Observation of Large CP Violation in the Neutral B Meson System Using $B^0 \to J/\psi\, K_L$ Decay

Masashi Yokoyama
Department of Physics, University of Tokyo

December 2001

Abstract

We present a measurement of the CP asymmetry parameter $\sin 2\phi_1$ at the KEKB asymmetric-energy e^+e^- collider. Using a data sample of 29.1 fb^{-1} collected at the $\Upsilon(4S)$ resonance with the Belle detector, we reconstruct 523 candidate events in decays of neutral B mesons to the CP even eigenstate $J/\psi K_L$. The flavor of the accompanying B meson is identified from its decay products. An unbinned maximum likelihood fit is carried out to extract $\sin 2\phi_1$ from the asymmetry in the distribution of the time intervals between the two B meson decay points. We obtain

$$\sin 2\phi_1 = 1.31^{+0.19}_{-0.23}(stat.) \pm 0.12(syst.).$$

We have also measured $\sin 2\phi_1$ using 36 candidate events in decays of neutral B mesons to $J/\psi K^{*0}$, followed by K^{*0} decays into $K_S\pi^0$. We use the angular information of decay daughters of the B mesons to separate CP odd and even components in the final state. Using the fraction of CP odd component $f_{\rm odd} = 0.19 \pm 0.06$ which is determined from a full-angular analysis of B decays into $J/\psi K^*$, we obtain

$$\sin 2\phi_1 = 0.97^{+1.38}_{-1.40}(stat.) \pm 0.19(syst.).$$

We conclude that we have observed CP violation in the neutral B meson system. This is the first observation of CP violation outside the kaon system.

Acknowledgements

There is no doubt that this thesis is not a product of my personal work alone. In fact, it is the result of enormous efforts by lots of people. I would like to acknowledge everyone who helped me directly or indirectly to complete this long and exciting journey, but I have realized it may need almost an infinite list to do so. Here I would like to express my gratitude to those who especially impressed me, but my heartfelt thanks also go to all of those who I cannot include their name (intentionally or unintentionally).

First of all, I would like to express my deepest gratitude to my supervisor, Prof. Hiroaki Aihara, for giving me an opportunity for the research in the wonderful world of high energy physics and for his support and guidance for my research activity.

I am grateful to all the students in our group at the University of Tokyo. I will never forget the early days of our group, when I enjoyed my life with T. Higuchi, J. Tanaka, T. Nakadaira, and T. Tomura at our old building. It was my great fortune to have worked with people who possess remarkable talent like them. I enjoyed very much to talk with younger students who helped me a lot, T. Matsubara, H. Kawai, N. Uozaki, and Y. Yamashita.

I am deeply thankful to the members of SVD group. J. Haba, M. Tanaka, Y. Yamada, and T. Tsuboyama for unforgettable memories of works at KEK machine shop and Tsukuba-B4; M. Hazumi and H. Ishino for their management of SVD software group; M. Różańska, A. Bożek, and H. W. Zhao for their help in the SVD alignment work; T. Karim and T. Hara for their help in SVD software development; Y. Ushiroda and K. Kawasaki for their hard work for Belle and SVD; Z. Natkaniec for his sense of humor which encouraged me in my hardest days at underground Tsukuba.

My warmest gratitude goes to the members of physics analysis groups. Y. Sakai and M. Hazumi for their coordination of ICPV group; K. Hanagaki and K. Sumisawa for their leadership in CP-fit microgroup; K. Abe, D. Marlow, M. Yamaga, and S. E. Vahsen for their preceding work, help and advice in $J/\psi K_L$ analysis; R. Itoh for his excellent analysis of $J/\psi K^*$ mode.

I would like to thank all the other graduate students in the Belle collaboration. K. Hara and H. Miyake for the memory of summer 2000; T. Tomoto, H. Kakuno, and S. Nishida for their friendship and help in the life at KEK; K. Suzuki for his surprizing marriage; J. Kaneko for the memory of the travel to Lyon; C. C. Wang

and M. C. Chang for their help in SVD replacement; R. Abe, F. Ohno, Y. Harada for their recent work for SVD upgrade.

I would like to express my sincere gratitude to the secretaries of the collaboration, R. Ohta(Suzuki)-san and C. Imai-san. I am also deeply thankful to the secretaries of the University of Tokyo, Kono-san, Hikami-san, and Enomoto-san for their help and enjoyable conversations we made during my hard days.

My special thanks go to Hiroyasu Tajima, who was the first JOSHU of our lab, the leader of SVD group, and who led the analysis team for *CP* violation measurement. He taught me a lot of things related to particle detectors, physics, data analysis, and human life. Almost all my research works over five years were done under his strong and powerful guidance. I am sure that this thesis did not exist without his help and encouragement. I hope he can make good achievement in his new career, but actually I am sure he will make it as he did in the past.

At last, but not the least, I would like to thank my family who have been supported and encouraged me for a long time. My parents, Motoya and Toshiko Yokoyama, have been continuously supporting me since I was born. I cannot thank them enough for their love they gave me for years. They kindly supported me to enter the graduate school, although they seem to have misunderstood my research program as famous Kamiokande for a while. My brother and sister, Kimiya and Asako, also supported me in various ways. Although we are living apart from each other, I believe we are always connected as a family.

My wife, Rika, has been always the largest source of vitality for me since we met at the high school about ten years ago. I really appreciate her patience and encouragement for my life as a graduate student in science, which is not so popular way to live. There was very tough time for both of us, but we also had many precious experiences which made our bonds of affection tighter. I hope it is tight enough to overcome any troubles in the future, but I hope more seriously not to have such a chance that we can test our relationship...

Now, it's over. Thank you all!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

We are living in a universe which is made of *matter*. However, according to the modern theories of cosmology, equal populations of *matter* and *anti-matter* should have been produced in the beginning of the universe. Then, a very natural and profound question arises; from where does the asymmetry between *matter* and *anti-matter* we are looking now come from?

A. Sakharov [1] pointed out that there are three essential elements to construct theories which can explain the excess of *matter* over *antimatter* in the universe:

1. departure from thermal equilibrium;
2. reactions that change baryon number have to occur in the early universe; and
3. the existence of ***CP* violation**.

Here, ***CP* violation** means there exists difference of properties between *matter* and *antimatter*. *CP* violation is thus one of essential elements in any attempt to understand the history of our universe.

Until 1964, physicist believed there was no difference between matter and anti-matter, despite antimatter has opposite-signed charge and internal quantum numbers. In other words, it was believed that there was *symmetry* between matter and antimatter, that is *CP* symmetry. The violation of *CP* symmetry was (totally unexpectedly) found in decays of neutral K mesons in 1964 [2]. Since then, an enormous effort has been done both theoretically and experimentally to reveal the origin of this phenomenon.

In 1973, Kobayashi and Maskawa (KM) [3] proposed a theory of quark mixing which can introduce *CP* violation within the framework of the *Standard Model* (SM) of elementary particle physics. They demonstrated that quark-flavor mixing matrix

with measurable complex phase introduces *CP* violation into quark interactions. This requirement is satisfied if there are at least six flavors of quarks, while only three types of quarks (u, d, s) were known when KM proposed their model. The fourth quark, charm or c, is discovered in 1974. The fifth and sixth quark (b and t) were discovered in 1977 and 1995, respectively. The KM model for *CP* violation is now considered to be an essential part of the SM and quark mixing matrix is now known as Kobayashi-Maskawa (KM) matrix.

However, it is not yet known whether the mechanism of *CP* violation realized in the nature is the one predicted by the KM model. There are other theoretical models that are consistent with presently-available experimental data, in spite of a considerable amount of experimental effort over past three decades. In fact, the mechanism of *CP* violation is one of the least experimentally tested part of the SM. Therefore, it is one of the most interesting topics of high energy physics today to test *CP* violating mechanism in the Standard Model, or to look for *CP* violation from beyond Standard Model sources. The key to the test of the KM mechanism is the measurement of the *Unitary Triangle*, which shows relations between the KM matrix elements. The KM model provides definitive predictions for three angles of the Unitary Triangle, ϕ_1, ϕ_2 and ϕ_3.[1] Any deviation from predicted values indicates contribution from physics beyond the SM.

In 1981, Sanda, Bigi and Carter [4] pointed out that the KM model predicted large *CP* violation in certain decays of B mesons, which consists of a b-quark and one of lighter quarks. Subsequent measurements of B meson properties [5, 6] indicated that B meson system provides very wealthy field to test *CP* violation theories, but the experimental study of *CP* violation in B meson decays requires very high-performance accelerator and detector due to its heavy mass, short lifetime, and small branching fractions to specific decay modes. The largest observable *CP* asymmetry effects are expected to show up in the difference of the partial decay rates between B^0 and $\overline{B}^0$ mesons to the same *CP* eigenstate. A resonance $\Upsilon(4S)$ is the best place to obtain pure samples of B^0 and $\overline{B}^0$ mesons. However, the measurement of *CP* asymmetry requires time-dependent measurement of B^0 and $\overline{B}^0$ decay rates, which is impossible with ordinary e^+e^- colliders.

An asymmetric e^+e^- collider at $\Upsilon(4S)$ resonance was proposed to produce $B^0/\overline{B}^0$ mesons boosted in the laboratory frame, allowing current vertex measurement technology to measure the time evolution of B^0 and $\overline{B}^0$ decays. Since each decay mode of B meson has a typical branching fraction of order 10^{-4}, it is also necessary to

[1] These angles are also known as β, α and γ, respectively.

produce a large number of $B^0\overline{B}^0$ pairs. The KEKB accelerator was designed and constructed to achieve these goals. The energy of electron(positron) beam is chosen to be 8.0(3.5) GeV, with which the Lorentz boost factor $\beta\gamma$ is 0.425 and the average B decay length is $\sim$200 μm. The design luminosity of the KEKB is 10^{34} cm^{-2}s^{-1}, corresponding to 10^8 $B^0\overline{B}^0$ pairs in a year.

A high-performance detector is also necessary to measure the *CP* asymmetry. Belle[2] detector was designed and constructed to have capability of measuring *CP* violation in B meson decays, *i.e.* efficient reconstruction of rare final states of B mesons, good vertex resolution and good particle identification.

Among three *CP* angles of the *Unitary Triangle*, ϕ_1 is expected to be the most accessible one, and its measurement is one of primary goals of the B-factory experiment. The decay $B^0 \to J/\psi K^0$ [3] is considered to be the cleanest channel for measurement of ϕ_1. Experimentally, it contains very clean signal of $J/\psi \to \ell\ell(\ell = e, \mu)$ decays, which enable us to reduce the background significantly. In addition, its decay amplitude is dominated by a single decay diagram, allowing a straightforward extraction of *CP* angle from the measured asymmetry.

There are two decay modes with opposite *CP* eigenvalues in the final state, depending on the *CP* eigenvalue of K^0. $B^0 \to J/\psi K_S$, with *CP* eigenvalue $\xi_f = -1$, is almost background free and called the "Golden Mode" for measuring ϕ_1. On the other hand, $B^0 \to J/\psi K_L$, with *CP* eigenvalue $\xi_f = +1$, provides an independent measurement of $\sin 2\phi_1$. Its importance goes beyond the increase in statistics. Because it has opposite *CP* eigenvalue to the $J/\psi K_S$ channel, $J/\psi K_L$ mode gives us a handle to understand the possible systematic bias in the $\sin 2\phi_1$ measurement. If there is such an effect, we should see opposite systematic effect between *CP* even and odd modes, since their asymmetry should be equal in magnitude but opposite in sign. This is an excellent check of the measurement procedure. For these reasons $B^0 \to J/\psi K_L$ mode is as important as the Golden Mode in $\sin 2\phi_1$ measurement, although it is expected to have more background due to experimental difficulty in K_L detection.

Apart from other charmonium+K^0 modes (like $\psi(2S)K^0$ or $\chi_{c1}K^0$), next mode to be used in the ϕ_1 measurement is $B^0 \to J/\psi K^{*0}$, where K^{*0} decays into K_S and

[2]The word "Belle" means "beauty" in French. *b*-quark is sometimes called *beauty*-quark. Also, "Belle" consists of a B, el (=electron) and le (=reversal of el=positron), symbolizing e^+e^- B-factory.

[3]Throughout this thesis, when a mode is quoted the inclusion of the charge conjugate mode is implied.

π^0. Although this decay mode is *CP* eigenstate and its decay proceeds with the same quark diagram as the Golden Mode, both decay products have spin 1 and the final state is a mixture of *CP* odd and even states. We must use angular information of decay products to separate contributions from different *CP* components.

In this thesis, we present a measurement of the *CP* asymmetry parameter $\sin 2\phi_1$ in $B^0 \to J/\psi\, K_L$ using data collected by the Belle detector at the KEKB asymmetric e^+e^- collider. We also present the first attempt to measure $\sin 2\phi_1$ using $B^0 \to J/\psi\, K^{*0}$ decays with decay angle information.

The outline of this thesis is as follows: Theoretical description of *CP* violation and experimental consideration to measure it in the neutral B meson system is provided in Chapter 2. Experimental apparatus used in this analysis is described in Chapter 3. The reconstruction procedure of candidates for $B^0 \to J/\psi\, K_L$ decays is explained in Chapter 4. The flavor tagging and vertex reconstruction procedure are described in Chapter 5. The *CP* asymmetry parameter $\sin 2\phi_1$ is extracted in Chapter 6. The check of possible systematics we made is described in Chapter 7. Finally, we discuss the result and conclude this thesis in Chapter 8.

Chapter 2

Phenomenology of CP Violation in the Neutral B Meson System

In this chapter, we provide necessary tools to describe CP violation in the neutral B meson system in the context of Kobayashi-Maskawa (KM) mechanism. We start from discussion about discrete symmetries in particle physics and discoveries of their violations. Then, we briefly review the development of theories of quark flavor mixing, which led us to the KM mechanism. We also discuss the formalism of time evolution of the neutral B meson system and mixing-induced CP violation. We then explain why we can expect large CP violation in the B meson system and what is necessary to experimentally observe such an effect. Finally, we review the experimental status of the test of the KM mechanism. More detailed discussion can be found, for example, in [7].

2.1 Symmetries in Physics

2.1.1 C, P, T, and CPT

Symmetries, like invariance under rotations, spatial translations, or temporal translations have been familiar in physics for hundreds of years. They often played very important roles to understand hidden laws behind observed phenomena. E. Noether proved [8] that every continuous symmetry results in an associated conserved quantity. Above three symmetries lead to the conservation of angular momentum, linear momentum and energy, respectively. These symmetries are called *continuous* because they corresponds to invariance under a continuous group of transformations.

There is another class of symmetries, *discrete* symmetries like invariance under

Parity (the reversal of the three spatial dimension coordinates). In the development of theories of elementary particles, discrete symmetries have been very valuable tools. Of particular interest in this field are three discrete symmetries:

- Parity(P), a transformation which reverses three-dimensional coordinates.
- Charge conjugation (C), an operation which replaces all particles with their anti-particles (and *vice versa*).
- Time reversal (T), inversion of the time coordinate.

If a state is an eigenstate of these symmetries (or combinations of them), it has an eigenvalue of $+1$ or -1.[1] We call the eigenstate with an eigenvalue of $+1(-1)$ even(odd) state. Particles like mesons or baryons, gauge bosons may have *intrinsic* eigenvalues of above symmetries. For example, pions have intrinsic eigenvalue of P -1 (usually written as pion has $P = -1$ or odd Parity).

The combination of above three discrete symmetries, CPT, has special importance in the quantum field theory. General principles of relativistic field theory require CPT should be conserved. This invariance of CPT is very fundamental consequence of relativistic field theory and often called *CPT theorem*. The CPT theorem assures the equality of masses, lifetimes and magnitude of charges between particles and anti-particles. Up to now, no CPT violation has been observed experimentally.

It had been naively believed in the early days of particle physics that individual symmetries, C, P, and T, would be also conserved in particle interactions. However, we eventually found the nature was more mysterious than we had believed – all these individual symmetries were proved to be violated.

2.1.2 Discovery of P violation

There was a problem in the particle physics in 1950s, which was called "θ–τ puzzle". Two particles, θ and τ, [2] were found through their weak decays into

$$\theta^+ \to \pi^+\pi^0, \tag{2.1}$$

$$\tau^+ \to \pi^+\pi^+\pi^-. \tag{2.2}$$

[1] Exactly speaking, the eigenvalues have an arbitrary phase $e^{i\theta}$. Usually we adopt $\theta = 0$ so that the eigenvalues become ± 1.

