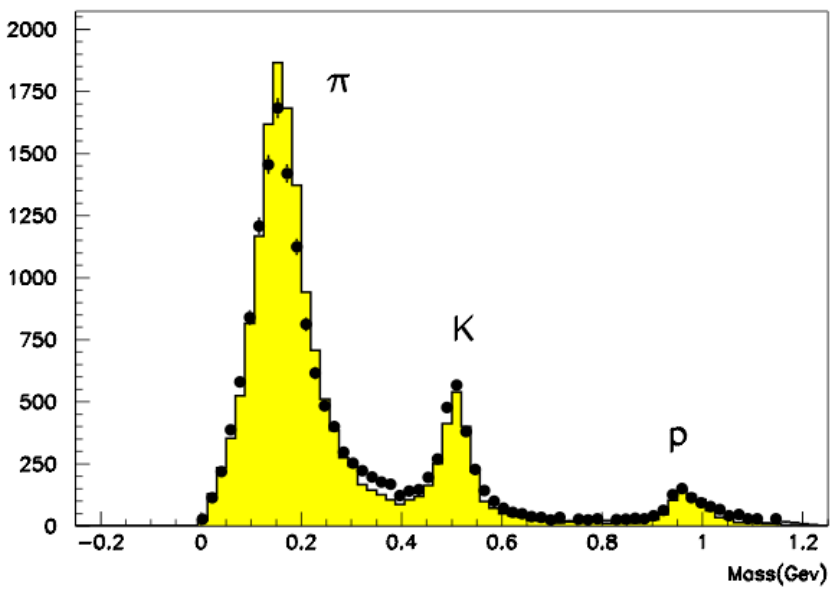


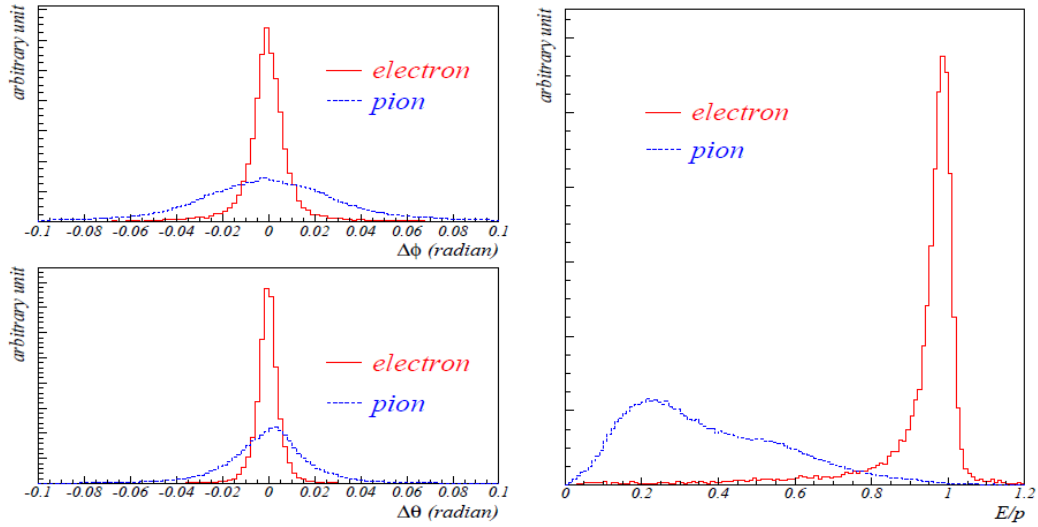
Figure 2.17: Mass distribution from TOF measurements for particle momentum below 1.2 GeV/c.

2.2.7 Electromagnetic Calorimeter (ECL)

The ECL is used for detecting photons from B meson decays. The main information has high efficiency and good resolutions in energy and position. Most of these photons have low energies relatively, thus, it is important for ECL with good performance below 500

doi:10.6342/NTU201800725

MeV. In addition, electron identification depends on comparison of energy and momentum of charged particle which deposits in the ECL [15]. Some useful variables for electron identification are shown in Fig. 2.18 [18].



(a) $\Delta\phi$ and $\Delta\theta$ for electrons and pions.

(b) E/p for electrons and pions.

Figure 2.18: Useful variables for electron identification.

The constitution of the ECL has a barrel and two end-caps which locate at $z = -1.0$ m and $z = +2.0$ m from the IP. There are many CsI(Tl) crystals with silicon photodiode readout arranging a array on the ECL. In order to avoid photons passing through the gap between these crystals, the crystals on the barrel section has a small tilt in θ and ϕ , also the crystals with a small tilt in θ direction on the End-cap. Fig. 2.19 shows the construction of ECL and constructive parameters are listed in Table 2.6 [15].

Item	θ coverage	θ seg.	ϕ seg.	No. of crystals
Forward end-cap	$12.4° - 31.4°$	13	48−144	1152
Barrel	$32.2° - 128.7°$	46	144	6624
Backward end-cap	$130.7° - 155.1°$	10	64−144	960

Table 2.6: Constructive parameters of ECL.

When the ECL installed into the Belle detector, it had some calibration. In 1998, the calibration was monitored by using cosmic-rays. In 1999, the energy resolution was calibrated by the Bhabha and $e^+e^- \to \gamma\gamma$ events. The energy resolution reached to 1.7% for barrel ECL, 1.74% for forward ECL and 2.85% for backward ECL [15].

doi:10.6342/NTU201800725

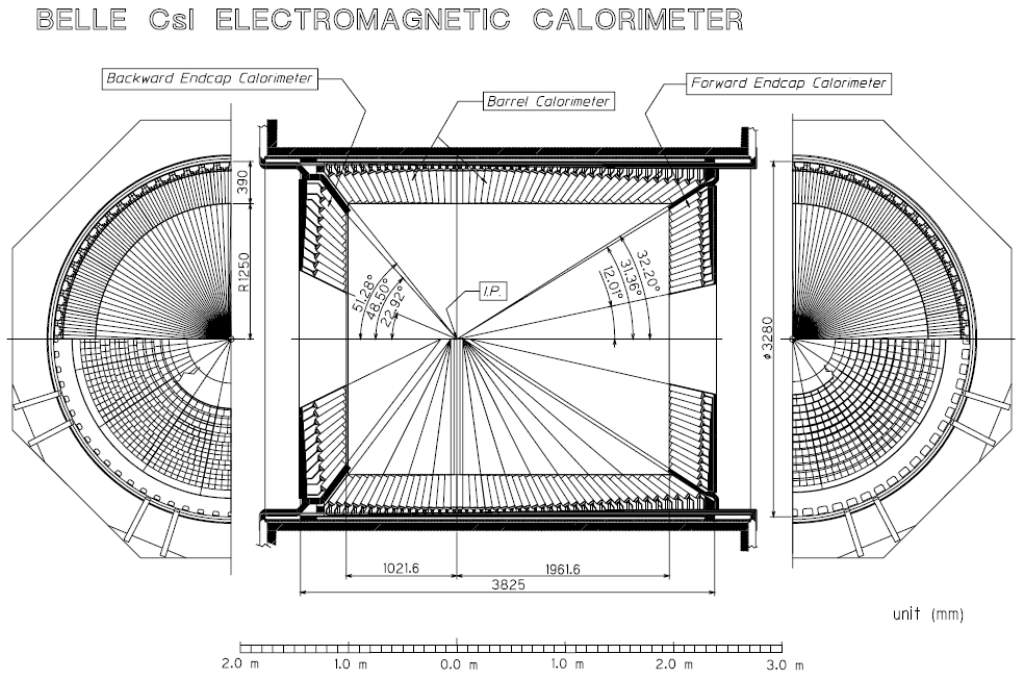


Figure 2.19: Construction of ECL.

2.2.8 K_L and Muon Detector (KLM)

The main function of KLM is that it can identify K_L's and muons with a momentum range over 600 MeV/c. The θ coverage of barrel-shaped region around the IP is $45° - 125°$, the end-caps in the forward and backward directions can extend θ coverage to $20° - 155°$. KLM consists of layers of charged particle detectors and iron plates with thickness of 4.7 cm. The iron plates and ECL provide interaction lengths for K_L's. When K_L interacts in the iron plates or ECL, it can produce a shower of ionizing particles. Using location of this shower can measure direction of K_L but the energy. Moreover, due to the multiple layers, it can distinguish between muons and other charged hadrons [15].

Each detector layer consists of the glass-electrode-resistive plate counters (RPCs). RPCs have two parallel plate electrodes which has bulk resistivity$\geq 10^{10}\Omega$ cm. There are gas-filled gaps between each plate electrodes. When an ionizing particle passes through the gap, the plates have a local discharge. According to local discharge, RPCs can record location and time of the ionizing particle. The construction of KLM is shown in Fig. 2.20 [15].

doi:10.6342/NTU201800725

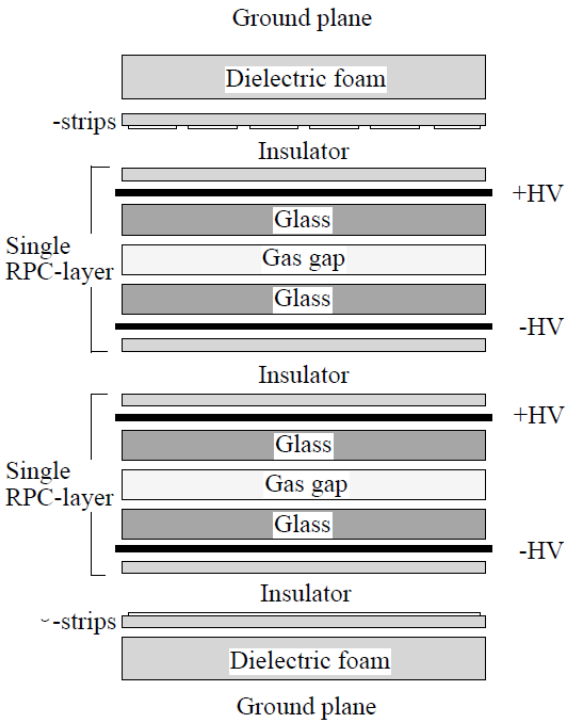


Figure 2.20: Cross-section of a KLM superlayer.

doi:10.6342/NTU201800725

doi:10.6342/NTU201800725

Chapter 3

B Event Reconstruction

3.1 Analysis Tools

In this section, it includes the software framework and some basic software for Belle experiment analysis. These tools are convenient to physics analysis.

3.1.1 BASF

In Belle experiment, Belle Analysis Framework (BASF), which can process event data is a software framework [19]. Fig. 3.1 shows the schematic view of the BASF architecture. The BASF consists of BASF user interface and BASF kernel. The user interface is separated from BASF kernel. When a message from users sends to user interface, the BASF kernel is controlled by user interface.

BASF has some important function: module and path structure, dynamic linking of modules, integrated event-by-event parallel processing capability on the SMP-sever, multi-language support for module and unified data access method by Panther. The Panther, which is a memory management system is used for data process. Users can write analysis codes or use software packages, these can be written as a module, and the module is plugged into BASF. The module is written in C++, C language or Fortran. When module links BASF kernel, it will process input data.

doi:10.6342/NTU201800725

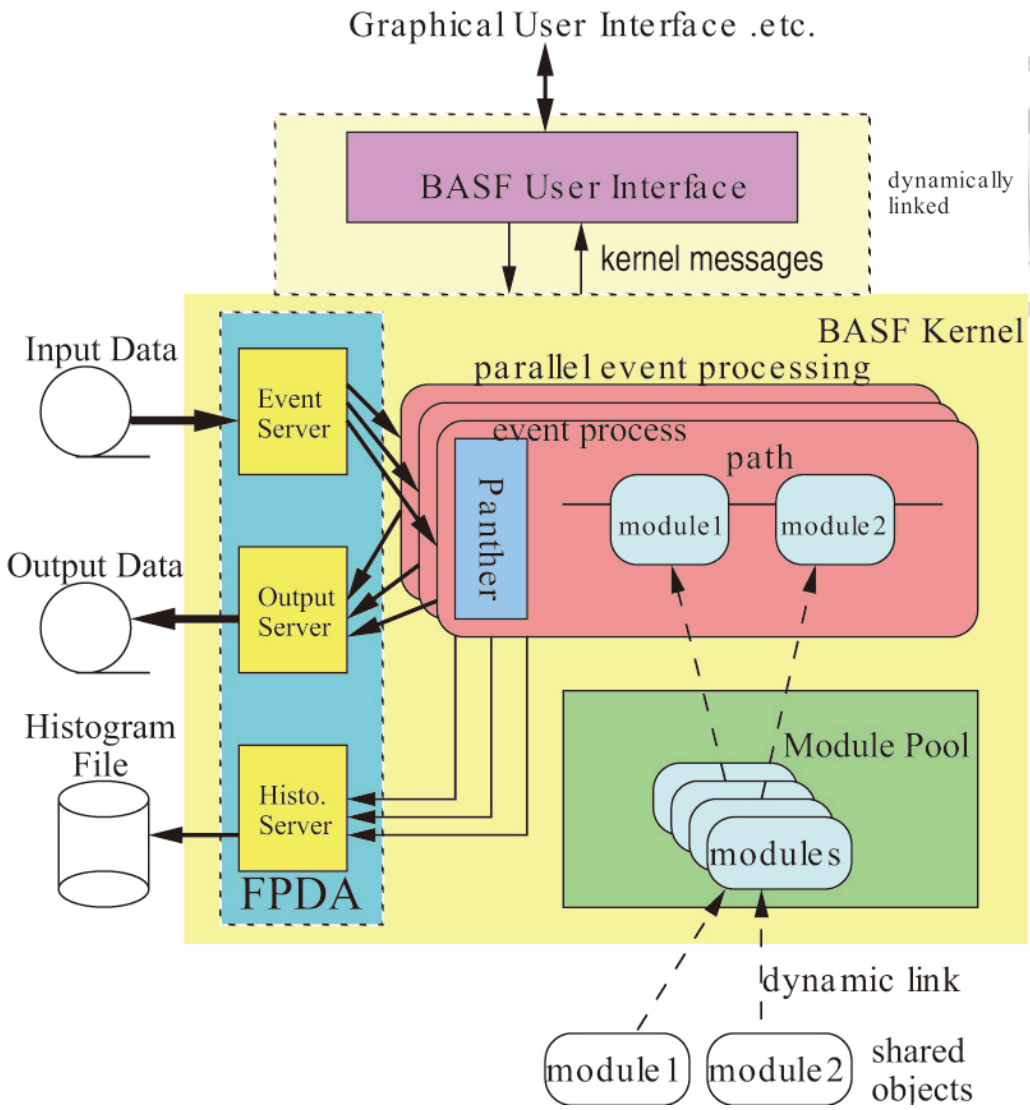


Figure 3.1: Schematic view of the BASF architecture.