[2] This is different particle from what is known as τ today (third-generation charged lepton), as is explained below.

They seemed to have absolutely the same mass and lifetime, but their final states have opposite parities (note that pions have $P = -1$). The θ and τ should have different parities so that they are different particles (that was why they were given different name!). Then, how could the nature give the same mass and lifetime to two distinct particles?

Lee and Yang solved this puzzle in 1956 [9]. They investigated the experiments performed in past and found the fact that the parity conservation had been tested extensively in the electromagnetic and strong interactions, but not in the weak interactions and it was possible parity is violated in the weak interaction. They proposed several experiments which could test the parity conservation in the weak interaction.

In 1957, Wu and her collaborators performed one of the proposed experiments [10]. They studied the angular distribution of β rays from polarized ^{60}Co nuclei. If parity were conserved, there must be no asymmetry in the distribution. They observed clear evidence for asymmetry, which was the first observation of the parity violation. Thus, "θ–τ *puzzle*" is no longer a puzzle, since θ^+ and τ^+ can be an *identical* particle, which is now called K^+.

Amazingly, it was found in successive experiments that parity is maximally violated in the weak interaction. An example is the decay of charged pion, $\pi^+ \to \mu^+\nu_\mu$. If parity were conserved, the muons from this decay would not exhibit any polarization along their direction of motion. It was found that the muons from $\pi^+ \to \mu^+\nu_\mu$ is maximally polarized and muons from $\pi^- \to \mu^-\overline{\nu}_\mu$ is oppositely polarized. This indicates that only

$$\pi^+ \to \mu^+\nu_{\mu L} \tag{2.3}$$

and

$$\pi^- \to \mu^-\overline{\nu}_{\mu R} \tag{2.4}$$

can take place in the nature, where the subscripts "L" and "R" denote the helicity of (anti-)neutrinos are left-handed and right-handed, respectively. [3]

This demonstrates that C is also violated, since charge conjugation of (2.3) will be

$$\pi^- \to \mu^-\overline{\nu}_{\mu L}, \tag{2.5}$$

which is not allowed.

[3]The helicity, H, is defined as the sign of spin along the direction of motion. We call particles with $H = +1(-1)$ right-(left-)handed.

However, product of C and P still seemed to be good symmetry; CP conjugation of (2.3) will be

$$\pi^- \to \mu^- \overline{\nu}_{\mu R}, \tag{2.6}$$

which is identical with (2.4) and thus allowed.

2.1.3 Neutral kaon system and *CP* invariance

There was an interesting history about the neutral kaons and CP invariance; existence of long-lived neutral kaon was predicted using CP invariance, whereas it was the long-lived neutral kaon system where CP violation was found!

Two species of the neutral kaon

Since the neutral kaon, K^0 carries a non-zero quantum number *strangeness*, K^0 should have an antiparticle $\overline{K}^0$, distinct from itself. Mixtures of the two states would be eigenstates of CP transformation:

$$|K_1^0> = \frac{|K^0> + |\overline{K}^0>}{\sqrt{2}} \tag{2.7}$$

$$|K_2^0> = \frac{|K^0> - |\overline{K}^0>}{\sqrt{2}} \tag{2.8}$$

The K_1^0 has $CP = +1$ and K_2^0 has $CP = -1$. [4] These states would not be so interesting if there were no interaction that relates K^0 and $\overline{K}^0$, but in reality, the K^0 and $\overline{K}^0$ can *mix*. *i.e.* there is a transition between two states through "box diagrams" (Figure 2.1) and an initially pure K^0 or $\overline{K}^0$ state will be a mixture of two states after a certain time. Then, K_1^0 and K_2^0 states can arise from pure K^0 or $\overline{K}^0$ states.

Since CP eigenvalue of two pion system is $+1$, only K_1^0 can decay into 2π state,

$$K_1 \to 2\pi, \tag{2.9}$$

$$K_2 \nrightarrow 2\pi. \tag{2.10}$$

Then, the leading non-leptonic decay channel for K_2 is

$$K_2 \to 3\pi. \tag{2.11}$$

[4] We take the convention $CP|K^0> = |\overline{K}^0>$.

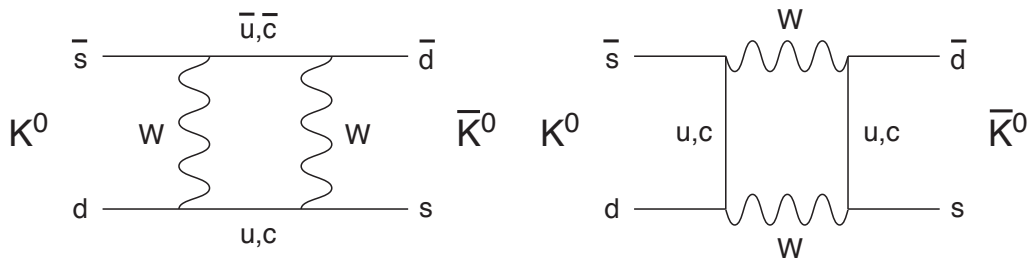

Figure 2.1: Box diagrams responsible for $K^0 - \overline{K}^0$ mixing

An accident produced huge difference in the property of two states; since $M_\pi \simeq$ 140 MeV and $M_K \simeq 500$ MeV, the phase space for Eq. (2.11) is very restricted compared to that of Eq. (2.9). Thus we expect the lifetime for K_2 to be much longer than that for K_1. Gell-Mann and Pais [11] predicted a long-lived meson based on above argument[5] in 1955, and Lederman and his collaborators [12] discovered the long-lived particle in 1956. Since K_1 and K_2 possess quite different lifetimes, they are usually referred to as K_S and K_L, respectively, where S [L] referring to short-lived [long-lived]. The difference of their lifetimes is huge;

$$\tau_{K_S} = (0.8940 \pm 0.0009) \times 10^{-10}\text{s} \tag{2.12}$$

$$\tau_{K_L} = (5.15 \pm 0.04) \times 10^{-8}\text{s} \tag{2.13}$$

Discovery of *CP* violation

As described in the previous section, violations of P and C were found in mid-1950s. However, the combined symmetry, CP, was still believed to be conserved. The observed violations of P and C exactly canceled in the combination of CP, in the case for the decay $\pi^+ \to \mu^+\nu_\mu$.

In 1964, Christenson, Chronin, Fitch and Turlay were studying decays of K^0 mesons using K_L beam from AGS at Brookheaven. When they were trying to set upper limit on the CP-violating $K_L \to 2\pi$ decay — they found the signal.

After exploring as many alternative interpretations as possible, they reported the observation of $K_L \to \pi^+\pi^-$ with tiny branching ratio [2],

$$\frac{\Gamma(K_L \to \pi^+\pi^-)}{\Gamma(K_L \to \text{all charged modes})} = (2.0 \pm 0.4) \times 10^{-3}. \tag{2.14}$$

[5] They originally assumed only C invariance.

This result implies that K_L state is not pure CP odd state K_2, but it contains a small admixture of a CP even component:

$$|K_L >= \frac{\epsilon|K_1 > +|K_2 >}{\sqrt{1+|\epsilon|^2}}. \tag{2.15}$$

From CPT theorem, K_S should contain CP odd component controlled by the same parameter ϵ:

$$|K_S >= \frac{|K_1 > +\epsilon|K_2 >}{\sqrt{1+|\epsilon|^2}}. \tag{2.16}$$

This was the first observation of CP violation.[6]

Since then, CP violation has been observed only in the weak interactions in neutral kaon system.

2.1.4 *CP* violation and cosmology

CP violation also plays an important role in cosmology. According to the modern theories of cosmology, our universe began with an explosion, called *Big Bang*. In the initial stage of the Big Bang, equal numbers of matter and anti-matter should have been produced. However, as we know, this is not true in the present universe; we see only matter on the earth, except eccentric scientists in the laboratory. Cosmological observations indicate that this is widely true in the universe, *i.e.* there is large imbalance between the amount of matter and anti-matter.

A. Sakharov [1] pointed out that there are three essential elements to construct theories which can explain the excess of matter over antimatter in the universe.

1. We need some interactions which violates *baryon number*, that is defined as $+1$ for baryons and -1 for anti-baryons. Up to now, there is no experimental evidence for such interactions, but it could be realized in theories beyond the Standard Model (for example, by introducing supersymmetry).

2. CP invariance has to be broken. Otherwise for every baryon number changing transition, $N \to f$, there is its CP conjugate one, $\overline{N} \to \overline{f}$, with the equal transition rate and no net baryon number can be generated.

3. If we believe the CPT theorem, the above transitions have to proceed out of thermal equilibrium. In thermal equilibrium, time becomes irrelevant globally

[6] This CP violating parameter ϵ is so tiny that the argument of two neutral K mesons described above is still valid.

and *CPT* invariance reduces to *CP* symmetry, which has to be avoided as we have seen above.

Thus, *CP* violation is an essential key to understand the history of our universe.

2.2 Theories of Quark Flavor Mixing

Since the discovery of the *CP* violation, people have been exploring its mechanism and origin. Theorists proposed numbers of models to explain the mechanism of *CP* violation. In the framework of the Standard Model, the origin of *CP* violation is explained by a single complex phase in the quark mixing matrix. This mechanism was first proposed by Kobayashi and Maskawa in 1973. We first discuss quark mixing, which is proposed by Cabibbo and extended by Glashow, Illiopoulos, and Maiani (GIM). Then, the idea of KM will be presented.

2.2.1 Cabibbo angle and GIM mechanism

Early measurements of the semi-leptonic decay rates of strange particles, characterized by a change in strangeness $\Delta S = 1$, e.g.,

$$\Sigma^- \to n + e^- + \overline{\nu}_e \tag{2.17}$$

indicated that they were suppressed by a factor of about 20 compared with $\Delta S = 0$ decays, such as

$$n \to p + e^- + \overline{\nu}_e. \tag{2.18}$$

At the same time, it was noticed that the value of the coupling constant of weak interactions, called the Fermi constant G, deduced from the unsuppressed nuclear β-decay such as (2.18) was slightly less than that deduced from the muon decay.

These factors were successfully accounted for in the model proposed by N. Cabibbo [13] in 1963. In the next year, M. Gell-Mann interpreted the Cabibbo theory in the context of quark model [14]. He proposed that the d and s quark states participating in a weak interaction are not necessarily pure flavor eigenstates and they are "rotated" by a mixing angle, called the Cabibbo angle θ_c. Then, instead of d, mixture of d and s quark couples with u quark (note that only these three quarks were known at that moment), which is written in a from of doublet;

$$\begin{pmatrix} u \\ d' \end{pmatrix} = \begin{pmatrix} u \\ d\cos\theta_c + s\sin\theta_c \end{pmatrix}. \tag{2.19}$$

The weak coupling G is then exactly the same for all weak interactions. For a $\Delta S = 0$ decay, involving d and u quarks, the decay rate will be proportional to $G^2 \cos^2\theta_c$, while for a $\Delta S = 1$ transition involving s and u, it will be proportional to $G^2 \sin^2\theta_c$, where the experiments give $\sin\theta_c \simeq 0.22$ ($\cos\theta_c \simeq 0.98$). The Cabibbo theory could explain all (known) weak interactions with a single coupling constant, but still there were phenomena which could not be explained by this theory alone.

All neutral interactions (or *currents*) observed before 1973 were characterized by the selection rule $\Delta S = 0$. If strangeness could be changed by a neutral current, then the decays $K^0 \to \mu^+\mu^-$ and $K^+ \to \pi^+ e^+ e^-$ would be possible. However, very stringent limits existed on these decays and others requiring strangeness-changing neutral weak currents. These limits were so restrictive that even second order weak processes would violate them in the Cabibbo scheme of weak interactions. In 1970, Glashow, Illiopoulos, and Maiani (GIM) [15] showed that if in addition to the charged weak current changing an s quark into a u quark, there were another current changing an s quark into fourth quark, there would be a cancellation. This leads to the required suppression of the strangeness-changing neutral weak currents.

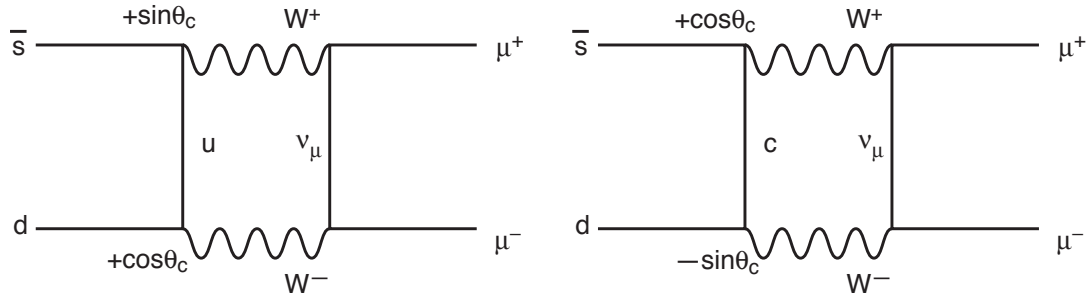

Figure 2.2: Two contributing diagrams to the decay $K_L^0 \to \mu^+\mu^-$. The second diagram contribution cancels most of the first.

Consider, for example, the decay $K_L^0 \to \mu^+\mu^-$ for which the rate was known to be extremely small. The decay can proceed through the diagrams shown in the Figure 2.2. Aside from other factors, the first diagram is proportional to $+\sin\theta_c \cos\theta_c$. The proposal of GIM was to add fourth c (stands for 'charm') quark which has charge $+\frac{2}{3}$ and form a doublet with the s, d combination orthogonal to d' in (2.19):

$$\begin{pmatrix} c \\ s' \end{pmatrix} = \begin{pmatrix} c \\ s\cos\theta_c - d\sin\theta_c \end{pmatrix}. \tag{2.20}$$

Then, when the $K_L^0 \to \mu^+\mu^-$ decay rate is calculated, there is a second term in

which a c quark appears in place of the u quark (the second diagram of Figure 2.2. This amplitude has a term proportional to $-\sin\theta_c\cos\theta_c$, just cancelling the previous term.

Now, the Cabibbo angle would be simply a rotation which mixes the d and s quarks:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \tag{2.21}$$

Another example is the K^0-$\overline{K}^0$ mixing, which proceeds with second order weak interaction shown in Fig. 2.1. If there were no contribution from c quark, the mixing rate is much larger than the observed value. One can estimate the mass of c-quark by the rates of $K_L \to \mu^+\mu^-$, K^0-$\overline{K}^0$ mixing and other rare processes. Calculation by M. K. Gaillard and B. W. Lee [16] predicted the mass of the c quark to be about 1.5–2 GeV.

The charm quark, c, was discovered in 1974[7] as J/ψ state. J/ψ is a bound state of $c\overline{c}$, with mass of 3.1 GeV/c^2. Thus, GIM mechanism could not only predict the existence of the fourth quark, but also the mass of the c quark, in advance of its experimental discovery!

2.2.2 Kobayashi-Maskawa mechanism

2.2.2.1 Cabibbo-Kobayashi-Maskawa matrix

In 1973, M. Kobayashi and T. Maskawa proposed a scheme that can explain *CP* violation within the framework of the Standard Model without introducing any additional interaction [3]. They extended the framework of the quark mixing from four quarks (or two generations) as Cabibbo-GIM proposed to general N generation case. Then, the mixing matrix becomes $N \times N$ unitary matrix, $\mathbf{V}$. The number of physics parameters can be counted as follows:

- A general $N \times N$ complex matrix has $2N^2$ real parameters.
- Unitarity imposes the conditions $V_{ij}V_{kj}^* = \delta_{ij}$, which give $N(N-1)/2$ complex constraints for $i \neq k$ and N real constraints for $i = k$. Therefore, there are N^2 remaining free parameters in $\mathbf{V}$.
- The phases of the quark fields can be rotated freely, except an overall phase. This removes $2N-1$ *relative* phases from $\mathbf{V}$, leaving $(N-1)^2$ independent

[7] This is the year the author was born.

physical parameters.

- Among them, there are $\frac{1}{2}N(N-1)$ rotation angles:

$$N_{angles} = \frac{1}{2}N(N-1).$$

- Then, the number of independent phases in V is:

$$N_{phases} = (N-1)^2 - \frac{1}{2}N(N-1) = \frac{1}{2}(N-1)(N-2).$$

For two quark generations, we have one angle and no phase, so mixing matrix is real. This is Cabibbo theory explained in the previous section. If there are three quark generations, we will have three angles and a complex phase.

It can be shown that one needs complex phase in coupling constant to violate CP . If the mixing matrix is complex, then the Standard Model Hamiltonian would contain complex terms. Kobayashi and Maskawa pointed out that complex phase can be introduced if there are three (or more) quark generations, based on above discussion. The subsequent discoveries of c, b and t quarks made the KM mechanism to be incorporated into the SM.

The quark mixing matrix in the Standard Model is now called the Cabibbo-Kobayashi-Maskawa (CKM) matrix:

$$\mathrm{V} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{2.22}$$

There are many ways to express the elements of $\mathbf{V}$ in terms of three rotation angles and one phase. One of popular parametrizations is that of Wolfenstein [17], written in the form of an expansion in $\lambda \equiv \sin\theta_c \simeq 0.22$:

$$\mathrm{V} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{2.23}$$

where A, ρ and η are real parameters of order unity. Experimentally, the parameters A and λ can be determined from tree-level decays and rather well known:

$$A = 0.84 \pm 0.04, \qquad \lambda = 0.2196 \pm 0.0023 \tag{2.24}$$

while ρ and η are not determined precisely, since its determination requires the measurement of V_{ub} and V_{td} which are of order λ^3. We will return to the determination of these parameters at the end of this chapter.

2.2.2.2 Unitary triangle

The unitarity of the CKM matrix leads to the following relations:

$$V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0 \quad (2.25)$$
$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \quad (2.26)$$
$$V_{td}V_{cd}^* + V_{ts}V_{cs}^* + V_{tb}V_{cb}^* = 0 \quad (2.27)$$
$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \quad (2.28)$$
$$V_{td}V_{ud}^* + V_{ts}V_{us}^* + V_{tb}V_{ub}^* = 0 \quad (2.29)$$
$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad (2.30)$$

Since CKM matrix elements are complex, these relations imply that they form triangles in a complex plane. These triangles are often referred to as the *unitary triangles*. There are six such triangles and it can be shown that all of them possess the same area;

$$\frac{J}{2} = \frac{1}{2}\left|\mathrm{Im}(V_{ij}V_{kl}V_{il}^*V_{kj}^*)\right|, \quad i \neq k, j \neq l. \quad (2.31)$$

where i, k and j, l denote up-type and down-type quarks, respectively. It can also be shown that this quantity is invariant under phase redefinitions of the quark fields. In the Wolfenstein parametrization, we obtain

$$J \simeq A^2\lambda^6\eta, \quad (2.32)$$

which is of order 10^{-4} for $\lambda \simeq 0.22$ and $A \simeq 0.8$.

Among six unitary triangles, two triangles which represent Eq. (2.25) and Eq. (2.26) have extremely 'squashed' shapes: two sides are of order λ, the third one of order λ^5. Next two triangles, which represent Eq. (2.27) and Eq. (2.28), are also rather squashed: two sides are of order λ^2, the third one of order λ^4. The last two triangles have the sides that are all of the same order, namely λ^3, and naturally large angles. One of such triangles is shown in Fig. 2.3. This triangle, which represents Eq. 2.30, is useful from phenomenological point of view, since it contains the most-poorly known elements of the KM matrix.

It is convenient to rescale the triangle by dividing all sides by $V_{cd}V_{cb}^*$. Then the rescaled triangle has the vertices at $(0,0)$, $(1,0)$, and $(\overline{\rho},\overline{\eta})$, where $\overline{\rho}$ and $\overline{\eta}$ are related to the Wolfenstein parameters ρ and η as:

$$\overline{\rho} = \left(1 - \frac{\lambda^2}{2}\right)\rho, \quad \overline{\eta} = \left(1 - \frac{\lambda^2}{2}\right)\eta \quad (2.33)$$

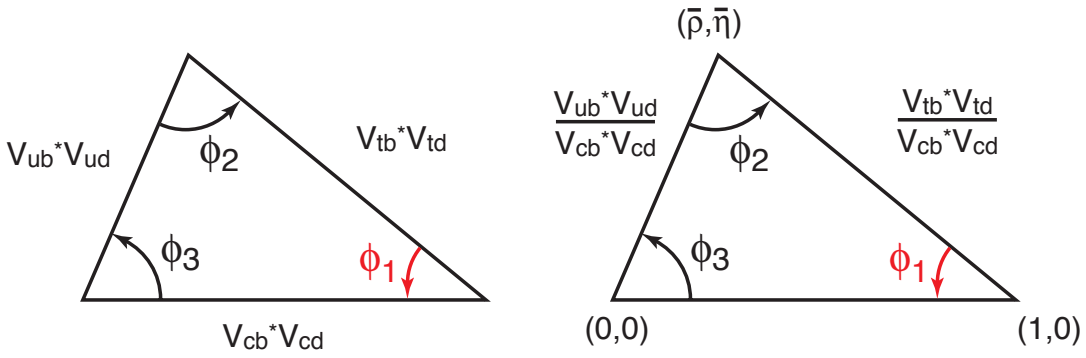

Figure 2.3: The unitary triangle with CKM matrix elements (left) and its rescaled form (right).

To test the KM mechanism, we must measure the sides and angles of the unitary triangle and check if it really forms a closed triangle or not. Especially, *CP* is violated if the area of the triangle does not vanish. This requires all the three angles of the unitary triangle shown in Fig. 2.3:

$$\begin{aligned} \phi_1 &= \pi - \arg\left(\frac{-V_{tb}V_{td}^*}{-V_{cb}V_{cd}^*}\right) \\ \phi_2 &= \arg\left(\frac{V_{tb}V_{td}^*}{-V_{ub}V_{ud}^*}\right) \\ \phi_3 &= \arg\left(\frac{V_{ub}V_{ud}^*}{-V_{cb}V_{cd}^*}\right) \end{aligned} \tag{2.34}$$

should have non-zero value. B meson system is considered to be the most suitable place to test the KM mechanism, since all the parameters appear in the unitary triangle are closely related to the dynamics of B mesons.

2.3 *CP* Violation in the Neutral B Meson System

In this section, we discuss the *CP* violation in the neutral B meson system. First, we describe the time evolution of neutral B meson. Then, we consider B meson decays into a *CP* eigenstate and possible existence of *CP* violation in the process. Finally, *CP* asymmetry in $B^0 \to J/\psi\, K^{(*)0}$ decays and its relation with KM parameters are discussed.

2.3.1 Time evolution of neutral B meson system

Let us consider the time evolution of the B^0–$\overline{B}^0$ system. Considering only $|B^0\rangle$ and $|\overline{B}^0\rangle$ states, the initial state ($|\Psi(0)\rangle$) is written as a linear combination of B^0 and $\overline{B}^0$ alone: $|\Psi(0)\rangle = a(0)|B^0\rangle + b(0)|\overline{B}^0\rangle$. Similarly to K^0–$\overline{K}^0$ system, B^0–$\overline{B}^0$ mixing occurs mainly via box diagrams shown in Fig. 2.4. We can write the time evolution of this system as

$$i\hbar\frac{\partial}{\partial t}\Psi(t) = \mathcal{H}\Psi(t), \tag{2.35}$$

where $\Psi(t)$ is written in B^0–$\overline{B}^0$ space as

$$\Psi(t) = \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}. \tag{2.36}$$

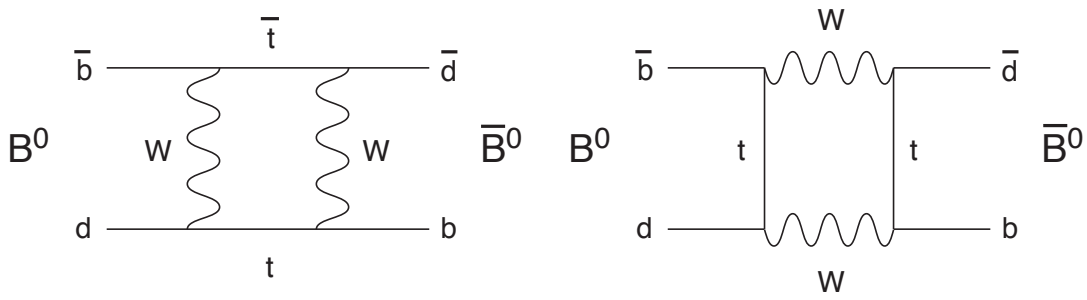

Figure 2.4: Box diagrams responsible for B^0–$\overline{B}^0$ mixing.

The matrix $\mathcal{H}$ is written as

$$\mathcal{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}, \tag{2.37}$$

where $\mathbf{M}$ and $\mathbf{\Gamma}$, known as mass and decay matrices, describe the dispersive and absorptive parts of B^0–$\overline{B}^0$ mixing. These matrices are hermitian, and CPT invariance requires $M_{11} = M_{22} \equiv M$ and $\Gamma_{11} = \Gamma_{22} \equiv \Gamma$, where M and Γ are the mass and decay width of the flavor eigenstates, B^0 and $\overline{B}^0$.

The two mass eigenstates, B_1 and B_2 can be written as

$$|B_1\rangle = p|B^0\rangle + q|\overline{B}^0\rangle, \tag{2.38}$$

$$|B_2\rangle = p|B^0\rangle - q|\overline{B}^0\rangle, \tag{2.39}$$

where

$$\frac{q}{p} = +\sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{2.40}$$

The time evolution of the mass eigenstates is expressed as

$$|B_1(t)\rangle = e^{-i(M_1 - \frac{i}{2}\Gamma_1)t}|B_1(0)\rangle \tag{2.41}$$

$$|B_2(t)\rangle = e^{-i(M_2 - \frac{i}{2}\Gamma_2)t}|B_2(0)\rangle. \tag{2.42}$$

We obtain the time evolution of initially pure $|B^0\rangle$ state by inverting Eq. (2.38), (2.39) and substituting Eq. (2.41), (2.42):

$$\begin{aligned}
|B^0(t)\rangle &= \frac{1}{2p}(|B_1(t)\rangle + |B_2(t)\rangle) \\
&= \frac{1}{2p}\left(e^{-i(M_1 - \frac{i}{2}\Gamma_1)t}|B_1(0)\rangle + e^{-i(M_2 - \frac{i}{2}\Gamma_2)t}|B_2(0)\rangle\right) \\
&= \frac{1}{2p}\left[e^{-i(M_1 - \frac{i}{2}\Gamma_1)t}(p|B^0\rangle + q|\overline{B}^0\rangle) + e^{-i(M_2 - \frac{i}{2}\Gamma_2)t}(p|B^0\rangle - q|\overline{B}^0\rangle)\right] \\
&= f_+(t)|B^0\rangle + \frac{q}{p}f_-(t)|\overline{B}^0\rangle,
\end{aligned} \tag{2.43}$$

and likewise for $\overline{B}^0$:

$$\begin{aligned}
|\overline{B}^0(t)\rangle &= \frac{1}{2q}(|B_1(t)\rangle - |B_2(t)\rangle) \\
&= \frac{1}{2q}\left(e^{-i(M_1 - \frac{i}{2}\Gamma_1)t}|B_1(0)\rangle - e^{-i(M_2 - \frac{i}{2}\Gamma_2)t}|B_2(0)\rangle\right) \\
&= \frac{p}{q}f_-(t)|B^0\rangle + f_+(t)|\overline{B}^0\rangle,
\end{aligned} \tag{2.44}$$

where

$$\begin{aligned}
f_\pm(t) &= \frac{1}{2}\left[e^{-i(M_1 - \frac{i}{2}\Gamma_1)t} \pm e^{-i(M_2 - \frac{i}{2}\Gamma_2)t}\right] \\
&= \frac{1}{2}e^{-i(M_1 - \frac{i}{2}\Gamma_1)t}\left[1 \pm e^{-i(\Delta m + \frac{i}{2}\Delta\Gamma)t}\right]
\end{aligned} \tag{2.45}$$

and $\Delta m \equiv M_2 - M_1$ and $\Delta\Gamma \equiv \Gamma_1 - \Gamma_2$ are mass and decay width difference of two mass eigenstates.

Denoting by $A(f)$ and $\overline{A}(f)$ the amplitude for the decay of B^0 and $\overline{B}^0$, respectively, into a final state f, and by $\overline{\rho}(f)$ and $\rho(f)$ their ratios, *i.e.*

$$A(f) = \langle f|H|B^0\rangle, \quad \overline{A}(f) = \langle f|H|\overline{B}^0\rangle,$$

$$\overline{\rho}(f) = \frac{\overline{A}(f)}{A(f)} = \frac{1}{\rho(f)}, \tag{2.47}$$

the time-dependent decay rates are given by

$$\begin{aligned}\Gamma(B^0(t)\to f) &\propto e^{-\Gamma_1 t}|A(f)|^2 \\ &\times \left[K_+(t)+K_-(t)\left|\frac{q}{p}\right|^2|\overline{\rho}(f)|^2+2\mathrm{Re}\left[L^*(t)\left(\frac{q}{p}\right)\overline{\rho}(f)\right]\right] \quad (2.48)\end{aligned}$$

$$\begin{aligned}\Gamma(\overline{B}^0(t)\to f) &\propto e^{-\Gamma_1 t}|\overline{A}(f)|^2 \\ &\times \left[K_+(t)+K_-(t)\left|\frac{p}{q}\right|^2|\rho(f)|^2+2\mathrm{Re}\left[L^*(t)\left(\frac{p}{q}\right)\rho(f)\right]\right] \quad (2.49)\end{aligned}$$

where

$$\begin{aligned}|f_\pm(t)|^2 &= \frac{1}{4}e^{-\Gamma_1 t}K_\pm(t), \\ f_-(t)f_+^*(t) &= \frac{1}{4}e^{-\Gamma_1 t}L^*(t), \\ K_\pm(t) &= 1+e^{\Delta\Gamma t}\pm 2e^{\frac{1}{2}\Delta\Gamma t}\cos\Delta m t, \\ L^*(t) &= 1-e^{\Delta\Gamma t}+2ie^{\frac{1}{2}\Delta\Gamma t}\sin\Delta m t. \quad (2.53)\end{aligned}$$

2.3.2 *B* meson decaying into a *CP* eigenstate

In the B_d^0 meson system, $\Delta\Gamma/\Gamma$ is much smaller than unity, because the difference is produced by common decay channels to B^0 and $\overline{B}^0$ with branching fractions of order 10^{-3} or less. Thus the term containing $\Delta\Gamma/\Gamma$ is usually dropped and Eq. (2.48), (2.49) are written as

$$\begin{aligned}\Gamma(B^0(t)\to f) &\propto 2e^{-\Gamma t}|A(f)|^2\left[1+\left|\frac{q}{p}\right|^2|\overline{\rho}(f)|^2\right. \\ &+ \left(1-\left|\frac{q}{p}\right|^2|\overline{\rho}(f)|^2\right)\cos\Delta m_d t \\ &+ \left.2\mathrm{Im}\left(\frac{q}{p}\overline{\rho}(f)\right)\sin\Delta m_d t\right], \quad (2.54)\end{aligned}$$

$$\begin{aligned}\Gamma(\overline{B}^0(t)\to f) &\propto 2e^{-\Gamma t}|\overline{A}(f)|^2\left[1+\left|\frac{p}{q}\right|^2|\rho(f)|^2\right. \\ &+ \left(1-\left|\frac{p}{q}\right|^2|\rho(f)|^2\right)\cos\Delta m_d t \\ &+ \left.2\mathrm{Im}\left(\frac{p}{q}\rho(f)\right)\sin\Delta m_d t\right], \quad (2.55)\end{aligned}$$

where Δm_d is the mass-difference of two mass eigenstates of the B_d^0 meson system.

Also, $|q/p| \simeq 1$ is very good approximation in B_d^0 meson system. Considering a CP eigenstate (f_{CP}) as the final state and introducing

$$\lambda_{CP} \equiv \frac{q}{p} \cdot \overline{\rho}(f_{CP}) \tag{2.56}$$

for convenience, the time-dependent asymmetry is expressed as:

$$\begin{aligned} a_{CP}(t) &\equiv \frac{\Gamma(\overline{B}^0(t) \to f_{CP}) - \Gamma(B^0(t) \to f_{CP})}{\Gamma(\overline{B}^0(t) \to f_{CP}) + \Gamma(B^0(t) \to f_{CP})} \\ &= \frac{2\mathrm{Im}(\lambda_{CP}) \sin \Delta m_d t + (|\lambda_{CP}|^2 - 1) \cos \Delta m_d t}{1 + |\lambda_{CP}|^2}. \end{aligned} \tag{2.57}$$

CP asymmetry arises if $\mathrm{Im}(\lambda_{CP}) \neq 0$ or $|\lambda_{CP}| \neq 1$. The former means the phase in B^0-$\overline{B}^0$ mixing is different from that in the decay. This type of CP violation is often called *mixing-induced CP violation*. The latter would be realized if the sizes of the decay amplitudes are different, *i.e.* $|A(f_{CP})| \neq |\overline{A}(f_{CP})|$, assuming $|q/p| = 1$ holds. This type of CP violation occurs solely in the decay, even if there were no B^0-$\overline{B}^0$ mixing, and it is often called *direct CP violation*.