3.1.2 EvtGen

Event generator can simulate the decay process. There are many event generator packages in particle physics experiment such as EvtGen [20]. EvtGen, which is written in C++ is initiated by CLEO [21] and developed by BaBar [22]. EvtGen package provides a framework for the implementation of B mesons and other resonances decay. Thus, it suits for study many details such as semileptonic decays, CP-violating decays and sequential decays in B meson physics [23].

In addition, EvtGen can produce background simulation such as $q\bar{q}$ events for B meson decay study. It also can set a new decay process or a new particle for event generation. In

doi:10.6342/NTU201800725

this thesis, all modes of Monte Carlo run 77100 events.

3.1.3 GSIM

GEANT-based detector simulation module (GSIM), which is developed by CERN [24] is able to simulate the Belle detector. The GEANT system simulates the behavior of detector when particles pass through and act on detector. GEANT is designed for particle physics experiments originally, now it is applying to other sciences and engineering areas. The most important two functions of GEANT are [25]:

- When particles pass through, it can simulate detector response.
- The particle trajectories can be represented graphically.

In BASF, we produce events by EvtGen and set for GSIM, and then we can start decay analysis.

3.1.4 ROOT

The high energy physics experiments usually have a big data process. Thus, CERN developed the ROOT package for data analysis in 1990s. ROOT, which is written in C++ is an object-oriented framework [26] and it assembles many tool packages.

The rootfit, which is a part of ROOT packages developed by BaBar originally provides the main purpose such as modelling the distribution of events in particle physics experiments and it can simulate by Monte Carlo method [27] for these physical models. Furthermore, it has many mathematical tools for processing distribution of event models and curve fitting.

3.2 Blind Analysis

In experiments, expectancy bias [28], which makes the experimenter interferes the steps casually due to the expected result so that the observer gets a invalid consequence. Avoiding this situation, we adopt the blind analysis that is able to eliminate expectancy bias

doi:10.6342/NTU201800725

without looking at the answer. Therefore, we choose the analysis based on Monte Carlo method instead of the real data.

There are some different blind analysis methods in particle physics experiment. One of them is that we avoid looking the signal box of the real data. In following sections, this thesis will show the Monte Carlo method for signal and background study.

3.3 Data Sample

The KEKB accelerator generates the asymmetric energy (electron reaches 8.0 GeV and positron reaches 3.5 GeV) e^+e^- pairs, they collide each other and decay to $\Upsilon(4S)$ mesons with integrated luminosity 710 fb^{-1}. It corresponds to a total number of $B\bar{B}$ events with 771.581 million. These events are collected by the Belle detector and this thesis uses these data for decay analysis.

3.4 Particle Identification

Particle identification (PID) provides a information that we can distinguish the types of charged particles e, μ, π, K and p. In Belle experiment, the information from CDC, TOF, ACC ,ECL and KLM detectors of PID are calculated by likelihood. If we assume the particle track in Belle detector is an electron, muon, pion, kaon and proton respectively, the likelihoods are denoted by $\mathcal{L}_e$, $\mathcal{L}_\mu$, $\mathcal{L}_\pi$, $\mathcal{L}_K$ and $\mathcal{L}_p$ respectively. The definition of likelihood ratio $\mathcal{R}_{ij}$ for types of charged particle i and j is

$$\mathcal{R}_{ij} = \frac{\mathcal{L}_i}{\mathcal{L}_i + \mathcal{L}_j}. \tag{3.1}$$

The CDC, ACC and TOF get the likelihood for pion, kaon and proton. Moreover, ECL and KLM provide further information for electron and muon. We can observe the likelihood of track, and decide the track what type of charged particle like. Table 3.1 shows the PID in this thesis.

doi:10.6342/NTU201800725

	e	μ	π	K	p
$\mathcal{L}_e$	>0.9	none	<0.95	<0.95	<0.95
$\mathcal{L}_\mu$	none	>0.8	<0.95	<0.95	<0.95
$\mathcal{R}_{K\pi}$	none	none	<0.4	>0.6	none
$\mathcal{R}_{pK}$ and $\mathcal{R}_{p\pi}$	none	none	none	none	>0.6 and >0.6

Table 3.1: Summary of particle identification.

3.4.1 Electron Identification

Electron identification (eid) is able to identify the electron track from other hadron particle track. In Belle detector, there are two different ways for the eid information. First, electron can induce electromagnetic showers and pion induce hadron shower, the energy of electron is deposited in ECL. It has different energy deposition and shower shape. Second, in low momentum range, it is good to identify electron and hadron by dE/dx measurement in CDC detector.

3.4.2 Muon Identification

Muon identification (muid) is able to identify the muon track from other hadron particle track by the difference of interaction track in detector. Muon and hadron pass through the different numbers of KLM layers and they have different track trajectory.

3.5 Event Selection

This thesis will reconstruct the B meson candidates from the $B^0 \to l^+l^-l^+l^-$ four-body decay, there are four particles in final state. The signal events are reconstructed from two pairs of oppositely charged muons or electrons.

3.5.1 Charged Track Requirements

We use track informations of all charged particles, it includes e, μ, π, K, p that we do some constraints. The track deviations from IP must within ± 2.5 cm ($|dz| < 2.5$ cm) in z direction and ± 0.2 cm ($|dr| < 0.2$ cm) in transverse ($x - y$) plane. Besides, the transverse

doi:10.6342/NTU201800725

momentum ($p_T = \sqrt{p_x^2 + p_y^2}$) of charged particles must greater than 0.1 GeV/c ($p_T > 0.1$ GeV/c).

3.5.2 Multiple Candidates Selection

The signal candidates may have multiple candidates, the analysis code contains event id and χ^2, χ^2 represents the goodness of the vertex fitting [29].

$$\chi^2 = \sum_{i=1}^{N} \frac{[y_i - f(x_i, a, b)]^2}{\sigma_i^2} \tag{3.2}$$

Assuming we have N data (x_i, y_i) and known relation $y = f(x, a, b, ...)$, where y_i is a set of $\{y_1, y_2, ..., y_N\}$ with a set of uncertainty σ_i, and x_i is a set of measurement $\{x_1, x_2, ..., x_N\}$ that we know without uncertainty. Besides, parameters a, b,...are constants which we want to determine from our data. Equation (3.2) can find a, b that it minimizes the χ^2. If the event id same, it means multiple candidates. In order to remove multiple candidates, we compare the χ^2 in the same event id and we choose the candidate which has the minimum χ^2.

3.5.3 Bremsstrahlung Recovery

According to electrodynamics, when the electron accelerates due to the another charged particle, electron has electromagnetic radiation. Thus, electron loss some energy. This phenomenon is called the bremsstrahlung. A simple example is shown in Fig. 3.2, a high energy electron is accelerated by a atomic nucleus, the lost energy become photon leaving the electron.

For the $B^0 \to e^+e^-e^+e^-$ and $B^0 \to e^+e^-\mu^+\mu^-$ decay, we need to recover the lost energy for electrons. When electron radiates, the electromagnetic radiation is photon, which carry the energy leaving the electron. Therefore, if the angle between photon momentum and electron momentum less than 0.05 rad, we assume the photon from this electron. Fig. 3.3 shows the bremsstrahlung recovery for $B^0 \to e^+e^-e^+e^-$ and $B^0 \to e^+e^-\mu^+\mu^-$ decay in ΔE and M_{bc} distributions, we can see the signal Monte Carlo become concentrate, the

doi:10.6342/NTU201800725

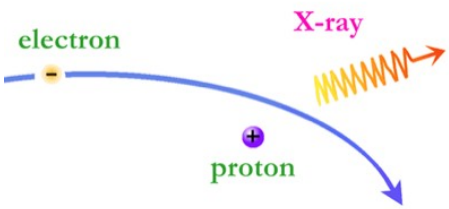
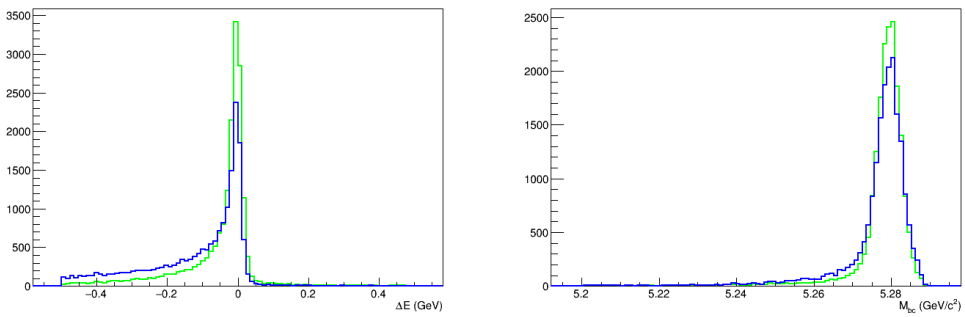
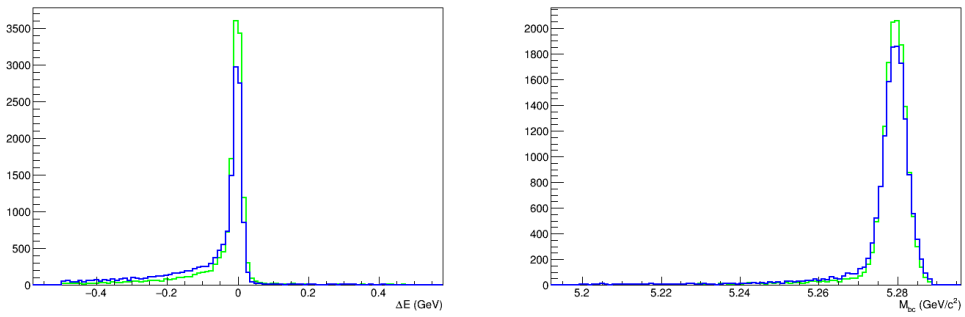


Figure 3.2: A simple example for bremsstrahlung [30].

tail of signal is less than before it do the bremsstrahlung recovery.



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



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

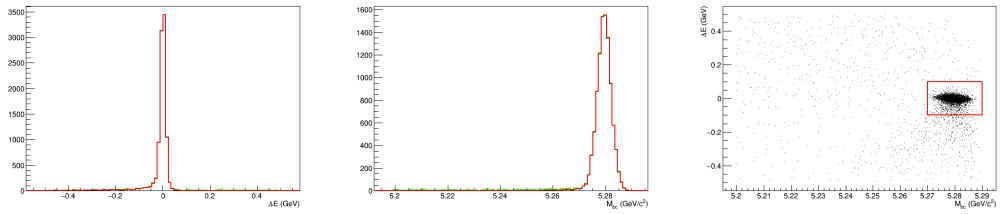
Figure 3.3: Bremsstrahlung recovery for $B^0 \to e^+e^-e^+e^-$ and $B^0 \to e^+e^-\mu^+\mu^-$ decay, the green lines represent after bremsstrahlung recovery and blue lines represent before bremsstrahlung recovery.

3.5.4 Signal Box

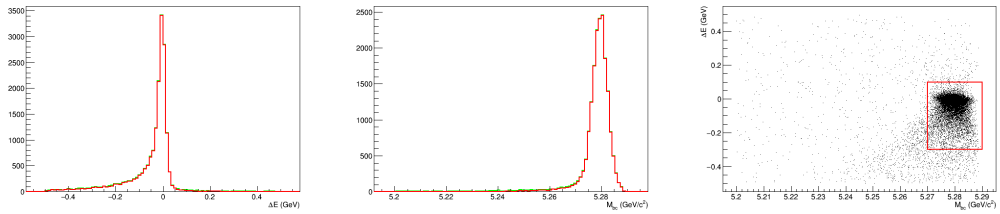
We reconstruct the B meson from 4-vector of all final state particles and do some selections. Furthermore, it often use two dynamic variables to identify B mesons: the energy difference $\Delta E = E_B - E_{beam}$ and the beam constrained mass $M_{bc} = \sqrt{E_{beam}^2 - |\vec{p}_B|^2}$,

doi:10.6342/NTU201800725

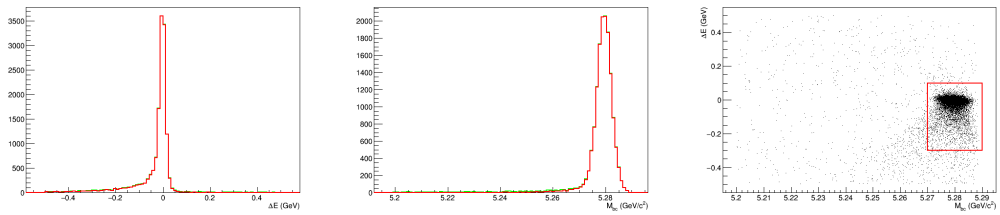
where E_B and $\vec{p}_B$ are the energy and momentum of B meson candidates in center-of-mass (CM) frame. This thesis requires $-0.1(-0.3) < \Delta E < 0.1$ GeV and $5.27 < M_{bc} < 5.29$ GeV/c^2 for $B^0 \to \mu^+\mu^-\mu^+\mu^-$ ($B^0 \to e^+e^-e^+e^-$ and $B^0 \to e^+e^-\mu^+\mu^-$). Fig. 3.4 shows ΔE and M_{bc} distributions for each mode.



(a) $B^0 \to \mu^+\mu^-\mu^+\mu^-$.



(b) $B^0 \to e^+e^-e^+e^-$.



(c) $B^0 \to e^+e^-\mu^+\mu^-$.

Figure 3.4: ΔE and M_{bc} distributions for $B^0 \to l^+l^-l^+l^-$, the red lines are true events for signal Monte Carlo in one dimension figures.

Finally, the summary of event selections in this section are shown in Table 3.2. ϵ_{sig} is signal efficiency:

$$\epsilon_{sig} = \frac{N_{remain}}{N_{generated}}. \tag{3.3}$$

Where N_{remain} is residual true events in signal MC after the cuts and $N_{generated}$ is 77100 in this thesis, true events are decided by Monte Carlo truth matching. We calculate ϵ_{sig} after we did the requirement for each step. In bremsstrahlung recovery step, ϵ_{sig} will not change.

doi:10.6342/NTU201800725

$B^0 \to \mu^+\mu^-\mu^+\mu^-$			
Step	Selection	Requirement	ϵ_{sig}
Step 1	Charged Track	$\lvert dr\rvert < 0.2$ cm $\lvert dz\rvert < 2.5$ cm $p_T > 0.1$ GeV/c	-†
Step 2	Multiple Candidates	compare event id and χ^2	12.82%
Step 3	Signal Box	$-0.1 < \Delta E < 0.1$ GeV $5.27 < M_{bc} < 5.29$ GeV/c^2	12.23%
$B^0 \to e^+e^-e^+e^-$			
Step	Selection	Requirement	ϵ_{sig}
Step 1	Charged Track	$\lvert dr\rvert < 0.2$ cm $\lvert dz\rvert < 2.5$ cm $p_T > 0.1$ GeV/c	-†
Step 2	Multiple Candidates	compare event id and χ^2	23.03%
Step 3	Bremsstrahlung Recovery	momentum angle < 0.05 rad	-
Step 4	Signal Box	$-0.3 < \Delta E < 0.1$ GeV $5.27 < M_{bc} < 5.29$ GeV/c^2	21.42%
$B^0 \to e^+e^-\mu^+\mu^-$			
Step	Selection	Requirement	ϵ_{sig}
Step 1	Charged Track	$\lvert dr\rvert < 0.2$ cm $\lvert dz\rvert < 2.5$ cm $p_T > 0.1$ GeV/c	-†
Step 2	Multiple Candidates	compare event id and χ^2	18.70%
Step 3	Bremsstrahlung Recovery	momentum angle < 0.05 rad	-
Step 4	Signal Box	$-0.3 < \Delta E < 0.1$ GeV $5.27 < M_{bc} < 5.29$ GeV/c^2	17.83%

Table 3.2: Summary of event selections, †: It includes multiple candidates.