Direct CP violation is realized if there are more than one Feynman diagrams describing the decay with weak decay phases. If there is only one amplitude (or more than one but with the same weak phase) contributing to the decay to the final state then $|\overline{\rho}(f)| = 1$ and Eqs. (2.54),(2.55) and (2.57) become as:

$$\Gamma(B^0(t) \to f_{CP}) = e^{-\Gamma t} |A(f_{CP})|^2 [1 - \mathrm{Im}(\lambda_{CP}) \cdot \sin \Delta m_d t], \tag{2.58}$$

$$\Gamma(\overline{B}^0(t) \to f_{CP}) = e^{-\Gamma t} |A(f_{CP})|^2 [1 + \mathrm{Im}(\lambda_{CP}) \cdot \sin \Delta m_d t], \tag{2.59}$$

$$a_{CP}(t) = \mathrm{Im}(\lambda_{CP}) \cdot \sin \Delta m_d t. \tag{2.60}$$

2.3.3 *CP* asymmetry in $B^0 \to J/\psi\, K^{(*)0}$

Among many possible B decays into CP eigenstates, $B^0 \to J/\psi\, K^{(*)0}$ modes are considered as the most important ones because:

- The decays $J/\psi \to \ell^+\ell^-$ provide a striking signature,
- They have relatively large branching fractions [18],

$$\mathrm{Br}(B^0 \to J/\psi\, K^0) = (8.9 \pm 1.2) \times 10^{-4} \tag{2.61}$$

$$\mathrm{Br}(B^0 \to J/\psi\, K^{*0}) = (15.0 \pm 1.7) \times 10^{-4} \tag{2.62}$$

- Large *CP* violation effect is expected from KM mechanism.
- Negligible theoretical uncertainty when translating the measured quantity into fundamental KM parameters.

where the last two items are explained below.

These decays are driven by $b \to c\overline{c}s$ and are dominated by tree diagram shown in the left side of Figure 2.5. The contamination from the loop diagram (which is often called *penguin* diagram), shown in the right side of Figure 2.5, is extremely small. Furthermore, the penguin diagram has up to very small corrections the same weak phase as the tree diagram. Therefore it is quite unlikely to exhibit direct *CP* violation and

$$|\overline{A}(J/\psi\, K^0)| = |A(J/\psi\, K^0)|, \quad |\overline{\rho}(J/\psi\, K^0)| = 1 \tag{2.63}$$

holds with very good accuracy.

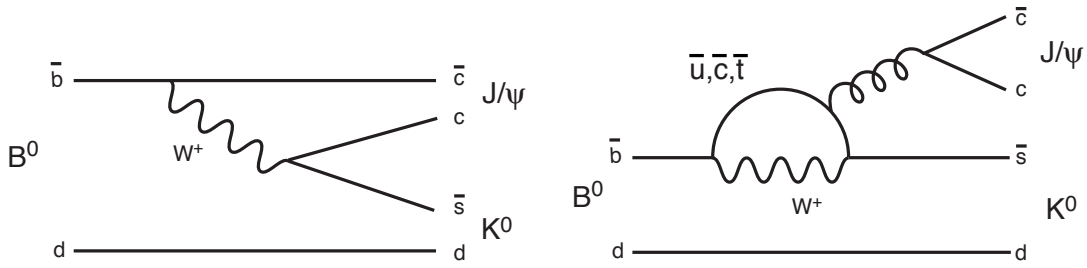

Figure 2.5: Tree (left) and penguin (right) diagrams for the decay $B^0 \to J/\psi\, K^0$.

Let us examine how the *CP* asymmetry in $J/\psi\, K^0$ mode can be related to the KM parameter. In the KM model, $\overline{\rho}(J/\psi\, K^0)$ is expressed with CKM matrix elements as

$$\overline{\rho}(J/\psi\, K^0) = \frac{\overline{A}(J/\psi\, K^0)}{A(J/\psi\, K^0)} = \frac{V_{cb}V_{cs}^*}{V_{cb}^*V_{cs}}. \tag{2.64}$$

Next we estimate q/p. With Γ_{12}/M_{12} being calculated to be $\sim 10^{-2}$,

$$\left(\frac{q}{p}\right)_{B_d} = \sqrt{\frac{M_{12}^*}{M_{12}}} + \mathcal{O}\left(\frac{\Gamma_{12}}{M_{12}}\right)\Bigg|_{B_d} \simeq \frac{V_{tb}^*V_{td}}{V_{tb}V_{td}^*}. \tag{2.65}$$

We have additional term in the case of final states containing K^0, because of K^0-$\overline{K}^0$ oscillation. This additional factor is expressed as

$$\left(\frac{q}{p}\right)_K = \frac{V_{cs}V_{cd}^*}{V_{cs}^*V_{cd}}. \tag{2.66}$$

Finally, we have an extra sign according to the *CP* eigenvalue of the final state, ξ_f. Then, we find

$$\begin{aligned}
\lambda_{J/\psi K^0} &= \xi_f \left(\frac{q}{p}\right)_{B_d} \left(\frac{q}{p}\right)_K \frac{\overline{A}(J/\psi K^0)}{A(J/\psi K^0)} \\
&\simeq \xi_f \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \cdot \frac{V_{cs} V_{cd}^*}{V_{cs}^* V_{cd}} \cdot \frac{V_{cb} V_{cs}^*}{V_{cb}^* V_{cs}} \\
&= \xi_f \frac{V_{tb}^* V_{td}}{V_{cb}^* V_{cd}} \cdot \frac{V_{cb} V_{cd}^*}{V_{tb} V_{td}^*} \\
&= \xi_f e^{-2i\phi_1}. \qquad (2.67)
\end{aligned}$$

Therefore, we obtain

$$\mathrm{Im}\lambda_{J/\psi K^0} \simeq \xi_f \sin 2\phi_1, \qquad (2.68)$$

$$a_{CP}(t) = -\xi_f \sin 2\phi_1 \cdot \sin \Delta m_d t. \qquad (2.69)$$

i.e. the asymmetry in $B^0 \to J/\psi K^0$ mode is directly related to the angle of the unitary triangle, ϕ_1.

For $B^0 \to J/\psi K_S(K_L)$, the final state has a definite *CP* eigenvalue, $\xi_f = -1(+1)$. Experimentally, $B^0 \to J/\psi K_S(K_S \to \pi^+\pi^-)$ is almost background free and called the "Golden Mode" for measuring $\sin 2\phi_1$. On the other hand, $B^0 \to J/\psi K_L$ is expected to have more background due to experimental difficulty in K_L detection. However, this mode is as important as the Golden Mode to check the consistency of the measured $\sin 2\phi_1$ with different *CP* eigenvalues of the final states.[8] Because it has *CP* eigenvalue opposite to the $J/\psi K_S$ mode, the $J/\psi K_L$ mode gives us a handle to understand the possible systematic bias on the $\sin 2\phi_1$ measurement. If there is such an effect, we should see an opposite systematic effect between *CP* even and odd modes, since their asymmetry should be equal in magnitude but opposite in sign. This is very important and useful check of our understanding of the detector response and the measurement procedure. In addition, it provides new experimental possibilities such as a comparison between $J/\psi K_S$ and $J/\psi K_L$ [19].

For $B^0 \to J/\psi K^{*0}(K^{*0} \to K_S\pi^0)$, since both J/ψ and K^{*0} are vector meson, the *CP* eigenvalue of the final state depends on the angular momentum: $\xi_f = +1$ for $l = 0, 2$ and $\xi_f = -1$ for $l = 1$. This implies that summing over l even and odd combinations will tend to wash out the total asymmetry. Fortunately, final states with different *CP* quantum numbers can be separated by measuring angular

[8] This mode must be more interesting for experimentalists than G.M. because it has more background and is more challenging!

distributions [20]. The detail of angular distribution analysis and measurement of *CP* asymmetry in $B \to J/\psi K^*$ mode is given in Ref. [21].

We conclude that $B^0 \to J/\psi K_L$ and $J/\psi K^{*0}$ modes are very promising and important to measure *CP* asymmetry, which is directly related to the parameters of KM matrix.

2.4 How to Measure the *CP* Asymmetry in the *B* Meson System

We have discussed possible *CP* violation in the B meson system. In particular, we have seen that the time-dependent asymmetry observable in the $B^0 \to J/\psi K^{(*)0}$ modes are directly and cleanly related to the angle of the unitary triangle, ϕ_1. In this section, we discuss the experimental requirements to measure this asymmetry.

2.4.1 Asymmetric *B* factory

The study of B mesons can be done in three energy regions: 1) e^+e^- collision on $\Upsilon(4S)$ resonance [9], 2) e^+e^- collision on Z^0 resonance and 3) higher center-of-mass energy in hadron colliders. Table 2.1 shows the comparison of the b-quark production cross section and the ratio between b-quark and total hadronic cross section for the above three cases.

Table 2.1: Comparison of B meson production cross section and its ratio to the total hadronic cross section.

	$\sigma_b{}^*$	$\sigma_b/\sigma_{\rm had}$
Υ(4S)	1 nb	0.25
Z^0	6 nb	0.22
$p\overline{p}$ ($\sqrt{s} = 1.8$TeV)	$\simeq$100 μb	$\simeq 10^{-3}$

* b=barn: 1 barn $= 1 \times 10^{-28}$ cm^2

Among them, experiments at the $\Upsilon(4S)$ resonance have some attracting features;

[9] $\Upsilon(4S)$ is a bound state of b and $\overline{b}$ quarks with mass of about 10.58 GeV, just above the threshold of $B\overline{B}$ production ($M_B = 5.28$ GeV/c^2).

- The ratio between b-quark production and total hadronic cross section is the highest among all energy regions.

- $\Upsilon(4S)$ decays almost 100% into $B\overline{B}$ pair. Since the mass of $\Upsilon(4S)$ is just above the threshold to produce $B\overline{B}$ pair and below $B^*\overline{B}$ or $B_s\overline{B}_s$ threshold, only $B^0_d\overline{B}^0_d$ and $B^+_uB^-_u$ pairs are produced. The ratio is measured to be [18]

$$\frac{\Upsilon(4S) \to B^+B^-}{\Upsilon(4S) \to B^0\overline{B}^0} = 1.044 \pm 0.069^{+0.043}_{-0.045} \tag{2.70}$$

  which is consistent with (and often assumed to be) 1. This assures efficient production of $B^0_d/\overline{B}^0_d$ with little contamination from other b-flavored mesons and baryons.

- Because the $\Upsilon(4S)$ decays to only two particles, the daughter B mesons have a unique momentum in the $\Upsilon(4S)$ center-of-mass frame. This can be used to greatly suppress the contamination from continuum background (*i.e.* u,d,s production).

At the Z^0 resonance, many particles other than B mesons are produced from decays of Z^0, which make the clean reconstruction of an exclusive final state and flavor identification difficult. In addition, the number of available events is limited. Four experiments at LEP recorded more than 4 million hadronic Z^0 decays per experiment during six years of LEP operation at the Z^0 energy, in which about 1 million b-quark events were included. However, we need $\simeq 10^8$ B mesons for precise study of CP violation, as we discuss later. This is two orders of magnitude larger than what LEP achieved.

The largest disadvantage of hadron machines is the poor signal to noise ratio. In order to prevent the data acquisition system from saturation, the trigger system has to be selective enough to suppress the background events by several orders of magnitudes. Therefore it is very difficult to take full advantage of the very high production cross section in hadron machines relative to e^+e^- facilities. Furthermore, due to large multiplicity in the final state, the purity of reconstructed events is much lower and identification of the b flavor is more difficult than the e^+e^- environment.

The experiment at the $\Upsilon(4S)$ resonance seems to have fascinating features. Let us discuss further the potential of the experiment. In the decay $\Upsilon(4S) \to B^0\overline{B}^0$, the B meson pair is produced into a C odd configuration. Then, both B mesons starts to oscillate between B^0 and $\overline{B}^0$ states, as we described in Section 2.3.1. However, since the $B\overline{B}$ pair is produced into a single coherent quantum mechanical state,

their oscillations are highly correlated with each other. The B meson pair can never become a B^0B^0 or $\overline{B}^0\overline{B}^0$ pair at the same time as a consequence of Bose statistics. In other words, if one of the meson is detected to be a B^0 at some time t, the other one must be $\overline{B}^0$ at *that* time.

We can utilize this feature to measure CP asymmetry of B meson system in the experiment at $\Upsilon(4S)$. We can know the flavor of B meson which decays into a CP eigenstate (we call this B B_{CP}) at the time when the other B ($B_{\rm tag}$) decays by detecting their decays into flavor-specific modes, f_{tag}. For example, once one of the B mesons is tagged as a $\overline{B}^0$ by its semi-leptonic decay $\overline{B}^0 \to \ell^- + X$ at time t, we know that its pair-produced partner has to be a B^0 at that time; its amplitude for decaying into a CP eigenstate f_{CP} at time $t + \Delta t$ is given as

$$\begin{aligned} A(f,t,\Delta t) &= e^{-\frac{1}{2}\Gamma(t+\Delta t)} e^{-i\frac{1}{2}(M_1+M_2)(t+\Delta t)} \\ &\times \left[A(f)\cos\left(\frac{1}{2}\Delta m_d \Delta t\right) + i\frac{q}{p}\overline{A}(f)\sin\left(\frac{1}{2}\Delta m_d \Delta t\right)\right]. \quad (2.71) \end{aligned}$$

It can be shown that the same expression applies when B_{CP} decays before the decay of $B_{\rm tag}$ by an amount of Δt. Then the decay rates for $(B^0\overline{B}^0)_{\Upsilon(4S)} \to f_{tag}(t_1) + J/\psi K^0(t_2)$, where one B decays into f_{tag} at the time t_1 and the other into $J/\psi K^0$ at t_2, are expressed as

$$\Gamma((B^0\overline{B}^0)_{\Upsilon(4S)} \to J/\psi K^0(t_2), B^0(t_1)) \propto e^{-\Gamma(t_1+t_2)}[1 - \xi_f \sin 2\phi_1 \sin \Delta m_d \Delta t], \quad (2.72)$$

$$\Gamma((B^0\overline{B}^0)_{\Upsilon(4S)} \to J/\psi K^0(t_2), \overline{B}^0(t_1)) \propto e^{-\Gamma(t_1+t_2)}[1 + \xi_f \sin 2\phi_1 \sin \Delta m_d \Delta t], \quad (2.73)$$

where $\Delta t \equiv t_2 - t_1$. In a real experiment, we can measure only Δt. Integrating over $t_1 + t_2$ (which ranges from $|\Delta t|$ to $+\infty$), we end up with equations in the same forms as Eqs. (2.58) and (2.59).

These time evolutions are shown in Fig. 2.6 for $J/\psi K_L$ with $\sin 2\phi_1 = +0.7$. The possible range of Δt is from $-\infty$ to $+\infty$ and it is easily seen that the CP asymmetry vanishes in the time integrated rate. In order to observe the CP asymmetry, we need to measure the decay time difference, Δt, with a precision comparable to or better than the lifetime of the B^0 meson, which is known to be about 1.5 ps. Because B mesons decay inside the beampipe and we cannot measure its decay time directly, the decay time difference is measured in terms of the decay length difference. Then, it is clear that despite advantages listed above, the $\Upsilon(4S)$ system has a weakness in CP asymmetry measurement; since the energy is just above the threshold to produce $B\overline{B}$ pairs, the daughter B mesons travel only 30 μm in the $\Upsilon(4S)$ rest frame before

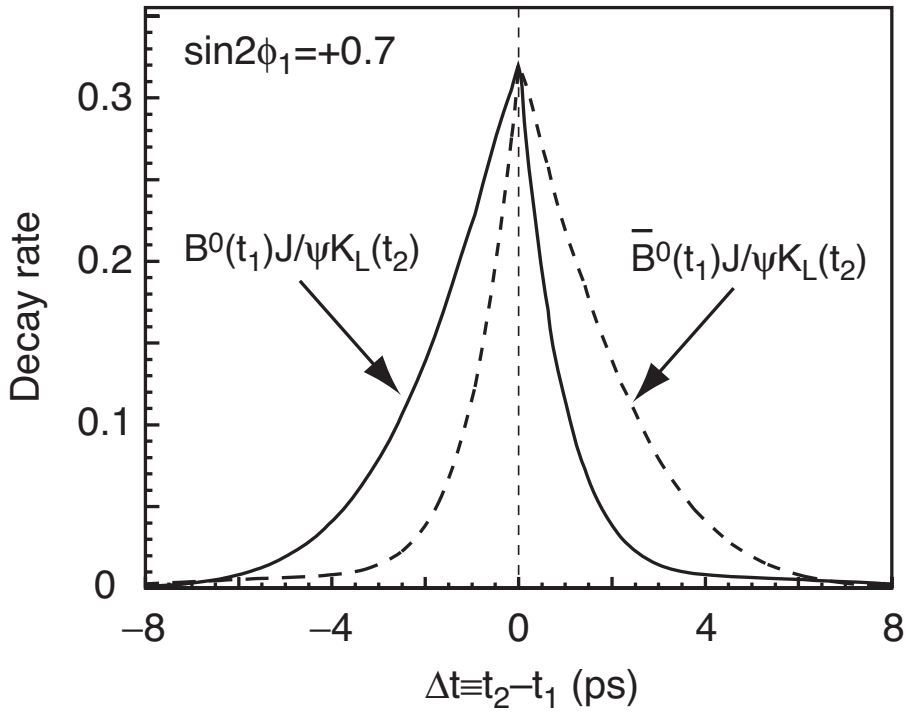

Figure 2.6: Calculated proper time distribution of $B^0(\overline{B}^0) \to J/\psi\, K_L$ with $\sin 2\phi_1 = +0.7$, $\tau_{B^0} = 1.548$ ps, and $\Delta m_d = 0.472$.

they decay! Presently available technology for the microvertex detectors does not provide us with necessary resolution.

The idea which provided the solution for this problem is to build an *asymmetric-energy* e^+e^- collider, *i.e.* a collider with different energy of e^+ and e^- beams. The asymmetric e^+e^- collider will produce a $\Upsilon(4S)$ boosted along the beam axis by a Lorentz factor:

$$(\beta\gamma)_{\Upsilon(4S)} = \frac{E_{\rm high} - E_{\rm low}}{M_{\Upsilon(4S)}} \tag{2.74}$$

which could make the decay length of the B mesons long enough to be measured with available vertex detectors.

Based on the idea described above, asymmetric colliders operating at $\Upsilon(4S)$ energy were designed and built in the last decade of 20th century.[10] Since these machines are designed to produce and detect lots of B mesons, they are called *asymmetric B-factories*.

[10]One was built at KEK (KEKB/Belle) and one at SLAC (PEP-II/*BABAR*).

2.4.2 Experimental consideration for measuring *CP* asymmetry

Now we consider the experimental requirements to measure *CP* asymmetry in $B^0 \to J/\psi K^{(*)0}$ decays at an asymmetric B-factory.

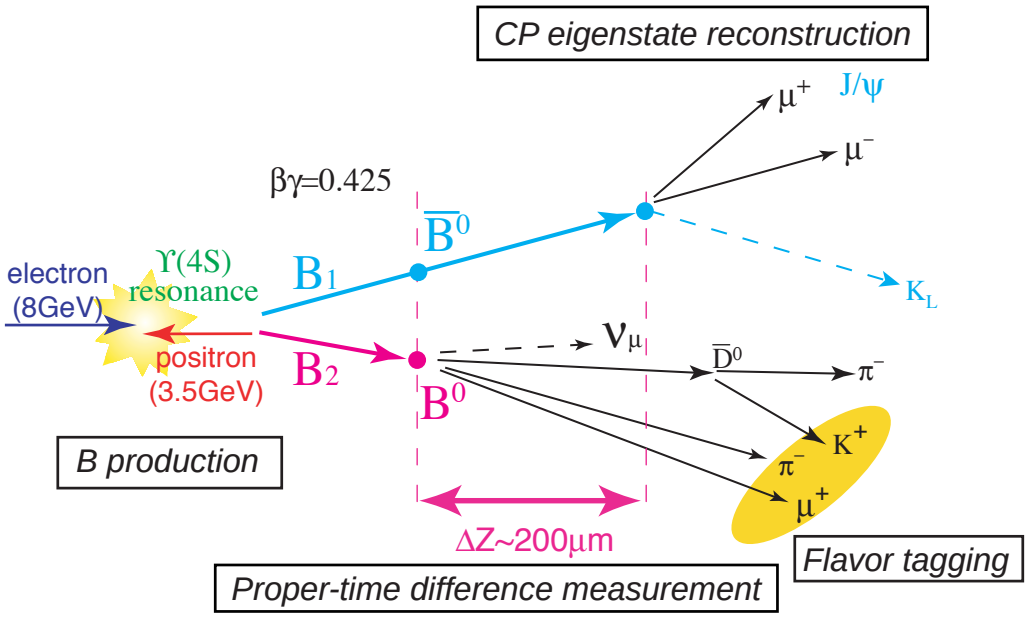

Figure 2.7: Illustration of event topology in KEK B-factory.

Figure 2.7 shows an illustrative view of the event topology at KEK B-factory. As is indicated in the figure, there are several indispensable experimental elements for the measurement of *CP* asymmetry. In this section, we briefly review the key ingredients of the experiment.

B meson production

First of all, we need to produce a large number of B mesons. Since the decay modes of our interest have branching fractions of only order of 10^{-5} or less (when we include the branching fractions to useful sub-decay modes), we have to produce many B mesons. Let us aim for 100 events, for example, of $B^0 \to J/\psi K_L$ to establish the asymmetry. J/ψ is reconstructed via $J/\psi \to \ell\ell$ ($\ell = e, \mu$). We assume the reconstruction efficiency $\epsilon_{\rm rec} = 20\%$ and flavor tagging efficiency $\epsilon_{\rm tag} = 30\%$.

Then, as a very rough estimate, we need:

$$\begin{aligned} N_B &= \frac{100}{Br(B^0 \to J/\psi\, K^0) \cdot Br(K^0 \to K_L) \cdot Br(J/\psi \to \ell^+\ell^-) \cdot \epsilon_{\rm rec} \cdot \epsilon_{\rm tag}} \\ &= \frac{100}{8.9 \times 10^{-4} \cdot 0.5 \cdot 0.12 \cdot 0.2 \cdot 0.3} \\ &\simeq 3 \times 10^7 \end{aligned} \tag{2.75}$$

B mesons. To produce such a large number of B mesons, we need very high-luminosity accelerator. At $\Upsilon(4S)$ resonance, $B\overline{B}$ production cross-section is about 1 nb and we need a luminosity of 3×10^{33} cm^{-2}s^{-1} if we want to detect CP asymmetry with one year of experiment (10^7 second). This is a luminosity ten times higher than that had achieved before the operation of new B-factories!

Reconstruction of *CP* eigenstates

High reconstruction efficiency of final states with low background is also necessary to maximize the sensitivity of the measurement. This requires efficient charged and neutral particle reconstruction with large solid-angle coverage.

For the reconstruction of $B^0 \to J/\psi\, K_L$ decays, we need to detect K_L. Because the lifetime of K_L is rather long ($\tau = 5.15\times10^{-8}$ sec), it rarely decays in the detector volume ($c\tau \simeq 15$ m). Since K_L is a neutral hadron, we cannot detect it with charged tracking detectors. Instead, we detect it as a *hadron shower*; when an incident hadron (K_L in our case) undergoes an inelastic nuclear collision with the matter, it produces secondary hadrons. The secondary hadrons then interact inelastically to produce a further hadron generation, making hadronic cascade shower. We actually do not need to measure precise energy of K_L's if the angular resolution is good enough ($\sim$ 30 mr), since it can be calculated from the kinematic constraints with the assumption of $B^0 \to J/\psi\, K_L$ decay. Such measurement can be realized with combination of an electromagnetic calorimeter and a crude hadron calorimeter (so-called "K_L-catcher"), as was first pointed out in [19].

Flavor tagging

Flavor tagging is another key for the measurement of CP asymmetry in B meson decays. As we saw, this can be done by detecting flavor-specific decays of the associated B meson.

The following information is useful to tag the flavor of the B meson:

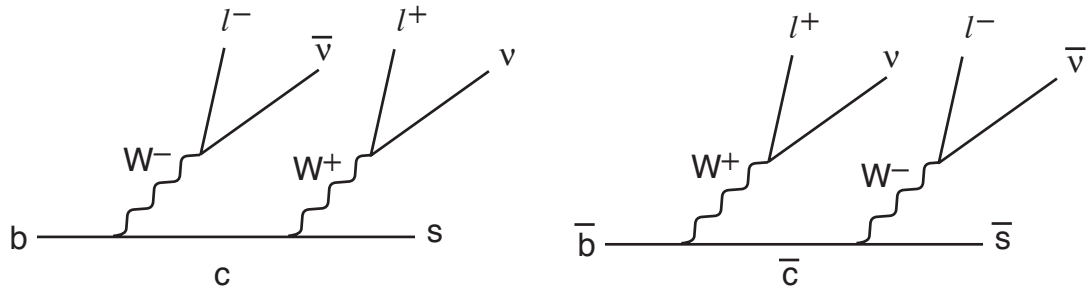

Figure 2.8: Feynman diagrams of $b(\overline{b})$-quark cascade decays useful for flavor tagging.

- The charge of leptons from semi-leptonic decays of B meson. As seen from Figure 2.8, the charge of leptons are correlated with the flavor of the primary b quarks;
- The charge of kaons from cascade decay $b \to c \to s$. Also seen from Figure 2.8, the flavor of kaons are correlated with the flavor of the primary b quarks;
- The charge of high momentum pions from $B^0 \to D^{(*)-} + (\pi^+, \rho^+, a_1^+, \text{etc.})$;
- The charge of low momentum pions from $B^0 \to D^{*-}X, D^{*-} \to \overline{D}^0\pi^-$ or $\overline{B}^0 \to D^{*+}X$, $D^{*+} \to D^0\pi^+$. Due to the fact that the mass of D^{*+} (2010 MeV) is very close to the sum of the mass of D^0 (1864 MeV) and π^+ (140 MeV), π^+ from D^{*+} has very low momentum and thus can be used to identify the existence of D^{*+}. The pions from the decay $D^* \to D\pi$ is customarily called *slow-pions* because of their soft momentum. The charge of slow-pion is correlated with the charge of D^* and consequently the flavor of the primary B meson, and thus used for flavor tagging.

Flavor tagging requires identification of particle species (e, μ and K) and therefore good particle identification capability is crucial.

Decay time difference measurement

As is mentioned in the last subsection, we have to measure the proper-time difference of two B mesons, Δt, with sufficiently good resolution to observe the CP asymmetry. Specifically, we need the proper time difference resolution comparable to the lifetime of B^0 or better. In the asymmetric B-factory, B mesons are produced with boost along the beam direction (which is usually chosen as z axis). Its Lorentz boost factor

is given in Eq. (2.74). Because B mesons have only about 330 MeV/c of momentum in the rest frame of $\Upsilon(4S)$, the separation of two B mesons from $\Upsilon(4S)$ is mostly in z (or beam) direction. The proper-time difference is approximately calculated as

$$\Delta t \simeq \frac{\Delta z}{(\beta\gamma)_{\Upsilon(4S)} \cdot c} \tag{2.76}$$

where Δz is defined as the difference of z position of the vertex of B_{CP} (z_{CP}) and that of $B_{\rm tag}$ ($z_{\rm tag}$); $\Delta z = z_{CP} - z_{\rm tag}$.

For the case of collision of 8 GeV e^- on 3.5 GeV e^+, which are beam energies at the KEKB, $(\beta\gamma)_{\Upsilon(4S)} = 0.425$ and the average Δz is about 200 μm for $\tau_{B^0} \simeq$ 1.5 ps. Therefore we need the decay vertex resolution $\sigma_{\Delta z} \lesssim 200$ μm to measure CP asymmetry, which can be achieved with (well designed and operated) microvertex detector.

Let us consider how the finite experimental precision affects the measurement of CP asymmetry. The asymmetry observed in an experiment, $A_{\rm obs}$ differs from the true asymmetry because of the presence of background events in the sample of reconstructed CP eigenstates, incorrect flavor assignment, and the finite proper-time resolution.

We can define the probability of incorrect flavor assignment (called wrong-tagging fraction), w. Then, the observed asymmetry is reduced by $(1 - 2w)$, because we have correctly-tagged events with a fraction of $(1 - w)$ and wrongly-tagged events, which have opposite sign of asymmetry, with w.

The observed asymmetry $A_{\rm obs}$ is expressed as

$$A_{\rm obs} = f_{\rm sig}(1 - 2w)d \sin 2\phi_1, \tag{2.77}$$

where $f_{\rm sig} = S/(S+B)$ is the fraction of signal events in the reconstructed events and d is the dilution factor to account for the effect of the finite proper-time resolution. The number of flavor-tagged events N is given by:

$$N = N_B(\mathcal{B}/f_{\rm sig})\epsilon_{\rm rec}\epsilon_{\rm tag}, \tag{2.78}$$

where N_B is the number of produced B^0 mesons, $(\mathcal{B}/f_{\rm sig})$ is the effective branching fraction for the decay mode considered, $\epsilon_{\rm rec}$ is the CP eigenstate reconstruction efficiency, and $\epsilon_{\rm tag}$ is the overall tagging efficiency, including wrong sign tags.

The statistical significance of the measured asymmetry s is defined as:

$$s = A_{\rm obs}/\delta A_{\rm obs}, \tag{2.79}$$

where $\delta A_{\rm obs}$ is the statistical error of the measured asymmetry and is roughly proportional to $1/\sqrt{N}$. Substituting Eqs. (2.77) and (2.78), s can be expressed as:

$$\begin{aligned} s &\simeq A_{\rm obs}\sqrt{N} \\ &= \sqrt{N_B \mathcal{B} \epsilon_{\rm rec} f_{\rm sig} \epsilon_{\rm tag} (1-2w)^2} \cdot d \cdot \sin 2\phi_1 \\ &= \frac{S}{\sqrt{S+B}} \cdot \sqrt{\epsilon_{\rm tag}(1-2w)^2} \cdot d \cdot \sin 2\phi_1. \end{aligned} \tag{2.80}$$

(Note $N_B \mathcal{B} \epsilon_{\rm rec} = S$.)

The first factor, $S/\sqrt{S+B}$, gives the dilution factor due to the existence of background. It represents a *figure of merit* for *CP* eigenstate reconstruction. We have to maximize this quantity to obtain the maximum statistical sensitivity.

The second factor originates from imperfect flavor tagging. $\epsilon_{\rm tag}(1-2w)^2$ is sometimes referred to as the *effective tagging efficiency*. The number of events required to observe the asymmetry with a certain significance is proportional to the effective tagging efficiency. We need to maximize the effective tagging efficiency to maximize the sensitivity of our measurement, which can be realized by reducing w and increasing $\epsilon_{\rm tag}$.

From Equation (2.80), we can see the importance of correct understanding of three quantities; signal fraction, wrong-tag fraction, and the effect of finite proper-time resolution. If we fail to correctly estimate these quantities, it directly causes bias in the $\sin 2\phi_1$ measurement. For example, if we overestimate the wrong-tag fractions, we will overestimate the true $\sin 2\phi_1$ (and *vice versa*). Therefore, it is of great importance to correctly estimate these three quantities in the $\sin 2\phi_1$ measurement.

In summary, we need:

- High luminosity ($\mathcal{O}(10^{33}$-$10^{34})$ cm^{-2}s^{-1}), asymmetric energy collider which can produce many B mesons;
- Particle detector with
  - high reconstruction efficiency of final states of interest with reasonably low background;

– good particle identification capability for flavor tagging;

– good vertex resolution for decay time measurement;

and good understanding of the performance,

to measure *CP* asymmetry in B meson decays. These are very challenging requirements for both the accelerator and the detector.