3.6 Background Suppression

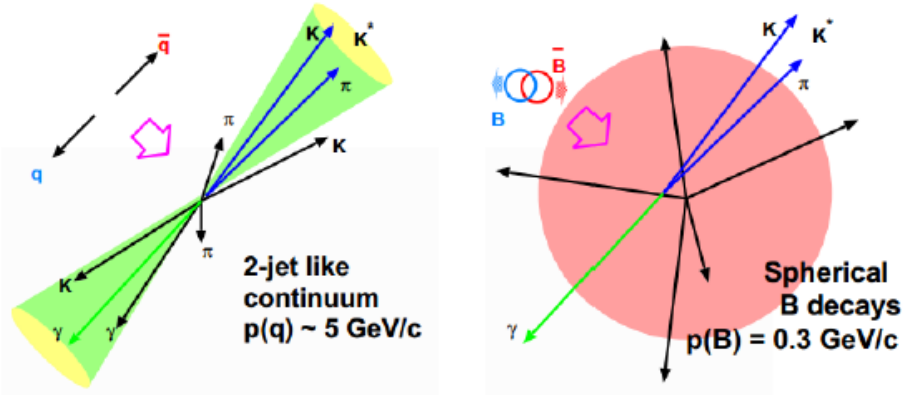
In this thesis, the analysis considers continuum background ($e^+e^- \to q\bar{q}$, where $q = u, d, s, c$ quarks) and generic B background ($b \to c$ transition, it includes mixed and charged decay).

In the following section we use Monte Carlo method to process the background, this thesis has 6 streams $q\bar{q}$ Monte Carlo and 6 streams generic B Monte Carlo for $B^0 \to \mu^+\mu^-\mu^+\mu^-$, $B^0 \to e^+e^-e^+e^-$ and $B^0 \to e^+e^-\mu^+\mu^-$. The analysis uses some variables for background suppression and it bases on algorithm and optimization.

doi:10.6342/NTU201800725

3.6.1 Continuum Background

Fig. 3.5 shows difference of event shape of $q\bar{q}$ pair event (continuum event) and B decay event. The shape of $q\bar{q}$ pair event is 2-jet-like and shape of B decay event is spherical-like, this property can help us to distinguish continuum event from B decay event. Therefore, some shape variables are used in this thesis and we introduce these variables.

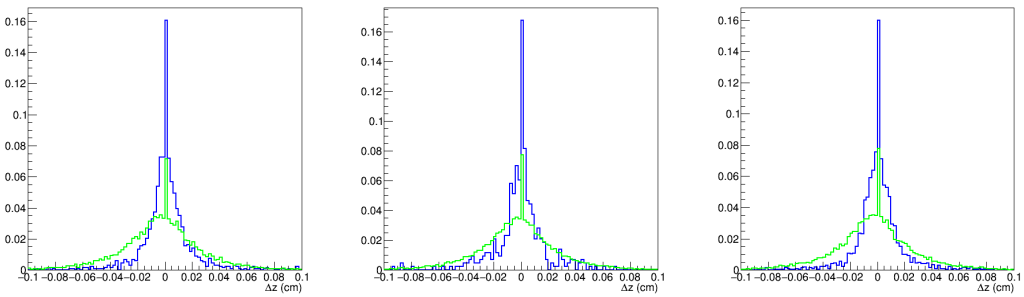


(a) $q\bar{q}$ pair event. (b) B decay event.

Figure 3.5: The difference of event shape.

- Δz

  The vertex difference between B candidate and the accompanying B is called Δz. The distribution of Δz of B decay events is broader than distribution of Δz of $q\bar{q}$ events due to color confinement in QCD. The distributions of Δz for $B^0 \to l^+l^-l^+l^-$ are shown in Fig. 3.6. There is a peak at $\Delta z = 0$ because we can't confirm Δz of the other side reconstructive B meson.



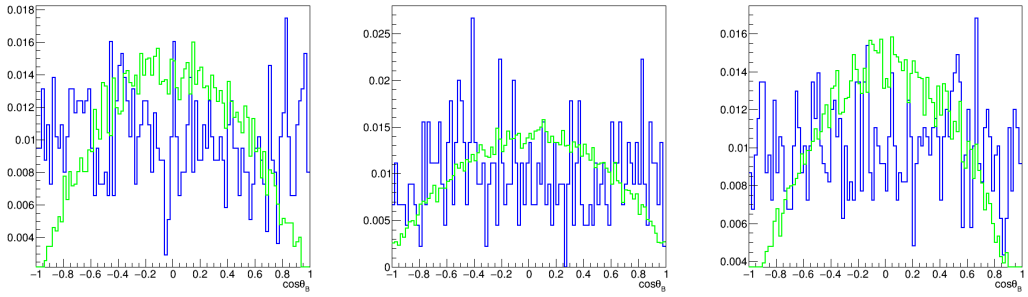
(a) $B^0 \to \mu^+\mu^-\mu^+\mu^-$. (b) $B^0 \to e^+e^-e^+e^-$. (c) $B^0 \to e^+e^-\mu^+\mu^-$.

Figure 3.6: The distributions of Δz for $B^0 \to l^+l^-l^+l^-$. The green lines are signal MC and blue lines are $q\bar{q}$ MC. The distributions had normalized.

doi:10.6342/NTU201800725

- $cos\theta_B$

  The angle between the beam direction and B flight direction in the $\Upsilon(4S)$ rest frame is called the θ_B. The distribution of $cos\theta_B$ conform to $1 - cos^2\theta_B$ for B decay events and uniform distribution for $q\bar{q}$ events. The distributions of $cos\theta_B$ for $B^0 \to l^+l^-l^+l^-$ are shown in Fig. 3.7.



(a) $B^0 \to \mu^+\mu^-\mu^+\mu^-$. (b) $B^0 \to e^+e^-e^+e^-$. (c) $B^0 \to e^+e^-\mu^+\mu^-$.

Figure 3.7: The distributions of $cos\theta_B$ for $B^0 \to l^+l^-l^+l^-$. The green lines are signal MC and blue lines are $q\bar{q}$ MC. The distributions had normalized.

- $cos\theta_T$

  The angle between the thrust axis ($\vec{n}$) of the B candidate and the remaining particles is called the thrust angle θ_T. The thrust angle ($\vec{n}$) is defined by the direction which maximize the $T(\vec{n})$:

$$T(\vec{n}) = \frac{\sum_{i=1}^{N} |\vec{P}_i \cdot \vec{n}|}{\sum_{i=1}^{N} |\vec{P}_i|}, \tag{3.4}$$

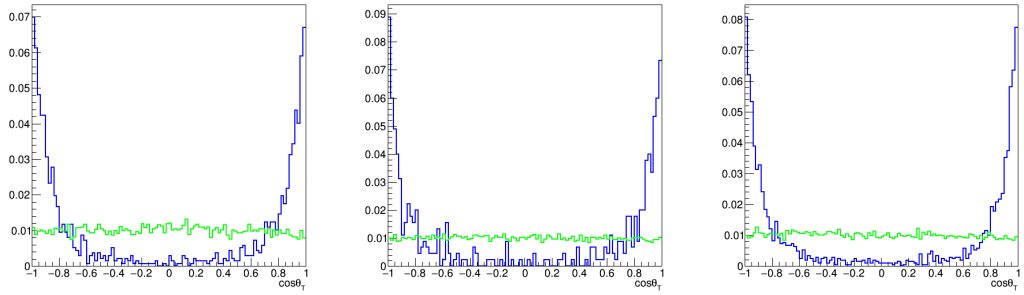
  where $\vec{P}_i$ is three-momentum of i-th daughter particle of B candidates and N is the number of daughter particles which are reconstructed to B candidates. The distributions of $cos\theta_T$ for $B^0 \to l^+l^-l^+l^-$ are shown in Fig. 3.8.

- Sphericity

  The ratio of total magnitude of transverse momentum to the total magnitude of momentum is defined as sphericity ($S_\perp$):

$$S_\perp = \frac{\sum |\vec{P}_t|}{\sum |\vec{P}|}, \tag{3.5}$$

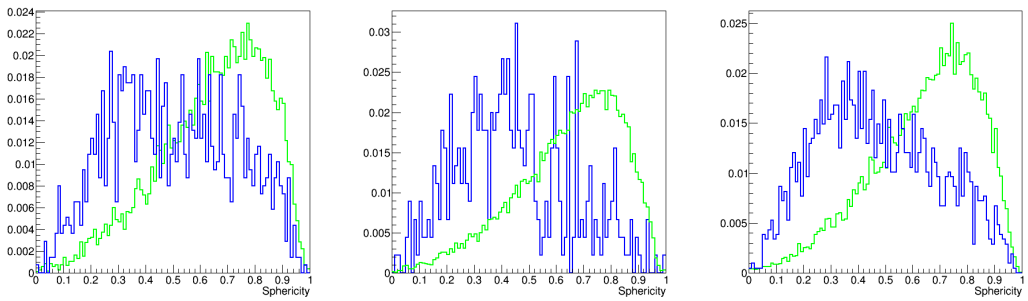
doi:10.6342/NTU201800725



(a) $B^0 \to \mu^+\mu^-\mu^+\mu^-$. (b) $B^0 \to e^+e^-e^+e^-$. (c) $B^0 \to e^+e^-\mu^+\mu^-$.

Figure 3.8: The distributions of $cos\theta_T$ for $B^0 \to l^+l^-l^+l^-$. The green lines are signal MC and blue lines are $q\bar{q}$ MC. The distributions had normalized.

where $\vec{P}_t$ is transverse momentum refer to thrust axis. Most particles of $q\bar{q}$ events fly along the thrust axis (Fig. 3.5(a)). Thus, transverse momentum of these particles is so small that $S_\perp$ near to 0. The distributions of sphericity for $B^0 \to l^+l^-l^+l^-$ are shown in Fig. 3.9.



(a) $B^0 \to \mu^+\mu^-\mu^+\mu^-$. (b) $B^0 \to e^+e^-e^+e^-$. (c) $B^0 \to e^+e^-\mu^+\mu^-$.

Figure 3.9: The distributions of sphericity for $B^0 \to l^+l^-l^+l^-$. The green lines are signal MC and blue lines are $q\bar{q}$ MC. The distributions had normalized.

- Kakuno Super Fox-Wolfram (KSFW)

  KSFW is another a set of variables for separating signal from $q\bar{q}$ events. It is defined as:

$$KSFW \equiv \sum_{l=0}^{4} R_l^{so} + \sum_{l=0}^{4} R_l^{oo} + \gamma \sum_{l=0}^{N_t} |P_t|, \tag{3.6}$$

  The superscript s means hadronic particles from reconstructed B meson, from other particles denotes o. P_t is transverse momentum, N_t is the number of tracks in a event, γ is Fisher coefficient.

  The mathematical form of R_l^{so} and R_l^{oo} are quite complicated. It provides many

doi:10.6342/NTU201800725

variables to separate the continuum background. In sum, these variables relate to the event shape.

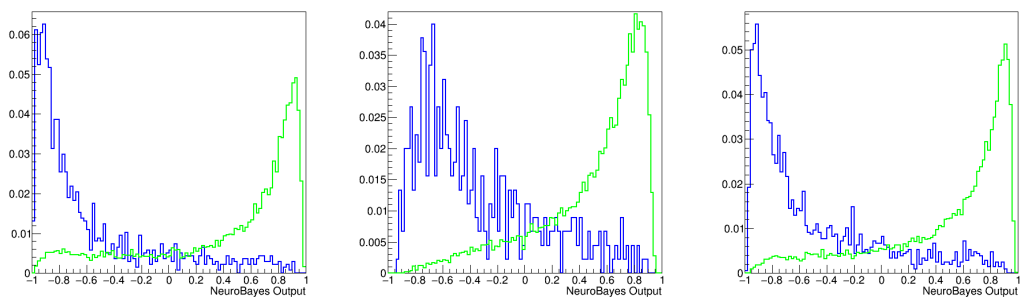
3.6.2 NeuroBayes

The NeuroBayes algorithm is a convenient tool for multivariate analysis. It includes neural network and bayesian statistics. Thus, it can yield a well performing algorithm. In the beginning we input variables what we need to training. After training, it output a variable. Final, we expert the variable and it is called NeuroBayes output. Using NeuroBayes output can help us to separate signal MC from background MC.

In this section we use NeuroBayes to process continuum background, so we input signal MC and $q\bar{q}$ MC for training. The used variables for training are list in Table 3.3. and the NeuroBayes outputs are shown in Fig. 3.10.

Variable
Δz
$cos\theta_B$
$cos\theta_T$
sphericity
KSFW variables

Table 3.3: Used variables for training.



(a) $B^0 \to \mu^+\mu^-\mu^+\mu^-$. (b) $B^0 \to e^+e^-e^+e^-$. (c) $B^0 \to e^+e^-\mu^+\mu^-$.

Figure 3.10: The NeuroBayes outputs for $B^0 \to l^+l^-l^+l^-$. The green lines are signal MC and blue lines are $q\bar{q}$ MC. The distributions had normalized.

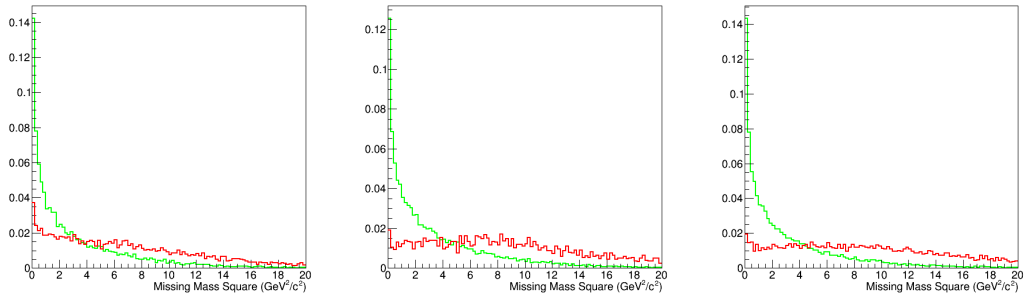
doi:10.6342/NTU201800725

3.6.3 Generic B Background

Comparing the number of background, the generic B background is main background. In order to separate signal MC from generic B MC, we choose missing mass square due to the neutrino. When we reconstruct the B candidate, it is impossible to catch the neutrino. If the background from some decay about lepton decay, the background includes the neutrino due to conservation of lepton number. The definition of missing mass square:

$$MM^2 = (\vec{P}_{beam} - \sum_n \vec{P}_n) \cdot (\vec{P}_{beam} - \sum_n \vec{P}_n), \tag{3.7}$$

where $\vec{P}_{beam}$ is a 4-vector and $\vec{P}_n$ also is a 4-vector for e, μ, π, K, p and photon. The distributions of missing mass square for $B^0 \to l^+l^-l^+l^-$ are shown in Fig. 3.11.



(a) $B^0 \to \mu^+\mu^-\mu^+\mu^-$. (b) $B^0 \to e^+e^-e^+e^-$. (c) $B^0 \to e^+e^-\mu^+\mu^-$.

Figure 3.11: The distributions of missing mass square for $B^0 \to l^+l^-l^+l^-$. The green lines are signal MC and red lines are generic B MC. The distributions had normalized.

3.6.4 Figure of Merit

To separate signal MC from background MC, we must find a good cut for NeuroBayes output and missing mass square. The Figure of Merit ($\mathcal{F}.\mathcal{O}.\mathcal{M}.$) can optimize the cut and maximize the signal-to-noise ratio. We calculate $\mathcal{F}.\mathcal{O}.\mathcal{M}.$ in signal box. There are many formulas for $\mathcal{F}.\mathcal{O}.\mathcal{M}.$, we use the formula in this thesis due to few signals [31]:

$$\mathcal{F}.\mathcal{O}.\mathcal{M}. = \frac{\epsilon_{sig}}{\frac{a}{2} + \sqrt{N_{bkg}}}, \tag{3.8}$$

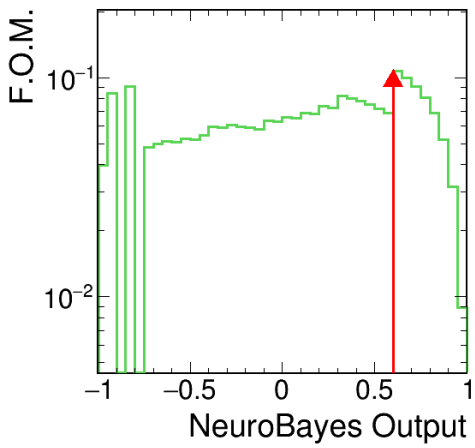
doi:10.6342/NTU201800725

where N_{bkg} is the number of background, a chooses 1.28 corresponding to an one-side Gaussian. The results of $\mathcal{F}.\mathcal{O}.\mathcal{M}.$ are shown in Fig. 3.12. and listed in Table 3.4.

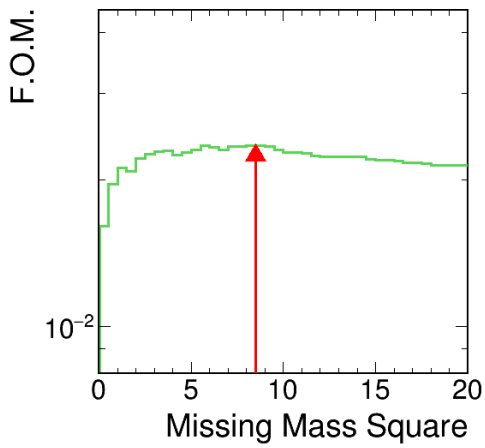
$B^0 \to \mu^+\mu^-\mu^+\mu^-$				
Variable	$\mathcal{F}.\mathcal{O}.\mathcal{M}.$	ϵ_{sig}	$N_{genericB}$	$N_{continuum}$
NeuroBayes output	>0.6	6.27%	1616	48
missing mass square	<8.5			
$B^0 \to e^+e^-e^+e^-$				
Variable	$\mathcal{F}.\mathcal{O}.\mathcal{M}.$	ϵ_{sig}	$N_{genericB}$	$N_{continuum}$
NeuroBayes output	>0.7	4.75%	129	3
missing mass square	<2.0			
$B^0 \to e^+e^-\mu^+\mu^-$				
Variable	$\mathcal{F}.\mathcal{O}.\mathcal{M}.$	ϵ_{sig}	$N_{genericB}$	$N_{continuum}$
NeuroBayes output	>0.8	4.80%	665	16
missing mass square	<4.0			

Table 3.4: Selections of figure of merit. The ϵ_{sig} are counted in signal box and number of background are counted in all region ($-0.5 < \Delta E < -0.5$ and $5.2 < M_{bc} < 5.29$) and number of background are 6 stream.

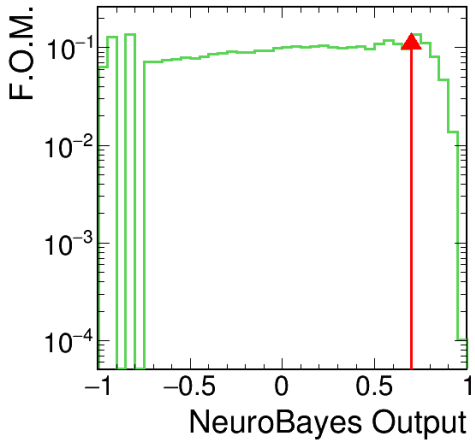
doi:10.6342/NTU201800725



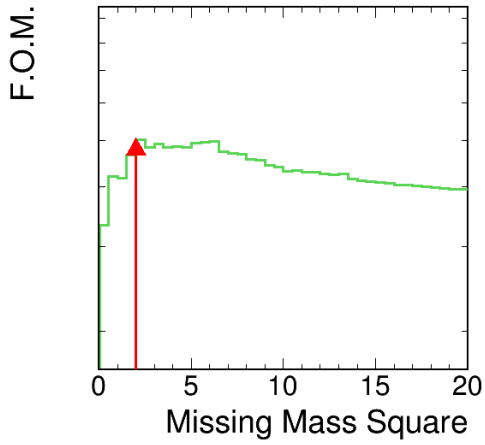
(a) NeuroBayes output for $B^0 \to \mu^+\mu^-\mu^+\mu^-$.



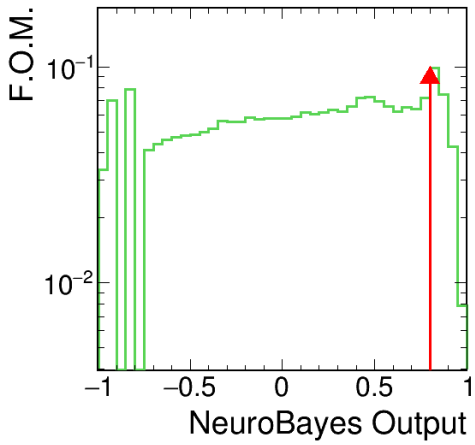
(b) Missing mass square for $B^0 \to \mu^+\mu^-\mu^+\mu^-$.



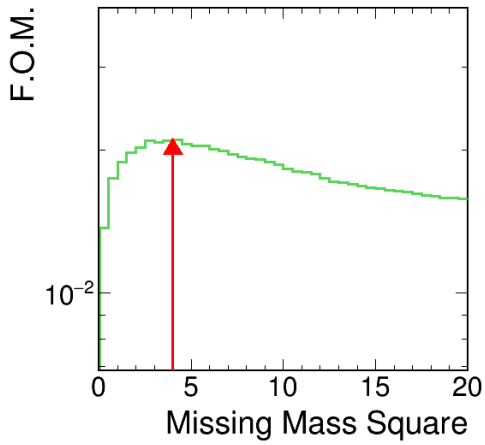
(c) NeuroBayes output for $B^0 \to e^+e^-e^+e^-$.



(d) Missing mass square for $B^0 \to e^+e^-e^+e^-$.



(e) NeuroBayes output for $B^0 \to e^+e^-\mu^+\mu^-$.



(f) Missing mass square for $B^0 \to e^+e^-\mu^+\mu^-$.

Figure 3.12: The $\mathcal{F}.\mathcal{O}.\mathcal{M}.$ results for NeuroBayes output and Missing mass square for each modes. We choose the maximal value of $\mathcal{F}.\mathcal{O}.\mathcal{M}.$ as the cuts.

doi:10.6342/NTU201800725

3.7 Signal Extraction

3.7.1 Introduction

To do the signal extraction, we use a two dimensional unbinned extended likelihood fit, which maximizes the likelihood function:

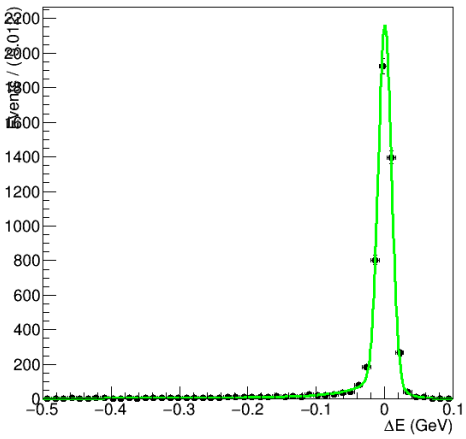
$$\mathcal{L} = \frac{e^{-N}}{N!} \prod_{i=1}^{N} [N_{sig} P^i_{sig}(M_{bc}, \Delta E) + N_{bkg} P^i_{bkg}(M_{bc}, \Delta E)], \tag{3.9}$$

where i means the i-th event, N, N_{sig} and N_{bkg} are the number of total events, signal events and background events respectively. P^i_{sig} (P^i_{bkg}) denotes the signal (background) probability density function for the i-th event with the two dimensional variables M_{bc} and ΔE. The numerical analysis is worked by RooFit.

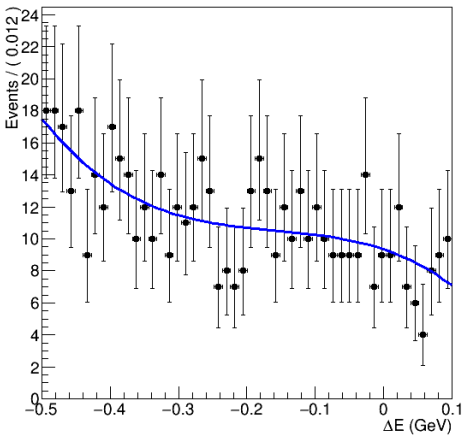
3.7.2 Modelling for Probability Density Function

In this thesis, we produce a product of ΔE and M_{bc} distribution for two dimensional fitting, fitting region are $-0.5 < \Delta E < 0.1$ GeV and $5.24 < M_{bc} < 5.29$ GeV/c^2. The two dimensional PDF can project on ΔE and M_{bc} components, and we use some functions to model the two components. A Gaussian function and a Crystal Ball function model the signal ΔE, a Crystal Ball function models the signal M_{bc} for each modes, a Chebyshev polynomial models the background ΔE and a Argus function models the background M_{bc} for each modes. The generic B background is much larger than continuum background. Thus, we combine two backgrounds for fitting. Fig. 3.13, 3.14, 3.15 and Table 3.5, 3.6 show the fitting results for each modes.

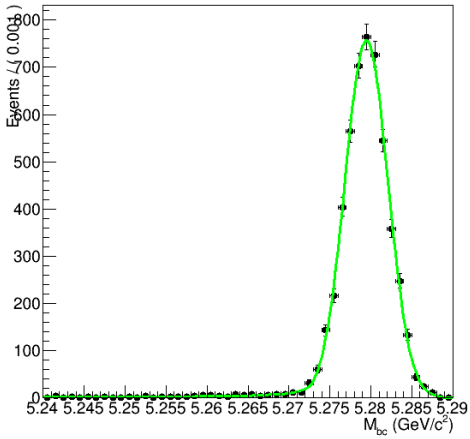
doi:10.6342/NTU201800725



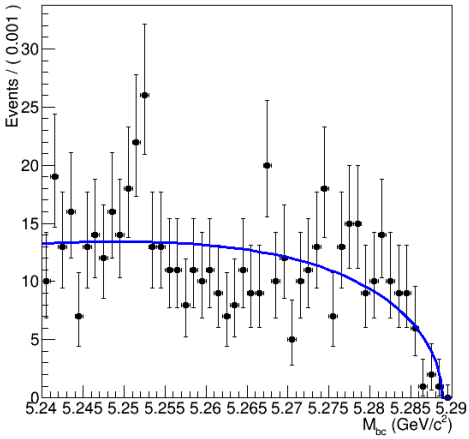
(a) Signal ΔE.



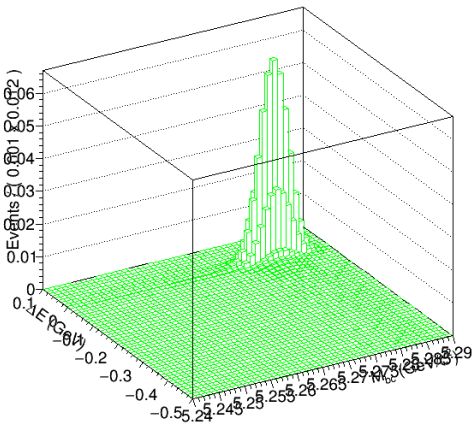
(b) Background ΔE.



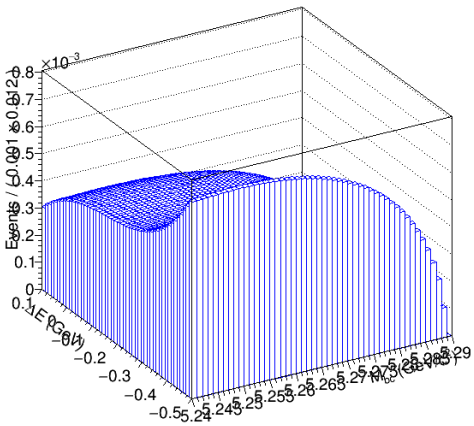
(c) Signal M_{bc}.



(d) Background M_{bc}.



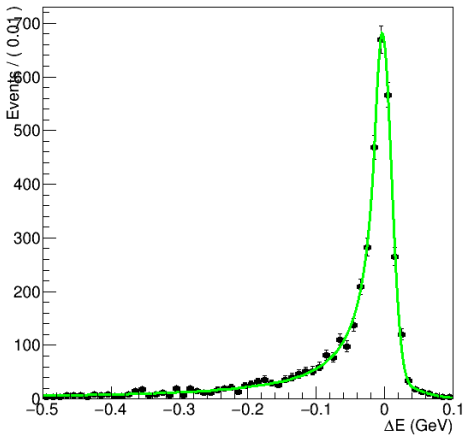
(e) 2 dimensional signal PDF.