2.5 Experimental Constraints on the Unitary Triangle

Here we review the present experimental constraints on the unitary triangle and the previous direct measurements of $\sin 2\phi_1$.

Four measurements, apart from the direct measurement of $\sin 2\phi_1$, restrict the possible range of variations of the $\overline{\rho}$ and $\overline{\eta}$ parameters:

- The relative rate of charmed and charmless b-hadron semi-leptonic decays which allows us to measure the ratio
$$\left|\frac{V_{ub}}{V_{cb}}\right| = \frac{\lambda}{1-\frac{\lambda^2}{2}}\sqrt{\overline{\rho}^2+\overline{\eta}^2} \simeq 0.09 \pm 0.02. \tag{2.81}$$

- The mass difference of two mass eigenstates of B^0_d-$\overline{B}^0_d$ system, Δm_d, which can be measured as the B^0_d-$\overline{B}^0_d$ oscillation frequency. Since the B^0_d-$\overline{B}^0_d$ oscillation occurs via box diagrams shown in Fig. 2.4, Δm_d is related to $|V_{td}|$ and expressed as
$$\Delta m_d = \frac{G_F^2 m_W^2}{6\pi^2}\eta_B B_{B_d} f_{B_d}^2 m_{B_d} S(m_t^2/m_W^2)|V_{td}V_{tb}^*|^2, \tag{2.82}$$
where η_B accounts for the QCD corrections, $S(m_t^2/m_W^2)$ is a function of the masses of top quark and W boson. The remaining factor, $B_{B_d}f_{B_d}^2$, encodes the information of non-perturbative QCD. Apart from $|V_{td}|$, the most uncertain parameter in this expression is $\sqrt{B_{B_d}}f_{B_d}$, which has a theoretical uncertainty of about 20%. Experimentally, Δm_d is measured to be
$$\Delta m_d = (0.472 \pm 0.017) \times 10^{12}\ \hbar\mathrm{s}^{-1}. \tag{2.83}$$

- One can reduce the theoretical uncertainty seen in the previous determination by taking ratio between Δm_d and the $B^0_s\overline{B}^0_s$ oscillation frequency Δm_s:
$$\frac{\Delta m_d}{\Delta m_s} = \frac{B_{B_d}f_{B_d}^2 m_{B_d}}{B_{B_s}f_{B_s}^2 m_{B_s}}\frac{|V_{td}|^2}{|V_{ts}|^2} \tag{2.84}$$

because the ratio $\sqrt{B_{B_d}} f_{B_d} / \sqrt{B_{B_s}} f_{B_s}$ is expected to be better determined than the individual quantities. Currently, only the lower bound is known for Δm_s

$$\Delta m_s > 14.9 \times 10^{12}\ \hbar \mathrm{s}^{-1} \quad (95\%\, C.L.) \tag{2.85}$$

- *CP* violation in the kaon system which is expressed by ϵ_K. This parameter constraints the $(\overline{\rho}, \overline{\eta})$ vertex of the unitary triangle on a hyperbola. Although ϵ_K is measured in good precision, the significance of the constraint is limited due to theoretical uncertainty.

Figure 2.9 shows the constraints on the vertex of unitary triangle in the $(\overline{\rho}, \overline{\eta})$ plane following from $|V_{ub}|/|V_{cb}|$, B mixing, and ϵ_K. These measurements set constraint on $\sin 2\phi_1$ and allowed region is [22]

$$0.47 < \sin 2\phi_1 < 0.89 \qquad (95\%\mathrm{C.L.}). \tag{2.86}$$

Several groups reported the results of direct measurement of $\sin 2\phi_1$. The OPAL collaboration at LEP collider first reported their result based on 24 $B^0 \to J/\psi\, K_S$ candidates with a purity of about 60%, selected from 4.4 million Z^0 decays [23]:

$$\mathrm{OPAL}: \qquad \sin 2\phi_1 = 3.2^{+1.8}_{-2.0}(stat.) \pm 0.5(syst.). \tag{2.87}$$

The CDF collaboration at Tevatron $p\bar{p}$ collider then reported the result from about 110 pb^{-1} of data collected at the center of mass energy 1.8 TeV [24]. They used about 400 $B^0 \to J/\psi\, K_S$ events with S/N ratio about 0.7, and obtained:

$$\mathrm{CDF}: \qquad \sin 2\phi_1 = 0.79^{+0.41}_{-0.44}(stat. + syst.).^{11} \tag{2.88}$$

The ALEPH collaboration at LEP also published the result using 23 $B^0 \to J/\psi\, K_S$ candidates with an estimated purity of 71% [26]:

$$\mathrm{ALEPH}: \qquad \sin 2\phi_1 = 0.84^{+0.82}_{-1.04}(stat.) \pm 0.16(syst.). \tag{2.89}$$

Those measurements were, although proving the feasibility of the $\sin 2\phi_1$ measurement (even in $p\overline{p}$ environment), far from over-constraining the unitary triangle and testing the validity of the KM mechanism. B-factories have been expected to provide precision measurements of $\sin 2\phi_1$, which are being realized as we describe in the rest of this thesis.

[11] Recently, the CDF collaboration has reported the updated result based on the same data set. Their latest preliminary value is $\sin 2\phi_1 = 0.91^{+0.37}_{-0.36}$ [25].

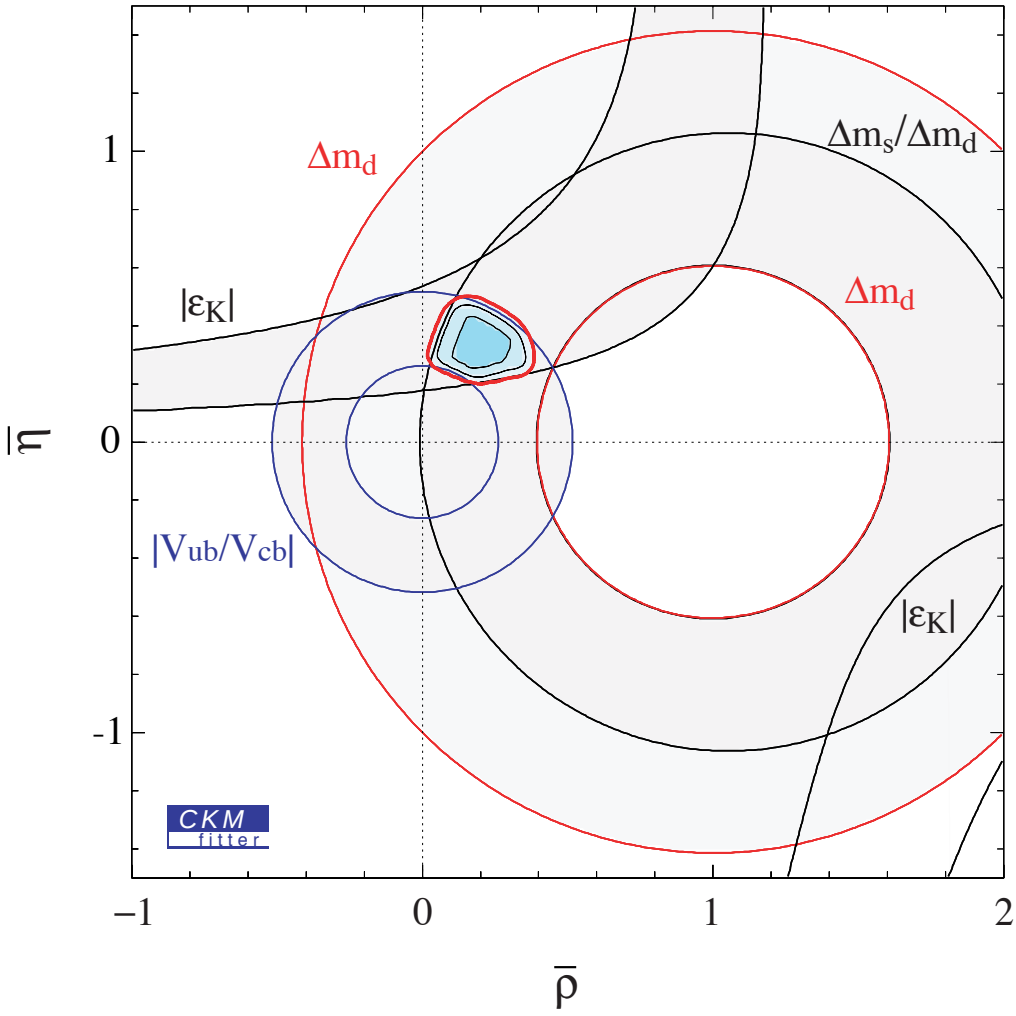

Figure 2.9: Constraints on $(\overline{\rho}, \overline{\eta})$ following from $|V_{ub}|/V_{cb}|$, B mixing, and ϵ_K. The shaded areas indicate the regions of 10%, 68%, and 95% C.L., respectively. Also shown are 95% C.L. contours of the individual constraints.

Chapter 3

Experimental Apparatus

This chapter describes the experimental apparatus used in the analysis. The B mesons used in this analysis are produced in electron-positron collisions at center-of-mass energy of 10.58 GeV (on the Υ(4S) resonance). The electron and positron beams are produced by the KEKB accelerator, which is described briefly in Section 3.1. The decay products of the B mesons are detected by the Belle detector, which is described briefly in Section 3.2 and in more detail in Reference [27, 28].

3.1 The KEKB Accelerator

KEKB [29] is an asymmetric energy e^+e^- collider of about 3 km circumference, which consists of two storage rings for 8 GeV e^- and 3.5 GeV e^+ and an injection linear accelerator for them. The center-of-mass energy is 10.58 GeV, which corresponds to the mass of the $\Upsilon(4S)$ resonance. The Lorentz boost factor $\beta\gamma \simeq 0.425$ corresponds to the average distance of the two B meson decay vertices of approximately 200 μm.

Figure 3.1 shows a schematic drawing of the KEKB accelerator. The KEKB has one interaction point where e^+ and e^- collide with a finite crossing angle of ± 11 mrad. The design luminosity is $10 \times 10^{33}\mathrm{cm}^{-2}\mathrm{sec}^{-1}$. The KEKB has achieved a peak luminosity[1] of $4.5 \times 10^{33}\mathrm{cm}^{-2}\mathrm{sec}^{-1}$. The KEKB delivered about 32.9 $\mathrm{fb}^{-1}$of data to Belle in about 2 years of operation.

[1] As of July 2001. This was the world record at that time.

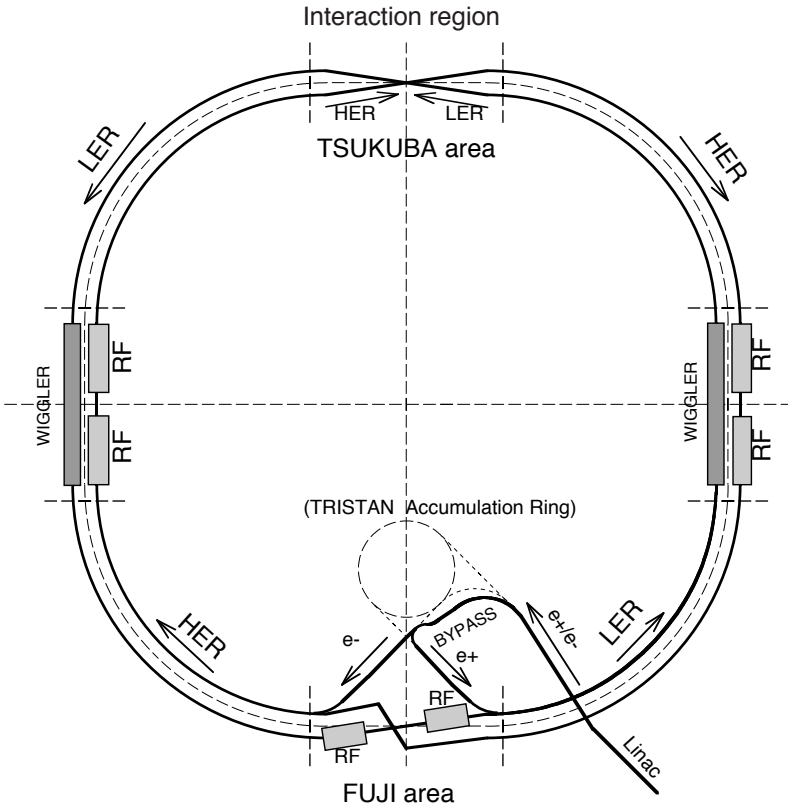

Figure 3.1: Schematic drawing of KEKB beamline

3.2 The Belle Detector

The Belle detector [28] is configured around a 1.5 T superconducting solenoid and iron structure surrounding the KEKB beams at the interaction region. Figures 3.2 and 3.3 show the configuration of the Belle detector. The Belle detector is designed to detect, identify and measure many particles, both charged and neutral, from e^+e^- interactions by using several types of sub-detectors.

B meson decay vertices are measured by a silicon vertex detector (SVD) situated just outside of a cylindrical beryllium beam pipe. Charged particle ($e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$, and $p/\bar{p}$) tracking is performed by the central drift chamber (CDC) with 1.5 T magnetic field provided by a superconducting solenoid. Particle identification is provided by dE/dx measurement in CDC, aerogel threshold Čerenkov counter (ACC) and time-of-flight counter (TOF) placed radially outside of CDC. Electromagnetic showers, which are used to identify γ's and electrons, are detected in an array of CsI(Tl) crystals (ECL) located inside the solenoid coil. Muons and K_L mesons are

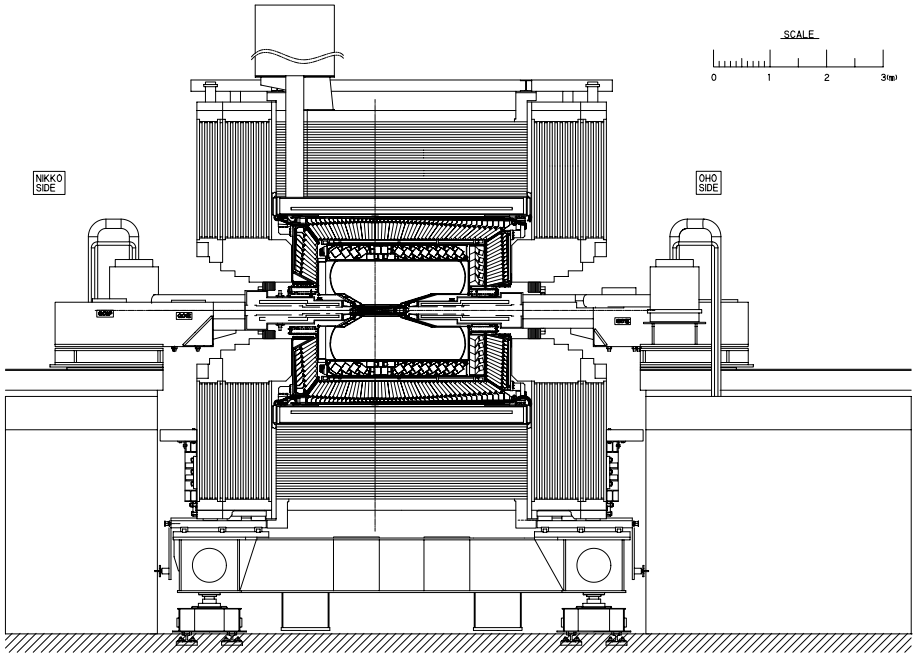

Figure 3.2: Side view of the Belle detector.

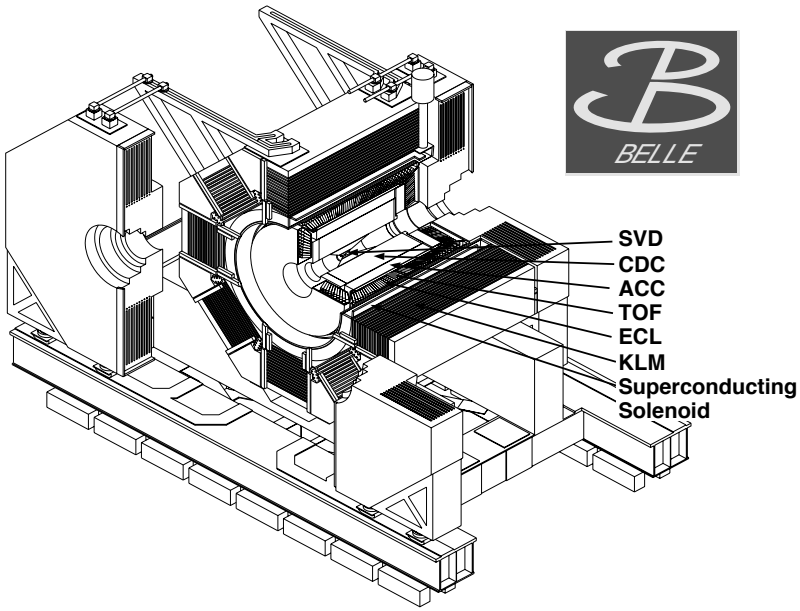

Figure 3.3: Configuration of Belle detector system.

identified by arrays of resistive plate counters interspersed in iron yoke (KLM). The detector covers the θ region extending from 17° to 150°. The part of the uncovered small-angle region is instrumented with a pair of BGO crystal arrays (EFC) placed on the surfaces of the cryostats in the forward and backward directions. The expected (or achieved) performance of the detector is summarized in Table 3.1. The design and performance of the major detector subsystems are described in the rest of this section.