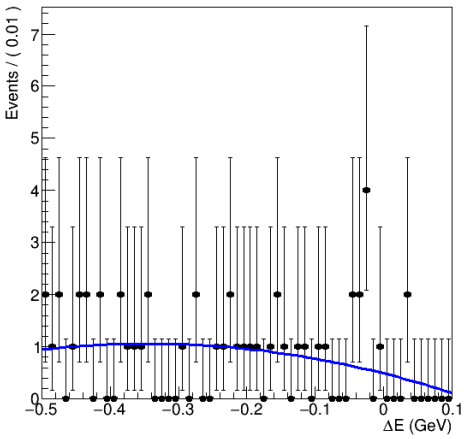
(f) 2 dimensional background PDF.

Figure 3.13: Fitting results for $B^0 \to \mu^+\mu^-\mu^+\mu^-$.

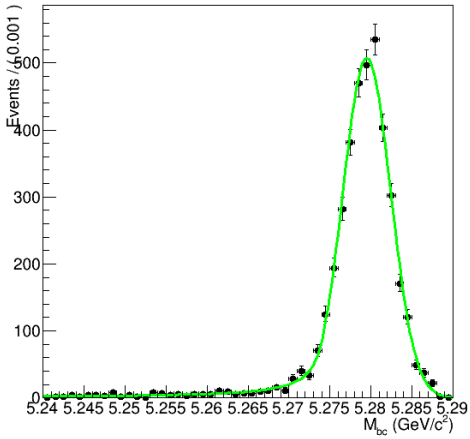
doi:10.6342/NTU201800725



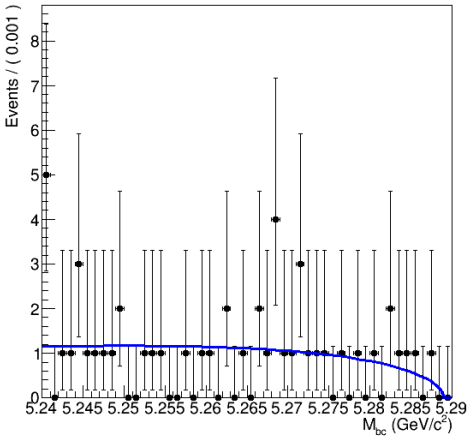
(a) Signal ΔE.



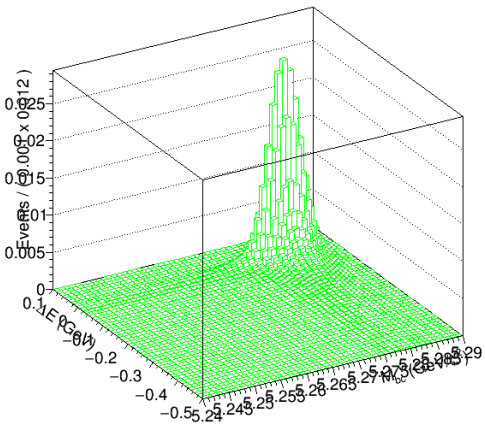
(b) Background ΔE.



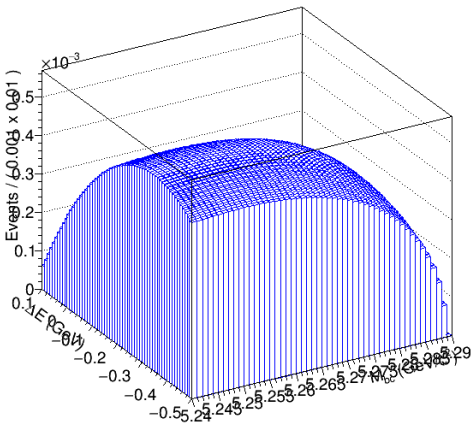
(c) Signal M_{bc}.



(d) Background M_{bc}.



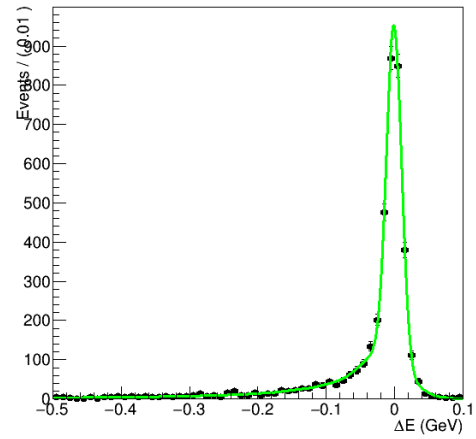
(e) 2 dimensional signal PDF.



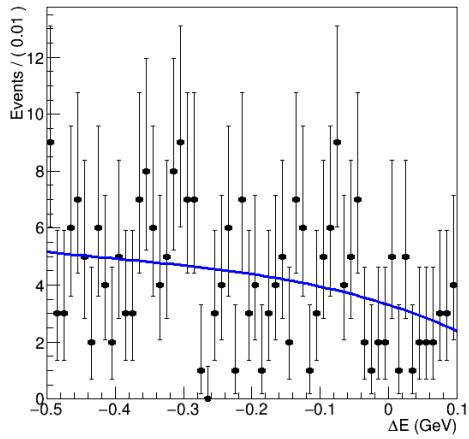
(f) 2 dimensional background PDF.

Figure 3.14: Fitting results for $B^0 \to e^+e^-e^+e^-$.

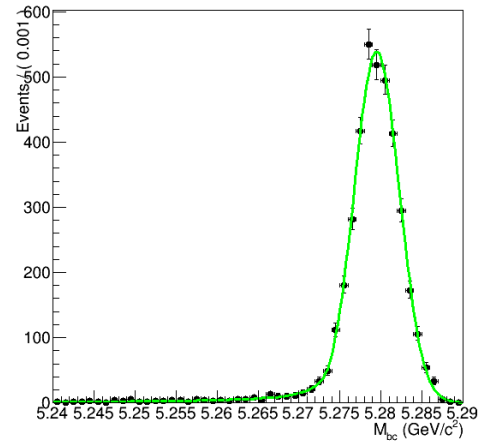
doi:10.6342/NTU201800725



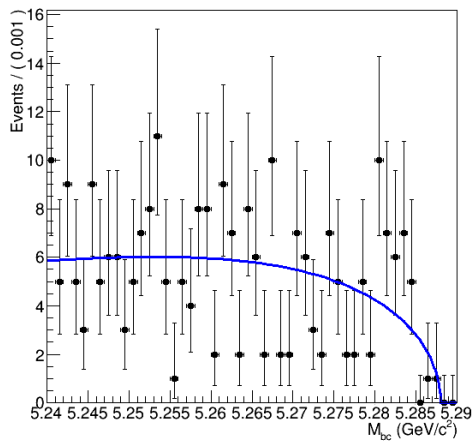
(a) Signal ΔE.



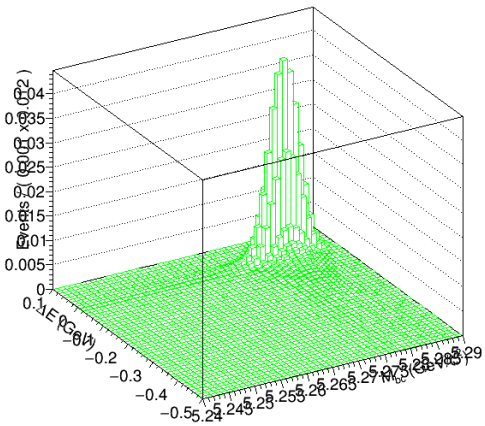
(b) Background ΔE.



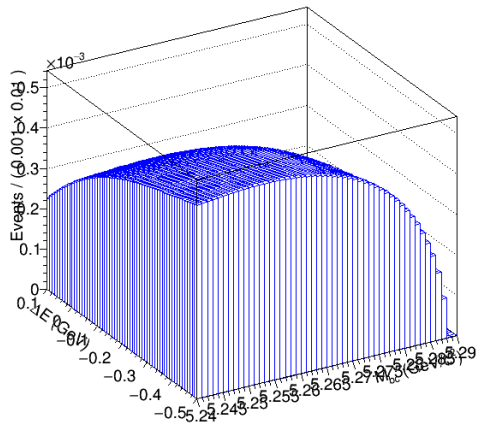
(c) Signal M_{bc}.



(d) Background M_{bc}.



(e) 2 dimensional signal PDF.



(f) 2 dimensional background PDF.

Figure 3.15: Fitting results for $B^0 \to e^+e^-\mu^+\mu^-$.

doi:10.6342/NTU201800725

ΔE for $B^0 \to \mu^+\mu^-\mu^+\mu^-$			
Function	Parameter	Value	Error
Gaussian	μ	5.7342×10^{-4}	1.71×10^{-4}
	σ	9.6711×10^{-3}	1.71×10^{-4}
Crystal Ball	μ	-1.4911×10^{-2}	2.96×10^{-3}
	σ	3.282×10^{-2}	2.37×10^{-3}
	α	9.3524×10^{-1}	1.23×10^{-1}
	n	1.2809	2.22×10^{-1}
Ratio $f^\dagger$		8.2343×10^{-1}	1.24×10^{-2}
M_{bc} for $B^0 \to \mu^+\mu^-\mu^+\mu^-$			
Function	Parameter	Value	Error
Crystal Ball	μ	5.2795	3.85×10^{-5}
	σ	2.6282×10^{-3}	2.93×10^{-5}
	α	2.4913	1.09×10^{-1}
	n	8.5434×10^{-1}	1.65×10^{-1}
ΔE for $B^0 \to e^+e^-e^+e^-$			
Function	Parameter	Value	Error
Gaussian	μ	2.4918×10^{-2}	1.07×10^{-2}
	σ	3.2236×10^{-2}	5.09×10^{-3}
Crystal Ball	μ	-3.1956×10^{-3}	5.45×10^{-4}
	σ	1.3324×10^{-2}	5.68×10^{-4}
	α	7.7762×10^{-1}	5.02×10^{-2}
	n	1.452	7.38×10^{-2}
Ratio $f^\dagger$		4.4505×10^{-2}	1.64×10^{-2}
M_{bc} for $B^0 \to e^+e^-e^+e^-$			
Function	Parameter	Value	Error
Crystal Ball	μ	5.2795	5.27×10^{-5}
	σ	2.912×10^{-3}	4.2×10^{-5}
	α	1.9216	8.63×10^{-2}
	n	1.2109	1.58×10^{-1}
ΔE for $B^0 \to e^+e^-\mu^+\mu^-$			
Function	Parameter	Value	Error
Gaussian	μ	-6.1077×10^{-4}	3.07×10^{-4}
	σ	1.1064×10^{-2}	3.58×10^{-4}
Crystal Ball	μ	-2.2177×10^{-2}	3.72×10^{-3}
	σ	3.567×10^{-2}	2.49×10^{-3}
	α	7.4946×10^{-1}	8.31×10^{-2}
	n	1.7967	2.89×10^{-1}
Ratio $f^\dagger$		6.1241×10^{-1}	2.17×10^{-2}
M_{bc} for $B^0 \to e^+e^-\mu^+\mu^-$			
Function	Parameter	Value	Error
Crystal Ball	μ	5.2795	4.78×10^{-5}
	σ	2.7547×10^{-3}	3.71×10^{-5}
	α	2.0672	8.77×10^{-2}
	n	1.2033	1.63×10^{-1}

Table 3.5: Modelling for signal Monte Carlo, †: $M = fF^{Gaussian} + (1-f)F^{CB}$.

doi:10.6342/NTU201800725

ΔE for $B^0 \to \mu^+\mu^-\mu^+\mu^-$			
Function	Parameter	Value	Error
Chebyshev	c_0	-3.6968×10^{-1}	7.91×10^{-2}
	c_1	6.9751×10^{-2}	6.76×10^{-2}
	c_2	-8.2878×10^{-2}	6.53×10^{-2}
M_{bc} for $B^0 \to \mu^+\mu^-\mu^+\mu^-$			
Function	Parameter	Value	Error
Argus	c	-3.3158×10	8.67
	m_0	5.2888	1.75×10^{-4}
ΔE for $B^0 \to e^+e^-e^+e^-$			
Function	Parameter	Value	Error
Chebyshev	c_0	-5.6099×10^{-1}	2.68×10^{-1}
	c_1	-2.8959×10^{-1}	3.25×10^{-1}
M_{bc} for $B^0 \to e^+e^-e^+e^-$			
Function	Parameter	Value	Error
Argus	c	-3.3638×10	2.99×10
	m_0	5.289	1.44×10^{-3}
ΔE for $B^0 \to e^+e^-\mu^+\mu^-$			
Function	Parameter	Value	Error
Chebyshev	c_0	-3.211×10^{-1}	1.16×10^{-1}
	c_1	-7.435×10^{-2}	1.08×10^{-1}
	c_2	-1.9084×10^{-2}	9.61×10^{-2}
M_{bc} for $B^0 \to e^+e^-\mu^+\mu^-$			
Function	Parameter	Value	Error
Argus	c	-3.7503×10	1.34×10
	m_0	5.288	4.48×10^{-4}

Table 3.6: Modelling for background Monte Carlo.

doi:10.6342/NTU201800725

Chapter 4

Control Sample Study

4.1 Introduction

In this thesis, we do the control sample for checking whether the analysis is correct or not. The method is that we calculate the calibration factor and branching fraction. We use the ΔE and M_{bc} distributions to find the difference between Monte Carol and Belle collected data so that we can get calibration factor. For control sample, this thesis focus on $B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$ and $B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$, according to Particle Data Group [32], branching fraction of $B^0 \to J/\psi(J/\psi \to l^+l^-)K^+\pi^-$ is the product of branching fraction of $B^0 \to J/\psi K^+\pi^-$ and $J/\psi \to l^+l^-$:

$$\begin{aligned}\mathcal{B}(B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-) &= \mathcal{B}(B^0 \to J/\psi K^+\pi^-) \times \mathcal{B}(J/\psi \to \mu^+\mu^-) \\ &= (1.15 \pm 0.05) \times 10^{-3} \times (5.961 \pm 0.033) \times 10^{-2} \\ &= (6.855 \pm 0.300) \times 10^{-5},\end{aligned} \tag{4.1}$$

$$\begin{aligned}\mathcal{B}(B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-) &= \mathcal{B}(B^0 \to J/\psi K^+\pi^-) \times \mathcal{B}(J/\psi \to e^+e^-) \\ &= (1.15 \pm 0.05) \times 10^{-3} \times (5.971 \pm 0.032) \times 10^{-2} \\ &= (6.867 \pm 0.301) \times 10^{-5}.\end{aligned} \tag{4.2}$$

All the Monte Carol and data run 1 stream in this chapter.

doi:10.6342/NTU201800725

4.2 Particle Identification

The PID of control sample is listed in Table 4.1.

	e	μ	π	K	p
$\mathcal{L}_e$	>0.9	none	<0.95	<0.95	<0.95
$\mathcal{L}_\mu$	none	>0.8	<0.95	<0.95	<0.95
$\mathcal{R}_{K\pi}$	none	none	<0.4	>0.6	none
$\mathcal{R}_{pK}$ and $\mathcal{R}_{p\pi}$	none	none	none	none	>0.6 and >0.6

Table 4.1: Summary of particle identification for control sample.