The Belle coordinate system is defined as follows:

x	horizontal, inward to the KEKB ring
y	vertical, upward
z	opposite of the positron beam direction
r	$\sqrt{x^2+y^2}$
θ	the polar angle from z axis
ϕ	the azimuth angle around z axis

3.2.1 Silicon Vertexing Detector (SVD)

As described in the previous chapter, the difference in B meson decay times, which is necessary to measure CP asymmetry in B decays, can be measured using the difference in the decay vertex positions. The main role of the Silicon Vertexing Detector (SVD) [30] is to provide precise measurements of the decay vertices of B mesons. As we discussed in Section 2.4.2, the required precision of the distance between the two B decay vertices is $\lesssim 200\mu$m.

Since most particles of interest in Belle have momenta less than 1 GeV/c the vertex resolution is dominated by the multiple-Coulomb scattering. This imposes strict constraints on the design of the detector. In particular, the innermost layer of the vertex detector must be placed as close to the interaction point as possible; the support structure must be low in mass; and the readout electronics must be placed outside of the tracking volume.

The design must also withstand large beam backgrounds. With the anticipated high luminosity operation of KEKB, the radiation dose to the detector due to beam background is measured to be about 10 krad/month. Radiation dose of this level could degrade the noise performance of the electronics and induce leakage currents in the silicon detectors.

Table 3.1: Performance of detector subsystems. p and p_t in GeV/c, E in GeV.

Detector	Type	Configuration	Readout	Performance
SVD	Double sided Si-strip	300 μm thick 3-layers r = 3.0-6.05 cm	ϕ: 40.96k θ: 40.96k	$\sigma_{\Delta z} \lesssim 200$ μm
CDC	Small-cell drift chamber	Anode: 50 layers Cathode: 3 layers r = 8-88 cm	A: 8.4k C: 1.8k	$\sigma_{r\phi} = 130$ μm $\sigma_{p_t}/p_t = 0.3\%\sqrt{p_t{}^2+1}$ $\sigma_{dE/dx} = 6\%$
ACC	Threshold Čerenkov $n = 1.01 \sim 1.03$ Silica aerogel	960 barrel 228 end-cap FM-PMT readout	1788	K/π separation: $1.2 < p < 3.5$ GeV/c
TOF/TSC	Plastic scintillator	ϕ segmentation: 128/64 r = 120 cm	128×2/64	$\sigma_t = 100$ ps K/π separation: $p < 1.2$ GeV/c
ECL	CsI(Tl) crystal	Barrel: r = 125-162 cm End-cap: $z = -102/+196$ cm	B: 6624 FE: 1152 BE: 960	$\sigma_E/E = 0.066\%/E$ $\oplus 0.81\%/E^{\frac{1}{4}} \oplus 1.34\%$
KLM	Glass resistive plate counter	14 layers	ϕ: 16k θ: 16k	$\Delta\phi = \Delta\theta = 30$ mrad for K_L 1% hadron fake rate for μ
EFC	BGO crystal	Segmentation: 32 in ϕ, 5 in θ	160 × 2	$\sigma_E/E = (0.3 \sim 1)\%/\sqrt{E}$
Beam pipe	Beryllium double-wall	inner r =20 mm 0.5/2.0/0.5 mm Be/He/Be		
Magnet	Superconducting	Inner radius 170 cm		B=1.5 T

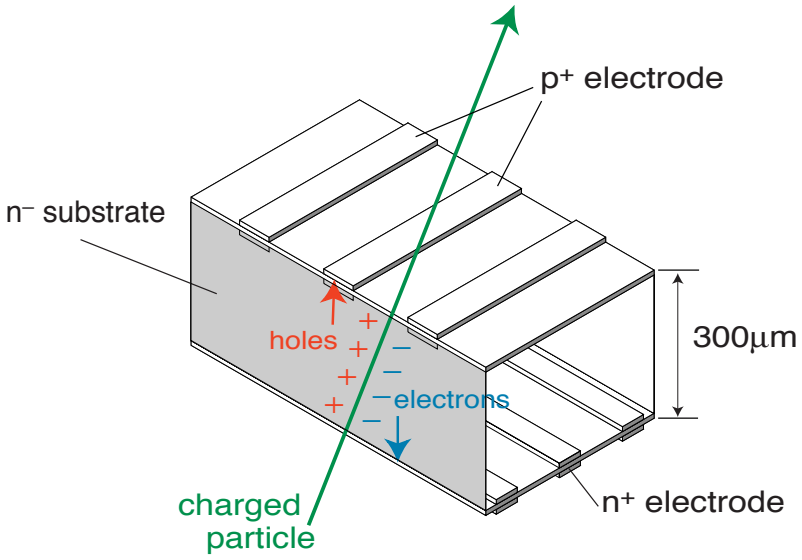

Figure 3.4: Schematic drawing of Double-sided Silicon Strip Detector(DSSD).

3.2.1.1 Silicon strip detector

Silicon strip detector is a large area p-n diode divided into narrow strips, each of which is read out by a separate electronic circuit. The detector consists of a highly doped shallow p^+ region on a low doped n^- substrate, and the backside of a highly doped n^+ layer. Usually, the reverse bias is applied to fully deplete the substrate.

Charged particles passing through the detector ionize atoms in the depletion region to produce electron-hole pairs. Generated electrons and holes are separated by the strong electric field and electrons (holes) will drift towards n^+ (p^+) electrode. The position of charged particle is given by the location of the strip carrying the signal. The signal from detector is read out as an amount of charge collected in the electrode. It is converted into voltage by a charge amplifier and sent to ADC.

The nominal thickness of a silicon strip detector is 300 μm. The energy necessary to produce an electron-hole pair is 3.6 eV in silicon. Minimum ionizing particle deposits about 80 keV of its energy in 300 μm thick Si detector and creates about 22000 electron-hole pairs. The strip pitch is as narrow as 50 μm, which enable us to measure the position of charged particle with precision of about 10 μm with analog signal readout.

Single-sided strip detectors make use of only one type of charge carrier, usually holes. By dividing the backside n^+ layer into strips and using electrons collected there, a second coordinate can be read out from the same wafer.

This is the principle of double-sided silicon strip detector (DSSD) shown in

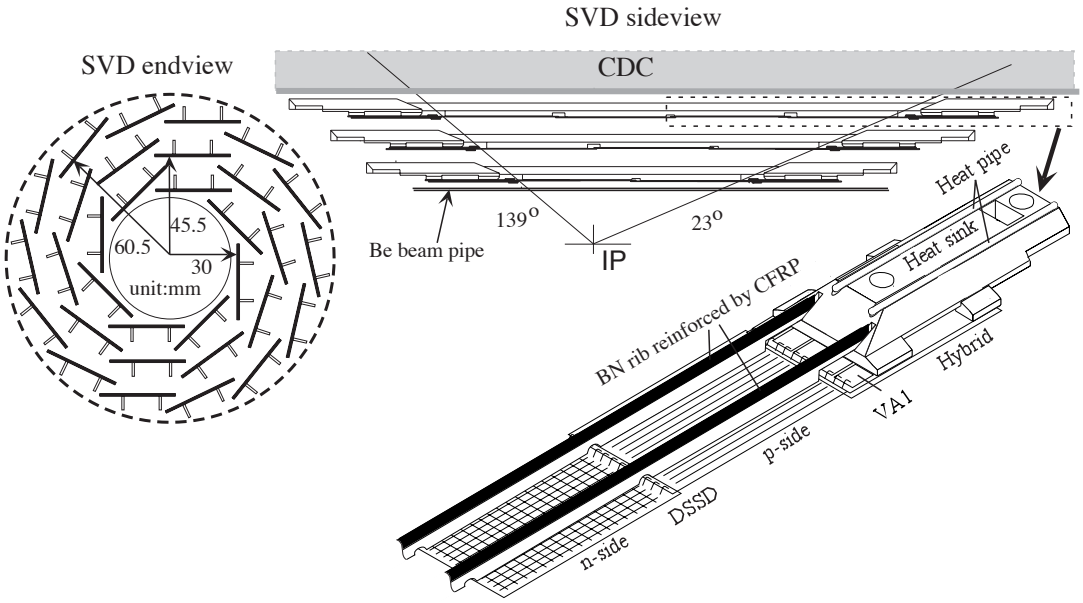

Figure 3.5: Configuration of the Belle SVD

Fig. 3.4. Since two-dimensional information can be obtained by one silicon wafer, one can minimize effects of multiple scattering by using DSSD.

3.2.1.2 Design overview

Figure 3.5 shows the geometrical configuration of the SVD. The SVD consists of three concentric cylindrical layers of silicon sensors and covers the angle range $23° < \theta < 139°$. This corresponds to 86% of the full solid angle. The radii of the three layers are 30.0 mm, 45.5 mm and 60.5 mm. The beampipe is a double-wall beryllium cylinder of 2.3 cm radius and 1 mm thick.

Each layer is constructed from independent ladders. Each ladder consists of double-sided silicon strip detectors (DSSDs) reinforced by boron-nitride support ribs. The design uses only a single type of DSSD, which reduces the cost of detector production, minimizes the amount of detector development work, and streamlines testing and bookkeeping during production. The benefit also extends to hybrid production and ladder assembly, where only a single type of hybrid is necessary and the design of the ladder assembly fixtures is greatly simplified [31].

The DSSDs are fabricated by Hamamatsu Photonics (HPK S6936), and were originally developed for the DELPHI micro-vertex detector [32]. The sensor size is $57.5 \times 33.5\text{mm}^2$. The strip pitch is 25 μm on the ϕ-side and 42 μm on the z-side.

The n-side strips are used for the z-coordinate measurement. A double-metal structure running parallel to z is employed to route the signals from orthogonal z-sense strips to the ends of the detector. Adjacent strips are connected to one readout trace on the z-side which gives an effective pitch of 84 μm. A p-stop structure is employed to isolate the sense strips. A relatively large thermal noise ($\sim 600e^-$) is observed due to the common-p-stop design.

Every other sense-strip is connected to readout electronics on the ϕ-side. Signals collected by floating strips are read out from adjacent strips by means of capacitive charge division. In total 102 DSSDs are used and the total number of readout channels is 81920.

Each ladder consists of two half-ladders that are electrically independent, but mechanically joined by support ribs. The support ribs are made of boron nitride sandwiched by carbon-fiber reinforced plastic to minimize the material while keeping enough stiffness. Total amount of material is about 0.5% of radiation length per layer.

Only two kinds of half-ladders, short and long, are used to construct all the ladders. A short half-ladder is made from a single DSSD while a long half-ladder is made of two DSSDs where strips on the p-side of one DSSD are wire-bonded to strips on the n-side of the other DSSD. This flipped design was chosen to maintain a good signal-to-noise (S/N) ratio in the half-ladders.

A ceramic-board preamplifier card (hybrid board) is mounted at one end of each half-ladder. Two hybrid boards were glued back-to-back to read both sides of DSSD. Since the preamplifier chips are a heat source, careful attention is paid to the thermal path way through the hybrid board, across glue joints, and into the heat-sink. The heat-sink is made of aluminum nitride, whose high thermal conductivity and low coefficient of thermal expansion make it an attractive option for both heat conduction and mechanical support. The thermal conductivity is further enhanced by embedding two heat-pipes in the heat-sink. Measurements indicate that the total temperature drop across the heat-sink will be 1.8 °C at normal preamplifier power levels.

Figure 3.6 shows a schematic view of the SVD readout system [33]. Signals from each side of DSSDs are read by electronics comprising VA1 front-end integrated circuits [34] mounted on ceramic hybrids. On each hybrid there are five 128-channel VA1 chips. The VA1 chip has an excellent noise performance of 200 e^- + 8 e^-/pF at a shaping time of 1.0 μs and reasonably good radiation tolerance up to a dose of 200 krad [35]. Analog signals from the VA1 chips are amplified and buffered in

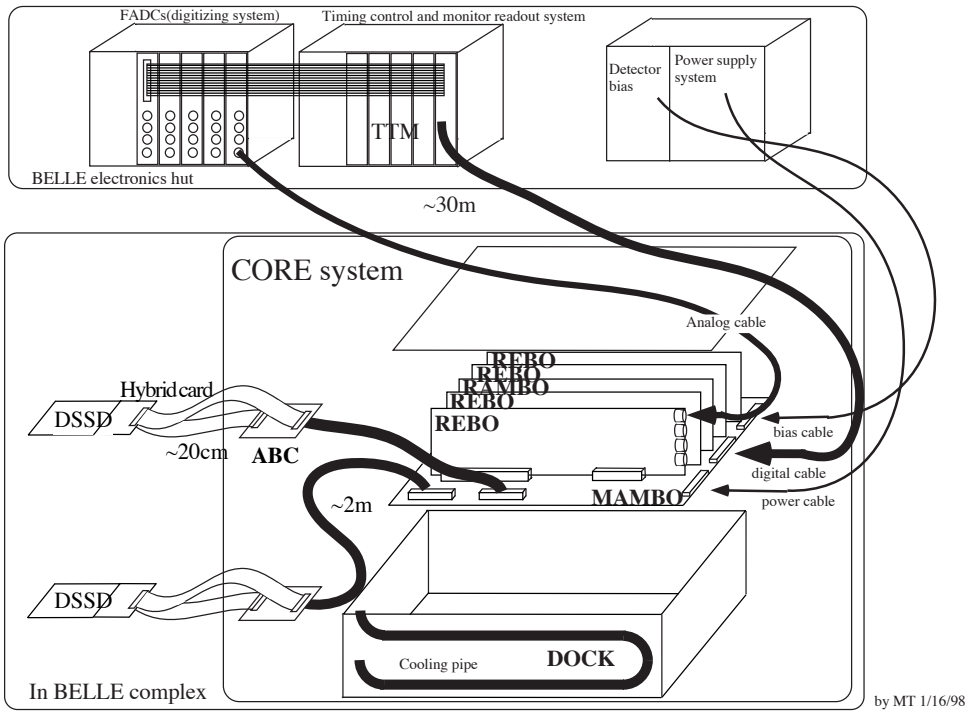

Figure 3.6: Schematic drawing of the SVD readout system

the repeater (CORE) system, which is located inside the Belle structure. The signal output from the repeater system is digitized by fast analog-to-digital converters (FADCs) located about 30 m away from the front-end electronics. Common mode noise subtraction, data sparsification and data formatting are performed online by digital signal processors in the FADC modules.

The first version of the SVD was unfortunately damaged by synchrotron radiation and replaced with a second version, with new inner layer ladders and Au-foils wrapped around the outer surface of the beam pipe, during the summer 1999 shutdown. This second version was used until summer 2000, when we installed the third version of the SVD. In the third version of the SVD, we used radiation-hard version of the VA1 chip developed by us, which is measured to be operational up to 1 Mrad [36]. The radiation hardness measurement and the development of radiation-hard frontend chip are summarized in Appendix E.