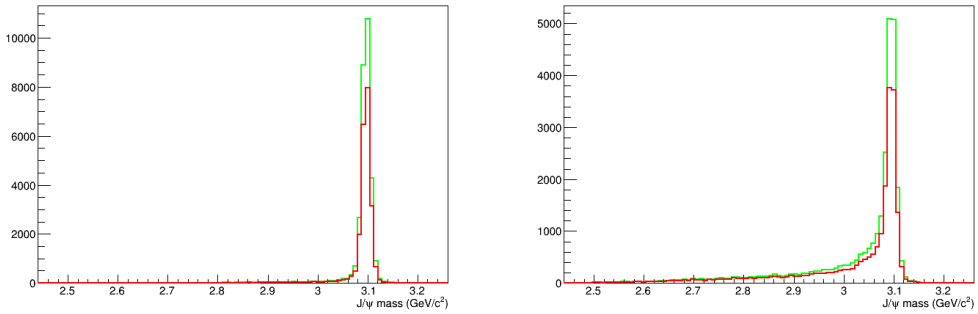
4.3 Event Selection

Table 4.2 lists event selections for control sample and Fig. 4.1 shows J/ψ mass distributions before mass requirement step.

$B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$			
Step	Selection	Requirement	ϵ_{sig}
Step 1	Charged Track	$\lvert dr\rvert < 0.2$ cm $\lvert dz\rvert < 2.5$ cm	-†
Step 2	Multiple Candidates	compare event id and χ^2	29.42%
Step 3	Mass Requirement	$3.05 < M(J/\psi) < 3.15$ GeV/c^2 [33]	27.73%
Step 4	Signal Box	$-0.1 < \Delta E < 0.1$ GeV $5.27 < M_{bc} < 5.29$ GeV/c^2	27.41%
$B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$			
Step	Selection	Requirement	ϵ_{sig}
Step 1	Charged Track	$\lvert dr\rvert < 0.2$ cm $\lvert dz\rvert < 2.5$ cm	-†
Step 2	Multiple Candidates	compare event id and χ^2	27.34%
Step 3	Bremsstrahlung Recovery	momentum angle < 0.05 rad	-
Step 4	Mass Requirement	$2.95 < M(J/\psi) < 3.15$ GeV/c^2 [33]	22.11%
Step 5	Signal Box	$-0.2 < \Delta E < 0.1$ GeV $5.27 < M_{bc} < 5.29$ GeV/c^2	21.79%

Table 4.2: Summary of event selections for control sample. †: It includes multiple candidates.

doi:10.6342/NTU201800725



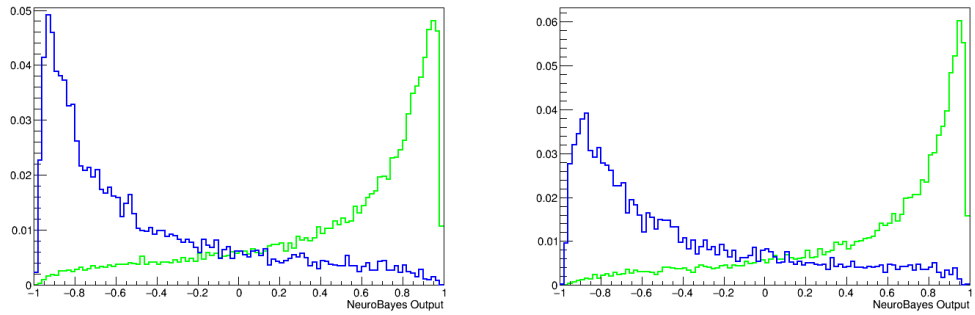
(a) $B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$. (b) $B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$.

Figure 4.1: J/ψ mass distributions, the red lines are true events for signal Monte Carlo of control sample modes.

4.4 Background Suppression

4.4.1 Continuum Background

We use same variables, which are listed in Table 3.3 to separate continuum background in control sample modes. The NeuroBayes outputs are shown in Fig. 4.2.



(a) $B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$. (b) $B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$.

Figure 4.2: The NeuroBayes outputs for control sample. The green lines are signal MC and blue lines are $q\bar{q}$ MC. The distributions had normalized.

The branching fraction of control sample is quite large, we consider the proper formula of $\mathcal{F.O.M.}$:

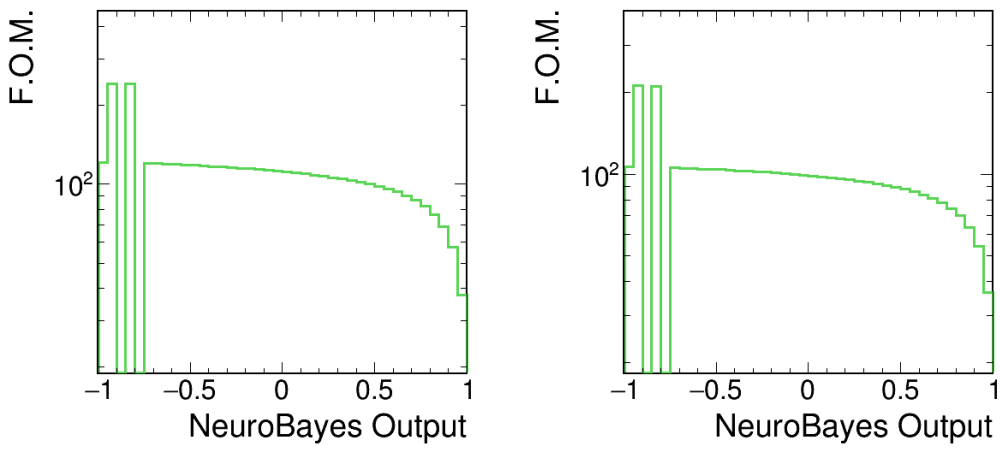
$$\mathcal{F.O.M.} = \frac{N_{sig}}{\sqrt{N_{sig} + N_{bkg}}}, \tag{4.3}$$

where N_{sig} is calculated by

$$N_{sig} = \epsilon_{sig} \times 7.71 \times 10^8 \times \mathcal{B}. \tag{4.4}$$

doi:10.6342/NTU201800725

Nevertheless, as shown in Fig. 4.3, we calculate the formula in signal box that we can't find the peak. Thus, for NeuroBayes output cuts of control sample, we choose 0 instinctively. We get values of signal efficiency after we cut the NeuroBayes output and we list it in Table 4.3.



(a) $B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$. (b) $B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$.

Figure 4.3: The $\mathcal{F}.\mathcal{O}.\mathcal{M}.$ of NeuroBayes outputs for control sample.

Mode	cut	ϵ_{sig}	$N_{genericB}$	$N_{continuum}$
$B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$	>0	22.91%	98753	1125
$B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$	>0	18.64%	83396	902

Table 4.3: NeuroBayes output cuts for control sample. The ϵ_{sig} are counted in signal box and number of background are counted in all region ($-0.5 < \Delta E < -0.5$ and $5.2 < M_{bc} < 5.29$).

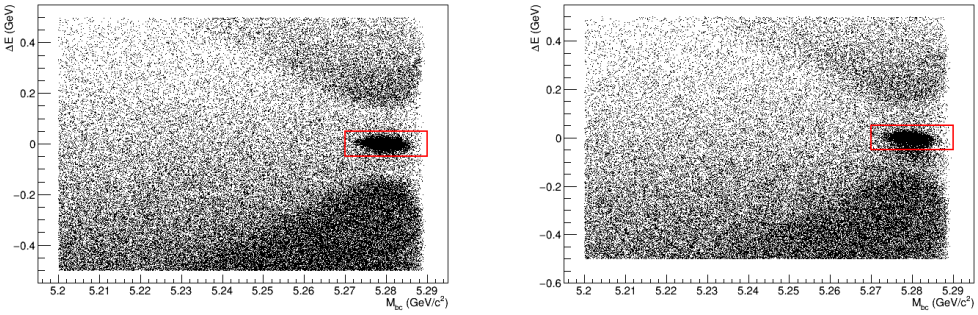
4.4.2 Generic B Background

As shown in Fig. 4.4 (a) and (b), there is a peak in signal box region of generic B background due to the large branching fraction of control sample modes. Thus, there are some signal event modes in this peak to be listed in Table 4.4. In order to eliminate peak, we use Monte Carlo truth matching to delete these modes, the results are shown in Fig. 4.4 (c) and (d), $N_{genericB} = 74492$ in Fig. 4.4 (c) and $N_{genericB} = 63791$ in Fig. 4.4 (d).

doi:10.6342/NTU201800725

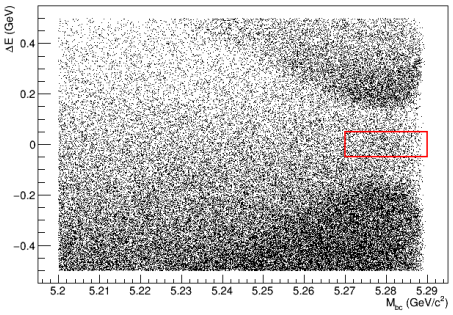
$B^0 \to J/\psi(J/\psi \to l^+l^-)K^+\pi^-$
$B^0 \to J/\psi(J/\psi \to l^+l^-)K^{*0}(K^{*0} \to K^+\pi^-)$
$B^0 \to J/\psi(J/\psi \to l^+l^-)K_2^{*0}(K_2^{*0} \to K^+\pi^-)$
$B^0 \to J/\psi(J/\psi \to l^+l^-)K_0^{*0}(K_0^{*0} \to K^+\pi^-)$
$B^0 \to J/\psi(J/\psi \to l^+l^-)(30343)((30343) \to K^+\pi^-)$

Table 4.4: Signal event modes in peak.

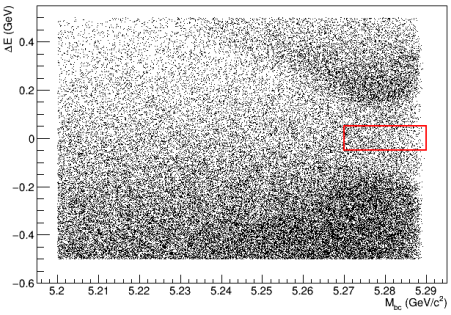


(a) $B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$.

(b) $B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$.



(c) $B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$.



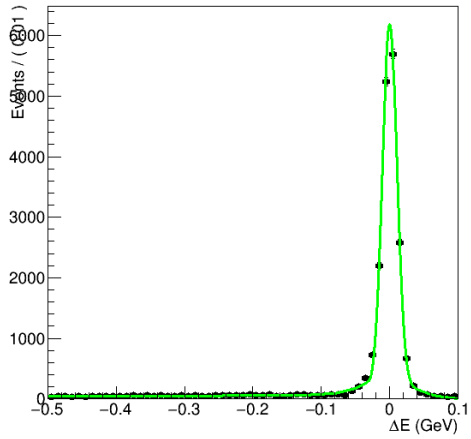
(d) $B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$.

Figure 4.4: Peak in generic B background.

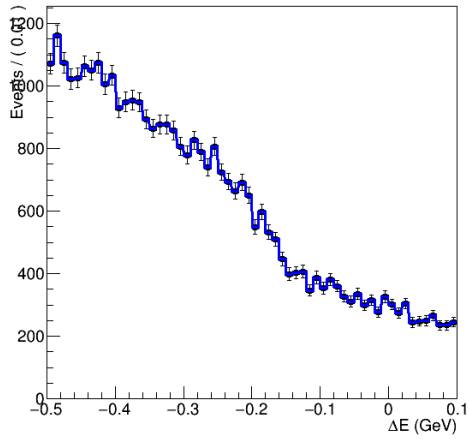
4.5 Signal Extraction

We use two dimensional fitting to fit signal MC and background MC, fitting region are $-0.5 < \Delta E < 0.1$ GeV and $5.24 < M_{bc} < 5.29$ GeV/c^2. We use a Gaussian function and a Crystal Ball function to model the signal ΔE, a Crystal Ball function model the signal M_{bc} for each control sample modes. For background, the generic B background is much larger than continuum background. Thus, we combine two backgrounds for fitting and we use a two dimensional histogram to model background for each control sample modes. The fitting results are shown in Fig. 4.5, 4.6 and Table 4.5.

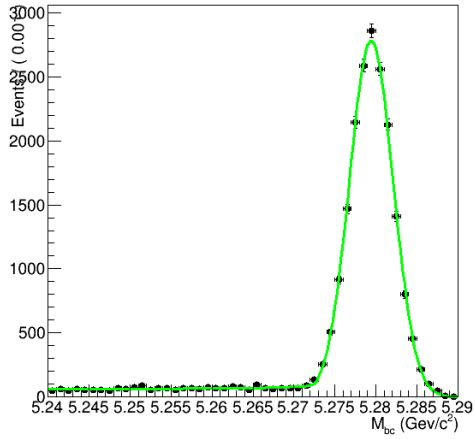
doi:10.6342/NTU201800725



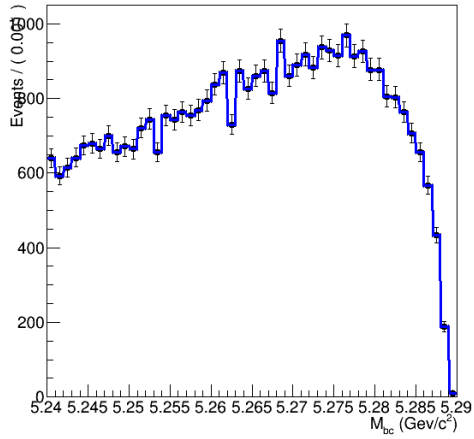
(a) Signal ΔE.



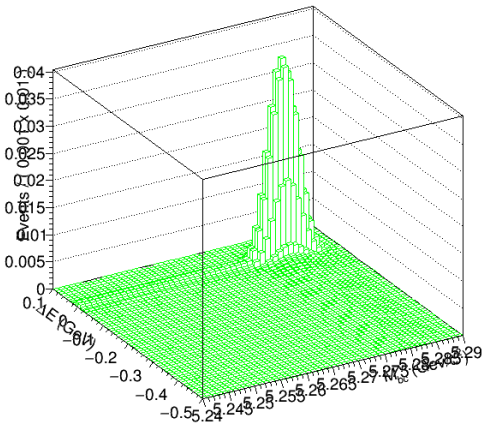
(b) Background ΔE.



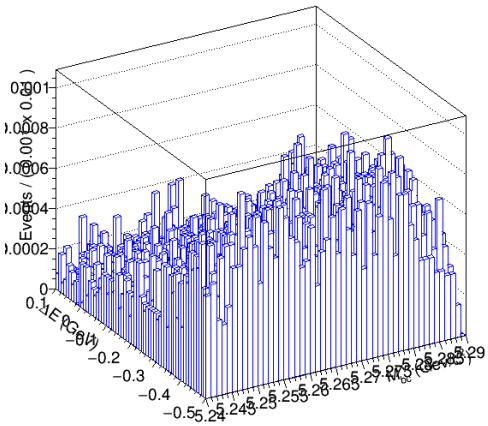
(c) Signal M_{bc}.



(d) Background M_{bc}.



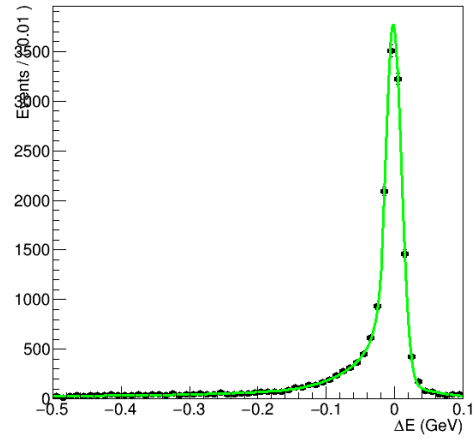
(e) 2 dimensional signal PDF.



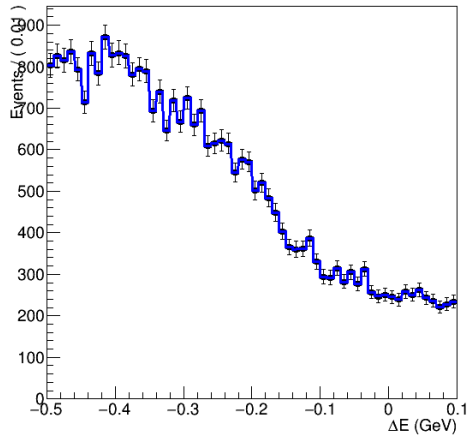
(f) 2 dimensional background PDF.

Figure 4.5: Fitting results for $B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$.

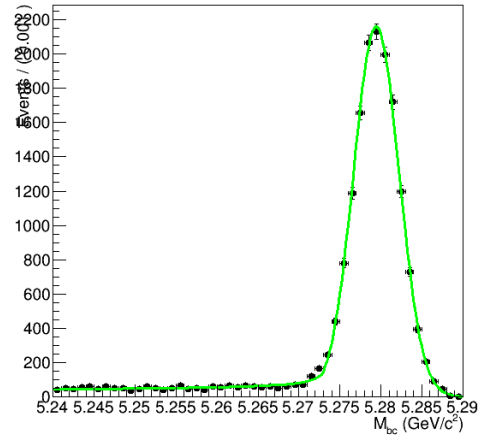
doi:10.6342/NTU201800725



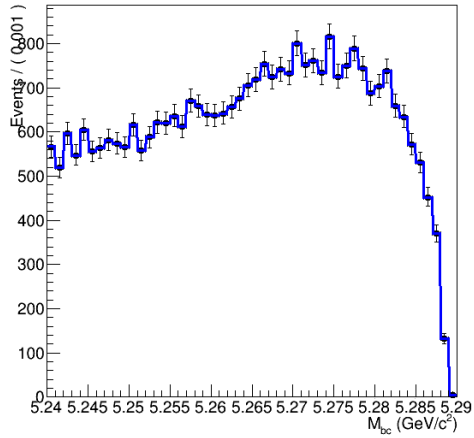
(a) Signal ΔE.



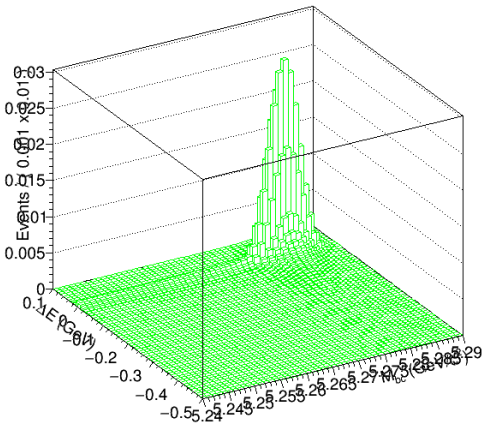
(b) Background ΔE.



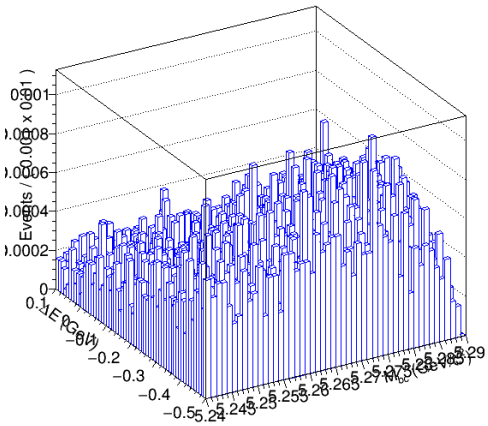
(c) Signal M_{bc}.



(d) Background M_{bc}.



(e) 2 dimensional signal PDF.



(f) 2 dimensional background PDF.

Figure 4.6: Fitting results for $B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$.

doi:10.6342/NTU201800725

ΔE for $B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$			
Function	Parameter	Value	Error
Gaussian	μ	4.503×10^{-3}	1.93×10^{-3}
	σ	3.7328×10^{-2}	1.19×10^{-3}
Crystal Ball	μ	4.205×10^{-4}	9.78×10^{-5}
	σ	1.0651×10^{-2}	1.13×10^{-4}
	α	2.5761	7.63×10^{-2}
	n	2.6807×10^{-1}	4.76×10^{-2}
Ratio $f^\dagger$		1.0614×10^{-1}	9.32×10^{-3}
M_{bc} for $B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$			
Function	Parameter	Value	Error
Crystal Ball	μ	5.2794	2.02×10^{-5}
	σ	2.6705×10^{-3}	1.52×10^{-5}
	α	2.5386	3.85×10^{-2}
	n	1.0874×10^{-1}	2.6×10^{-2}
ΔE for $B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$			
Function	Parameter	Value	Error
Gaussian	μ	-1.2161×10^{-2}	4×10^{-3}
	σ	5.9483×10^{-2}	2.8×10^{-3}
Crystal Ball	μ	-1.6315×10^{-3}	1.6×10^{-4}
	σ	1.1811×10^{-2}	1.6×10^{-4}
	α	1.4096	5.17×10^{-2}
	n	9.4797×10^{-1}	4.16×10^{-2}
Ratio $f^\dagger$		1.2256×10^{-1}	1.21×10^{-2}
M_{bc} for $B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$			
Function	Parameter	Value	Error
Crystal Ball	μ	5.2794	2.5×10^{-5}
	σ	2.7815×10^{-3}	2.01×10^{-5}
	α	2.2948	4.63×10^{-2}
	n	2.6927×10^{-1}	3.84×10^{-2}

Table 4.5: Modelling for signal Monte Carlo in control sample. †: $M = fF^{Gaussian} + (1-f)F^{CB}$.

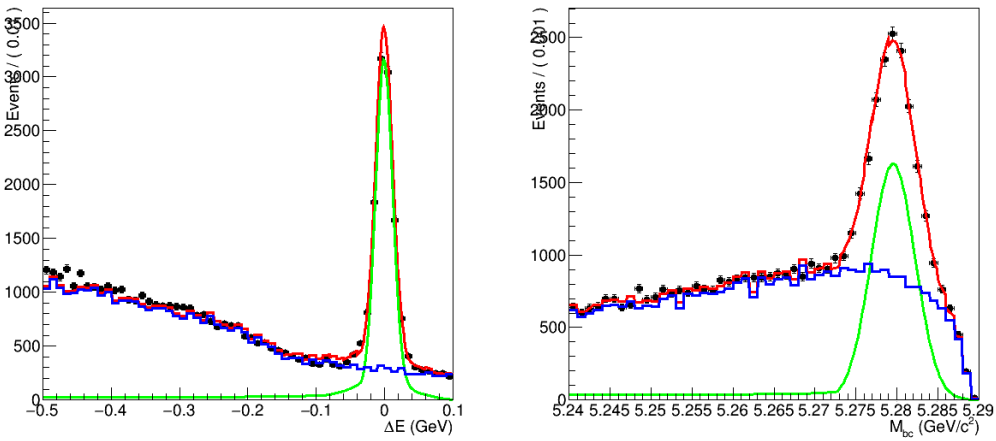
4.6 Calibration Factor

We will compare the difference between Monte Carlo and data by calculating calibration factor in this section. We use the same function in section 4.5 to fit data. When we fit the data, we float values of μ and σ and fix values of other parameters. We compare initial values and final values of μ and σ and calculate calibration factor.

doi:10.6342/NTU201800725

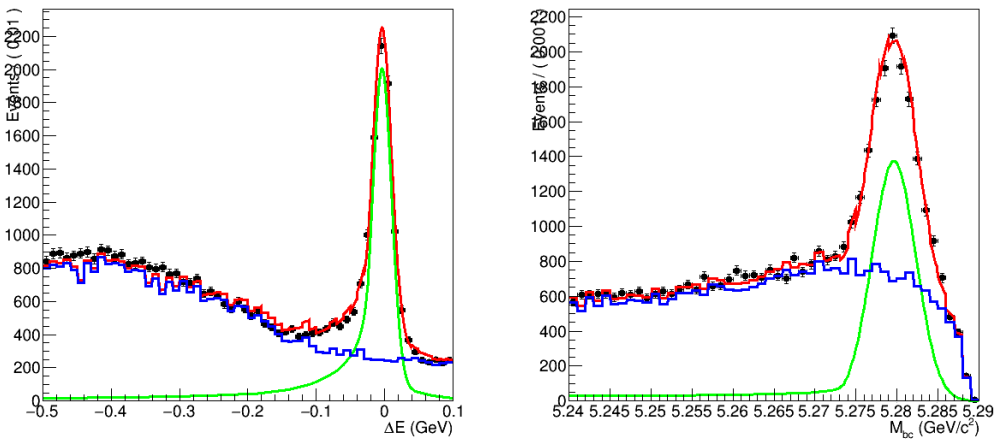
4.6.1 Calibration Factor Result

According to ratio in Table 4.5, we use two functions to fit ΔE, Crystal Ball function is the main composition in two functions. Thus, we float values of μ and σ of Crystal Ball function and fix values of other parameters in Table 4.5. The fitting results for data as shown in Fig. 4.7 and 4.8 and calibration factors are listed in Table 4.6.



(a) ΔE. (b) M_{bc}.

Figure 4.7: Fitting data for $B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$, The green lines are signal, blue lines are background and red lines are sum of them.



(a) ΔE. (b) M_{bc}.

Figure 4.8: Fitting data for $B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$, The green lines are signal, blue lines are background and red lines are sum of them.

doi:10.6342/NTU201800725

$B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$				
Parameter	Initial Value	Final Value	Error	Calibration Factor
$\mu_{\Delta E}$	4.205×10^{-4}	-5.9593×10^{-4}	1.55×10^{-4}	-1.01643×10^{-3} GeV
$\sigma_{\Delta E}$	1.0651×10^{-2}	1.2456×10^{-2}	1.38×10^{-4}	$+16.95\%$
$\mu_{M_{bc}}$	5.2794	5.2796	2.87×10^{-5}	$+2 \times 10^{-4}$ GeV/c^2
$\sigma_{M_{bc}}$	2.6705×10^{-3}	2.6573×10^{-3}	2.29×10^{-5}	-0.49%
$B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$				
Parameter	Initial Value	Final Value	Error	Calibration Factor
$\mu_{\Delta E}$	-1.6315×10^{-3}	-3.5379×10^{-3}	2.3×10^{-4}	-1.9064×10^{-3} GeV
$\sigma_{\Delta E}$	1.1811×10^{-2}	1.3886×10^{-2}	2.14×10^{-4}	$+17.57\%$
$\mu_{M_{bc}}$	5.2794	5.2797	3.3×10^{-5}	$+3 \times 10^{-4}$ GeV/c^2
$\sigma_{M_{bc}}$	2.7815×10^{-3}	2.6613×10^{-3}	2.63×10^{-5}	-4.32%

Table 4.6: Calibration factor result.

4.6.2 Branching Fraction

The data fitting results provide yield of number of signal and background. Using yield can calculate branching fraction of control sample modes. The yields of number are listed in Table 4.7.

$B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$		
Yield	Value	Error
$N_{datasig}$	1.2072×10^4	1.22×10^2
$N_{databkg}$	3.6184×10^4	1.97×10^2
$B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$		
Yield	Value	Error
$N_{datasig}$	1.0318×10^4	1.18×10^2
$N_{databkg}$	3.1134×10^4	1.86×10^2

Table 4.7: Yield of number for control sample.

And then, we calculate branching fraction, values of ϵ_{sig} are listed in Table 4.3:

$$\mathcal{B}(B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-) = \frac{N_{datasig}}{\epsilon_{sig} \times 7.71 \times 10^8} = (6.83 \pm 0.07) \times 10^{-5}, \tag{4.5}$$

$$\mathcal{B}(B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-) = \frac{N_{datasig}}{\epsilon_{sig} \times 7.71 \times 10^8} = (7.18 \pm 0.08) \times 10^{-5}. \tag{4.6}$$

doi:10.6342/NTU201800725

Chapter 5

Conclusion

5.1 Data Fitting

We use the fitting results in section 3.7.2 and calibration factor (We denote it by $\Delta\mu_{\Delta E,M_{bc}}$ and $\Delta\sigma_{\Delta E,M_{bc}}$.) results in section 4.6.1 to fit data of $B^0 \to l^+l^-l^+l^-$. Before we fit the data, we use calibration factor as the correction between Monte Carlo and data. Thus, we must modify μ and σ in signal function by calibration factor. $B^0 \to \mu^+\mu^-\mu^+\mu^-$ is modified by $B^0 \to J/\psi(J/\psi \to \mu^+\mu^-)K^+\pi^-$ and $B^0 \to e^+e^-e^+e^-$, $B^0 \to e^+e^-\mu^+\mu^-$ are modified by $B^0 \to J/\psi(J/\psi \to e^+e^-)K^+\pi^-$.

According to the ratio in Table 3.5, Gaussian function is main composition in two functions in ΔE of $B^0 \to \mu^+\mu^-\mu^+\mu^-$ and $B^0 \to e^+e^-\mu^+\mu^-$, Crystal Ball function is main composition in two functions in ΔE of $B^0 \to e^+e^-e^+e^-$. Thus, we modify μ and σ in Gaussian function of ΔE of $B^0 \to \mu^+\mu^-\mu^+\mu^-$ and $B^0 \to e^+e^-\mu^+\mu^-$ and modify μ and σ in Crystal Ball function of ΔE of $B^0 \to e^+e^-e^+e^-$. We list the results in Table 5.1.

We fix final values in Table 5.1 and values of other parameters in Table 3.5 and 3.6 to fit data, the fitting results are listed in Table 5.2 and Fig. 5.1.

doi:10.6342/NTU201800725

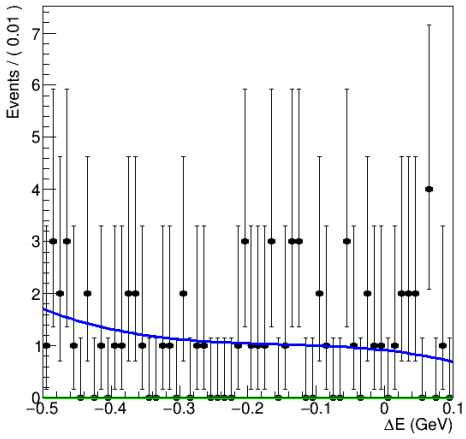
$B^0 \to \mu^+\mu^-\mu^+\mu^-$			
Parameter	Initial Value	Modification	Final Value
$\mu_{\Delta E}$	5.7342×10^{-4}	$\mu_{\Delta E} + \Delta\mu_{\Delta E}$	-4.4301×10^{-4}
$\sigma_{\Delta E}$	9.6711×10^{-3}	$\sigma_{\Delta E}(1 + \Delta\sigma_{\Delta E})$	1.1310×10^{-2}
$\mu_{M_{bc}}$	5.2795	$\mu_{M_{bc}} + \Delta\mu_{M_{bc}}$	5.2797
$\sigma_{M_{bc}}$	2.6282×10^{-3}	$\sigma_{M_{bc}}(1 + \Delta\sigma_{M_{bc}})$	2.6153×10^{-3}
$B^0 \to e^+e^-e^+e^-$			
Parameter	Initial Value	Modification	Final Value
$\mu_{\Delta E}$	-3.1956×10^{-3}	$\mu_{\Delta E} + \Delta\mu_{\Delta E}$	-5.102×10^{-3}
$\sigma_{\Delta E}$	1.3324×10^{-2}	$\sigma_{\Delta E}(1 + \Delta\sigma_{\Delta E})$	1.5665×10^{-2}
$\mu_{M_{bc}}$	5.2795	$\mu_{M_{bc}} + \Delta\mu_{M_{bc}}$	5.2798
$\sigma_{M_{bc}}$	2.912×10^{-3}	$\sigma_{M_{bc}}(1 + \Delta\sigma_{M_{bc}})$	2.7862×10^{-3}
$B^0 \to e^+e^-\mu^+\mu^-$			
Parameter	Initial Value	Modification	Final Value
$\mu_{\Delta E}$	-6.1077×10^{-4}	$\mu_{\Delta E} + \Delta\mu_{\Delta E}$	-2.5172×10^{-3}
$\sigma_{\Delta E}$	1.1064×10^{-2}	$\sigma_{\Delta E}(1 + \Delta\sigma_{\Delta E})$	1.3643×10^{-2}
$\mu_{M_{bc}}$	5.2795	$\mu_{M_{bc}} + \Delta\mu_{M_{bc}}$	5.2798
$\sigma_{M_{bc}}$	2.7547×10^{-3}	$\sigma_{M_{bc}}(1 + \Delta\sigma_{M_{bc}})$	2.6357×10^{-3}

Table 5.1: Modification by calibration factor.

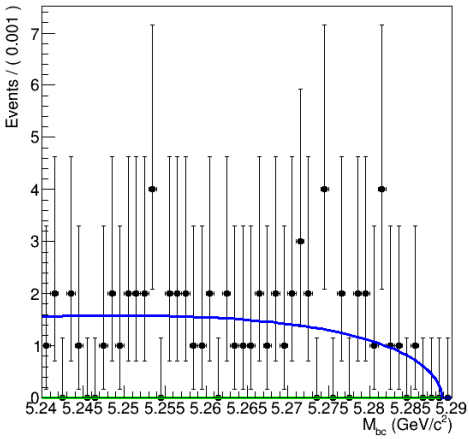
$B^0 \to \mu^+\mu^-\mu^+\mu^-$		
Yield	Value	Error
$N_{datasig}$	2.0314×10^{-6}	7.08
$N_{databkg}$	66.003	8.12
$B^0 \to e^+e^-e^+e^-$		
Yield	Value	Error
$N_{datasig}$	2.7301×10^{-1}	1.34
$N_{databkg}$	3.7318	2.22
$B^0 \to e^+e^-\mu^+\mu^-$		
Yield	Value	Error
$N_{datasig}$	1.3008×10^{-5}	8.57×10^{-1}
$N_{databkg}$	26.985	5.19

Table 5.2: Yield of number for $B^0 \to l^+l^-l^+l^-$.

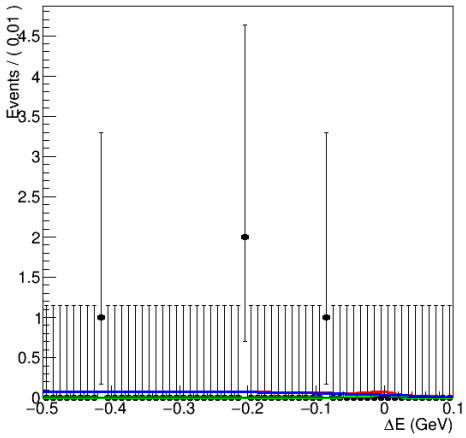
doi:10.6342/NTU201800725



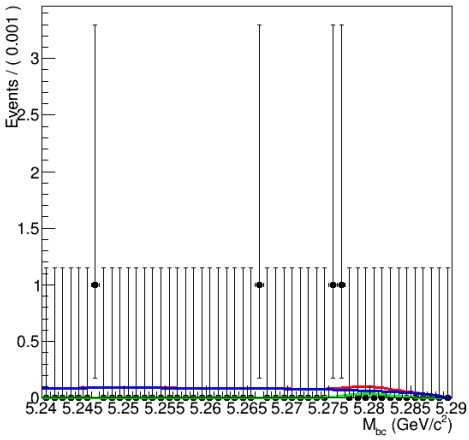
(a) ΔE in $B^0 \to \mu^+\mu^-\mu^+\mu^-$.



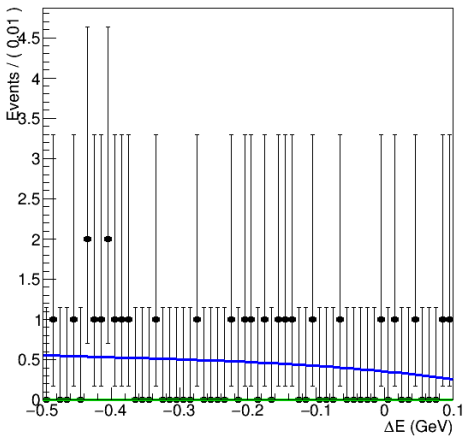
(b) M_{bc} in $B^0 \to \mu^+\mu^-\mu^+\mu^-$.



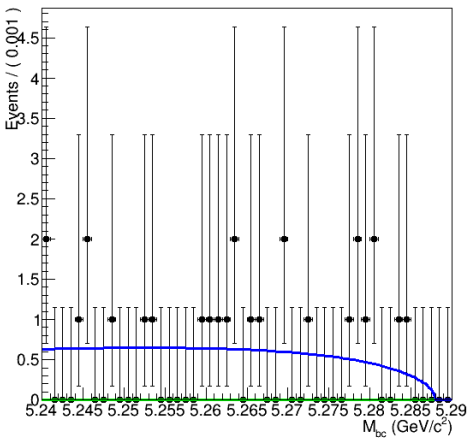
(c) ΔE in $B^0 \to e^+e^-e^+e^-$.



(d) M_{bc} in $B^0 \to e^+e^-e^+e^-$.



(e) ΔE in $B^0 \to e^+e^-\mu^+\mu^-$.



(f) M_{bc} in $B^0 \to e^+e^-\mu^+\mu^-$.

Figure 5.1: Fitting data for $B^0 \to l^+l^-l^+l^-$, The green lines are signal, blue lines are background and red lines are sum of them.

doi:10.6342/NTU201800725

5.2 Upper Limit Estimation

We use the Feldman-Cousins method [34] to estimate the upper limit of number (We denote it by N_{upper}) at 90% confidence interval. The method is that we input N_{data} and N_{MCbkg} in signal box and get N_{upper}, then we carry N_{upper} into equation (5.1) so that we can calculate the value of upper limit of branching fraction for $B^0 \to l^+l^-l^+l^-$ at 90% confidence interval, values of ϵ_{sig} are listed in Table 3.4. The upper limit results are listed in Table 5.3.

$$\mathcal{B}(B^0 \to l^+l^-l^+l^-) = \frac{N_{upper}}{\epsilon_{sig} \times 7.71 \times 10^8}. \tag{5.1}$$

N_{data}	N_{MCbkg}	N_{upper}	Upper Limit
7	6.33	6.205 at 90% C.L.	$\mathcal{B}(B^0 \to \mu^+\mu^-\mu^+\mu^-) < 1.28 \times 10^{-7}$ at 90% C.L.
1	1.33	3.055 at 90% C.L.	$\mathcal{B}(B^0 \to e^+e^-e^+e^-) < 8.34 \times 10^{-8}$ at 90% C.L.
4	7.50	1.885 at 90% C.L.	$\mathcal{B}(B^0 \to e^+e^-\mu^+\mu^-) < 5.09 \times 10^{-8}$ at 90% C.L.

Table 5.3: Upper limit of $\mathcal{B}(B^0 \to l^+l^-l^+l^-)$, N_{data} is number of data in signal box and N_{MCbkg} is number of background MC in signal box for 1 stream.

5.3 Conclusion

We have fitted data and estimated upper limit of branching fraction at 90% confidence interval. According to these results, we can't find any signal in data of each modes. In other words, we don't get significant anything beyond the Standard Model. These results are still under the Standard Model. The number of data in Belle II is several times larger than Belle I, so is the luminosity. However, branching fraction of $B^0 \to l^+l^-l^+l^-$ is too small to discover significant anything in Belle II.

doi:10.6342/NTU201800725

Bibliography

[1] HiggsTan. http://higgstan.com/standerd-model/.

[2] Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Elementary_particle_interactions_in_the_Standard_Model.png.

[3] K.A. Olive et al. (Particle Data Group). Review of Particle Physics. *Chin. Phys. C*, 38(090001), 2014.

[4] K. Abe et al. (Belle Collaboration). Observation of Large CP Violation in the Neutral B Meson System. *Phys. Rev. Lett.*, 87(091802), 2001.

[5] HyperPhysics. http://hyperphysics.phy-astr.gsu.edu/hbase/Particles/expar.html.

[6] Discovery of the Bottom Quark. http://www.fnal.gov/pub/inquiring/physics/discoveries/bottom_quark_pr.html.

[7] D. Besson and T. Skwarnicki. UPSILON SPECTROSCOPY: Transitions in the Bottomonium System. *Annu. Rev. Nucl. Part. Sci.*, 43:333–78, 1993.

[8] J. Beringer et al. (Particle Data Group). Review of Particle Physics. *Phys. Rev. D*, 86(010001), 2012.

[9] Brian Batell, Maxim Pospelov, and Adam Ritz. Multi-lepton Signatures of a Hidden Sector in Rare B Decays. *Phys. Rev. D*, 83(054005), 2011.

[10] S. V. Demidov and D. S. Gorbunov. Flavor violating processes with sgoldstino pair production. *Phys. Rev. D*, 85(077701), 2012.

doi:10.6342/NTU201800725

[11] Dmitri Melikhov and Nikolai Nikitin. Rare radiative leptonic decays $B_{d,s} \to l^+l^-\gamma$. *Phys. Rev. D*, 70(114028), 2004.

[12] R. Aaij et al. (LHCb Collaboration). Search for rare $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ decays. *Phys. Rev. Lett.*, 110(211801), 2013.

[13] KEKB. http://www2.kek.jp/accl/introKEKB/.

[14] Wikimedia Commons. https://commons.wikimedia.org/wiki/File:KEKB.png.

[15] A. Abashian et al. (Belle Collaboration). The Belle detector. *Nuclear Instruments and Methods in Physics Research A*, 479:117–232, 2002.

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

[17] Z. Natkaniec et al. (Belle SVD2 Group). Belle SVD2 Vertex Detector. *Nucl. Instrum. Methods Phys. Res. Sect. A*, 568, 1:269–273, 2006.

[18] K. Hanagaki et al. Electron Identification in Belle. *Nucl. Instrum. Methods Phys. Res. Sect. A*, 485:490–503, 2002.

[19] BASF Overview (Belle internal). http://belle.kek.jp/secured/wiki/doku.php?id=software:basf.

[20] The EvtGen package home page. http://www.slac.stanford.edu/~lange/EvtGen/.

[21] CLEO. https://www.classe.cornell.edu/Research/CLEO/.

[22] BaBar. http://www.slac.stanford.edu/BFROOT/.

[23] Anders Ryd et al. EvtGen A Monte Carlo Generator for B-Physics. http://evtgen.hepforge.org/doc/EvtGenGuide.pdf.

[24] CERN. https://home.cern/.

[25] CERN. *GEANT Detector Description and Simulation Tool*.

doi:10.6342/NTU201800725

[26] ROOT Users Guide. https://root.cern.ch/root/htmldoc/guides/users-guide/ROOTUsersGuide.html.

[27] Wikimedia. https://en.wikipedia.org/wiki/Monte_Carlo_method.

[28] Wikimedia. https://en.wikipedia.org/wiki/Observer-expectancy_effect.

[29] Ohio State University Department of Physics. https://www.physics.ohio-state.edu/~gan/teaching/spring04/Chapter6.pdf.

[30] UW Madison Department of Astronomy. http://www.astro.wisc.edu/~bank/.

[31] Giovanni Punzi. Sensitivity of Searches for New Signals and Its Optimization. *PHYSTAT2003, SLAC, Stanford*, 2003.

[32] C. Patrignani et al. (Particle Data Group). Review of Particle Physics. *Chin. Phys. C*, 40(100001), 2016.

[33] Ryosuke Itoh. Measurements of Polarization CP Asymmetry in $B \to J/\psi K^*$ decays. *Belle note 487*, 2001.

[34] Gary J. Feldman and Robert D. Cousins. A Unified Approach to the Classical Statistical Analysis of Small Signals. *Phys. Rev. D*, 57(3873), 1998.

doi:10.6342/NTU201800725