3.2.1.3 Performance

Good signal-to-noise (S/N) ratios and high strip yield are needed to ensure efficient matching between tracks detected in the central drift chamber (CDC) and hits

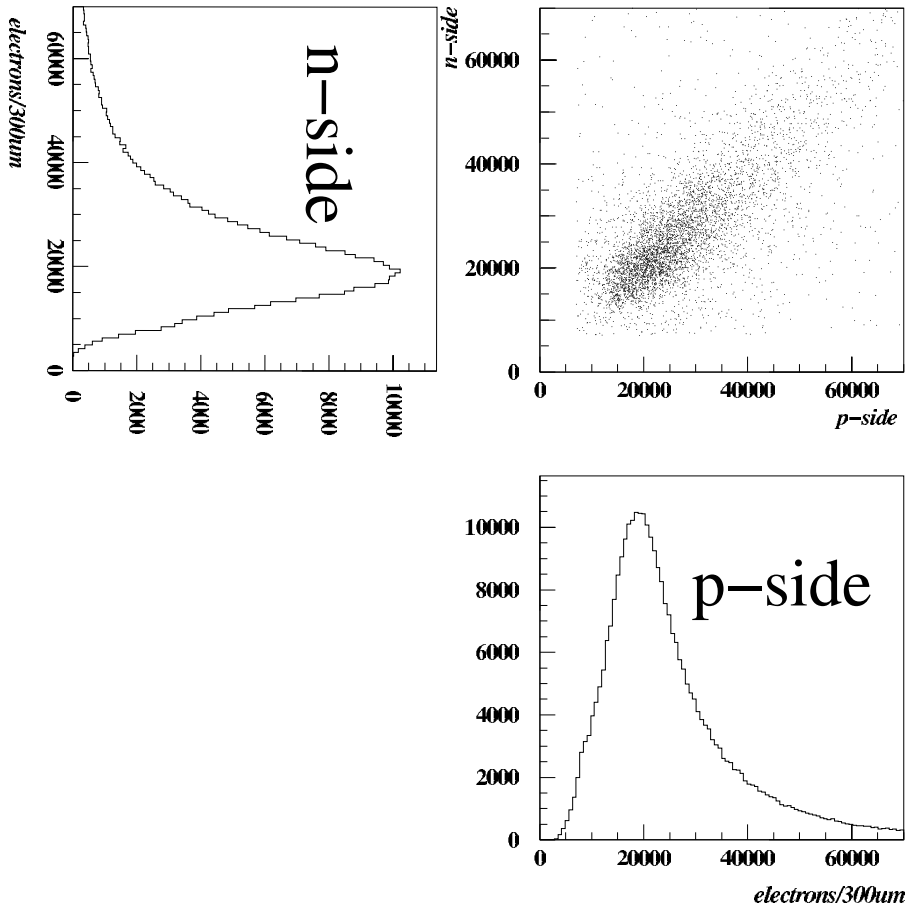

Figure 3.7: Normalized cluster energy distributions in the SVD

detected in the SVD. The normalized cluster energy distributions for minimum-ionizing particles from hadronic events are shown in Fig. 3.7, where the energy of each cluster is calculated from the sum of signals above 4σ of pedestal on each channel. The cluster energy is normalized to the same track path length in the DSSD (300μm). The measured most-probable peak height is approximately 19,000 e^-, which corresponds to the charge collection efficiency about 86%.

The noise level and the S/N ratio for each DSSD side (p, n, and $p+n$) are summarized in Table 3.2. Strip yields, which are defined to be the fraction of channels with S/N ratios larger than 10, are measured to be 98.8% on layer 1, 96.3% on layer 2, and 93.5% on layer 3.

The track-matching efficiency is defined as the probability that a CDC track within the SVD acceptance has associated SVD hits in at least two layers, and in at least one layer with both r-ϕ and r-z information. Tracks from K_S decays are

Table 3.2: Noise level and the S/N ratio for each DSSD side before irradiation.

side	capacitance(pF)	ENC (e^-)	S/N
p	7	400	47
n	22	1000	19
p+n	29	1100	17

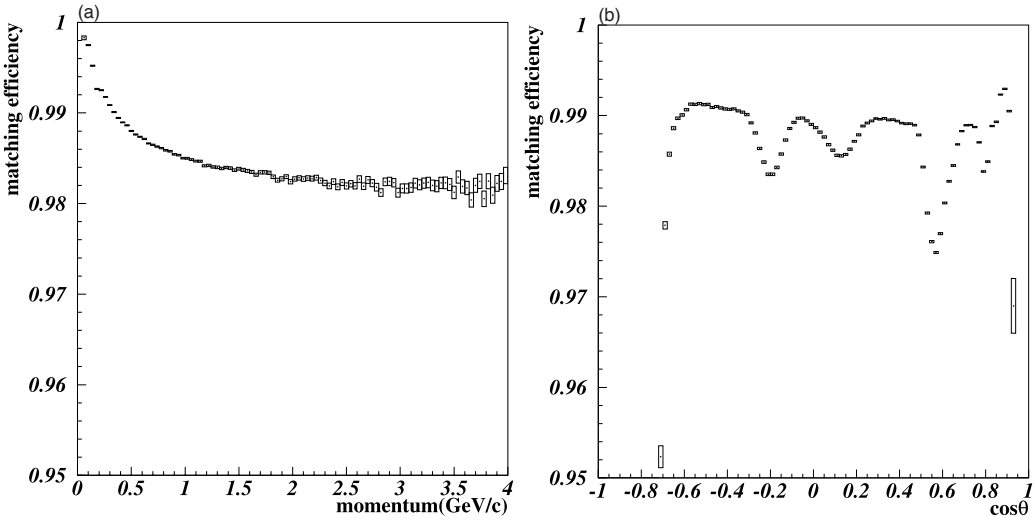

Figure 3.8: SVD-CDC Matching efficiency as a function of (a) momentum and (b) angle from the beam axis.

excluded since these tracks do not necessarily go through the SVD. Figure 3.8 (a) and (b) show the matching efficiency as a function of (a) momentum and (b) $\cos\theta$ for hadronic events. We observe several dips in the $\cos\theta$ distribution due to gaps between DSSDs in a ladder in the z direction. The average matching efficiency is found to be 98.7%, which agrees with the Monte Carlo simulation. Figure 3.9 shows the matching efficiency as a function of time. We observed a slight degradation of the matching efficiency due to the gain loss of the VA1 from radiation damage.

Precise alignment of the SVD is crucial to minimize the vertex resolution. Each DSSD is aligned in the SVD local coordinate first (internal alignment) and then the entire SVD is aligned with respect to the CDC which defines the Belle coordinate (global alignment). The detail of alignment procedure is described in Appendix A.

Good impact parameter resolution is essential for the measurement of time de-

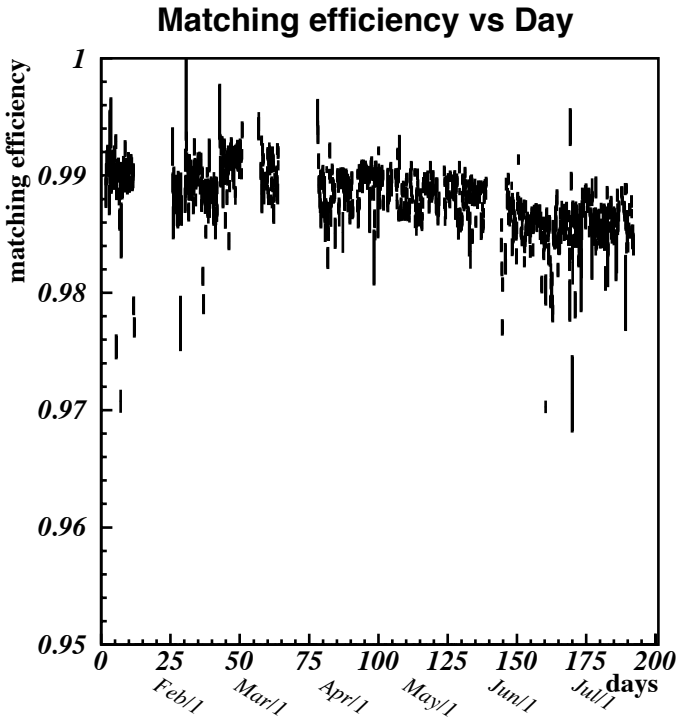

Figure 3.9: SVD-CDC Matching efficiency as a function of the date of data taking.

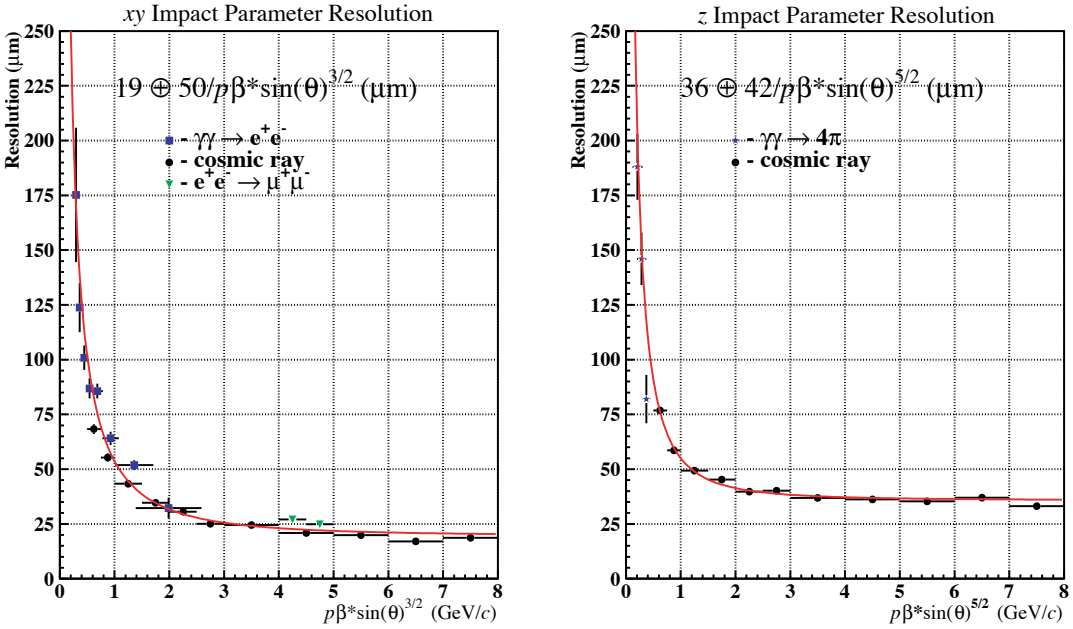

Figure 3.10: Impact parameter resolution for the r-ϕ plane and z direction.

pendent *CP* asymmetry in B meson decays. The impact parameter resolution depends on the alignment accuracy as well as the intrinsic resolution of the DSSD's. The impact parameter resolution σ_{ip} can be expressed in the following form:

$$\sigma_{ip} = a \oplus b/\tilde{p}, \tag{3.1}$$

where $\oplus$ stands for the quadratic sum, $\tilde{p}$ is *pseudo momentum* defined as

$$\tilde{p} = \begin{cases} p\beta \sin^{\frac{3}{2}}\theta & (r\phi) \\ p\beta \sin^{\frac{5}{2}}\theta & (z) \end{cases} \tag{3.2}$$

The first term of Eq. (3.1), a is determined by the SVD intrinsic resolution and the alignment accuracy. The second term, b reflects the effects of multiple Coulomb scattering in the material.

The impact parameter resolution is measured using the cosmic ray sample taken during beam collisions. The detail of the impact parameter resolution measurement is given in [37] and Appendix B. One cosmic ray is reconstructed as an incoming and an outgoing track. Since the two reconstructed tracks belong actually to the same cosmic ray, the difference of the track positions near the interaction point represents the resolution. The impact parameter resolution is calculated as the standard deviation of the difference of the two track positions divided by $\sqrt{2}$. A wide momentum range of cosmic rays can be used for study. Figure 3.10 shows the impact parameter resolutions, σ_{xy}, σ_z, for the xy plane and the z coordinate as a function of pseudo momentum.

The impact parameter resolution is also studied using tracks from collisions. Bhabha, dimuon, and $\gamma\gamma \to \rho^0\rho^0 \to \pi^+\pi^-\pi^+\pi^-$ events are used. The results obtained from these studies are also shown in Fig. 3.10. The momentum and angular dependence is well represented by the following formula:

$$\sigma_{xy} = 19 \oplus 50/(p\beta \sin^{3/2}\theta)\mu\mathrm{m}, \tag{3.3}$$

$$\sigma_z = 36 \oplus 42/(p\beta \sin^{5/2}\theta)\mu\mathrm{m}. \tag{3.4}$$

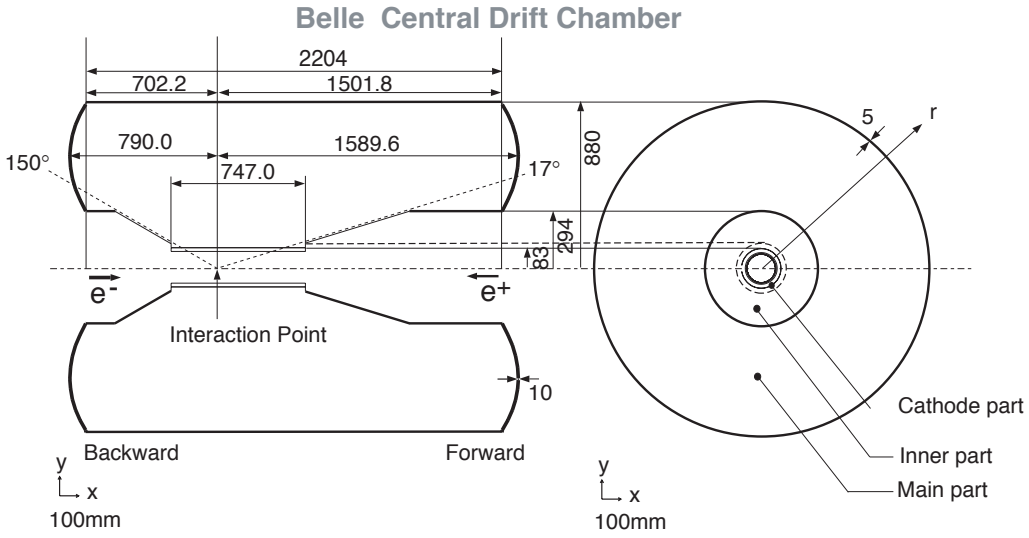

Figure 3.11: Schematic view of the CDC structure. The lengths in the figure are in units of mm

3.2.2 Central Drift Chamber (CDC)

The efficient reconstruction of charged particle tracks and precise determination of their momenta are essential ingredient to the experiment. Specifically, the physics goals of the Belle experiment require a transverse momentum (momentum component transverse to the beam axis, p_t) 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 $p_t \geq 100$ MeV/c in the polar angle region of $17^\circ \leq \theta \leq 150^\circ$. We can calculate p_t from the radius of curvature r as

$$p_t = 0.3Br \tag{3.5}$$

where p_t is in units of GeV/c, B in Tesla, and r in m.

The main purpose of the CDC [38] is to provide good momentum and position resolution for charged tracks. Since most B meson decay products have momenta below 1 GeV/c in Belle, the effects of multiple Coulomb scattering have great impact on the momentum resolution. Therefore, it is essential to reduce the amount of material in the tracking volume in order to obtain good momentum resolution. The charged particle tracking system is expected to provide an important information for the trigger system. It also provides precise energy loss (dE/dx) measurements for charged particles for the particle identification.

The structure of the CDC is shown in Fig. 3.11. It is a cylindrical chamber with inner radius 77 mm, outer radius 880 mm, and length 2400 mm and consists of fifty

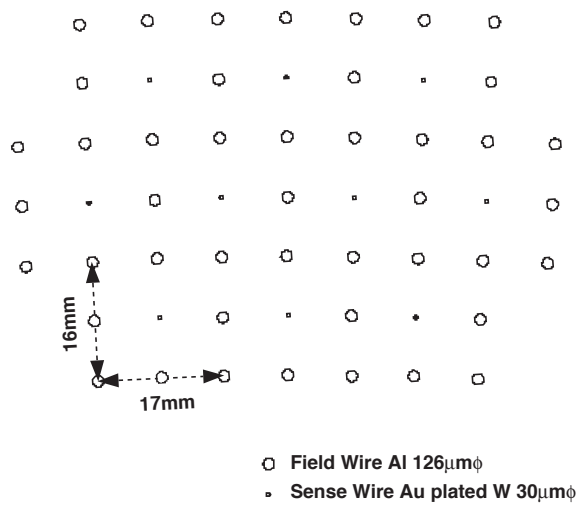

Figure 3.12: CDC cell structure.

sense wire layers and three cathode strip layers. Figure 3.12 shows the cell structure of the Belle CDC. The sense wire layers are grouped into eleven superlayers, where six of them are axial and five are small-angle stereo superlayers. Each super-layer consists of between three and six radial layers, all with the same number of drift cells in azimuthal direction. The small-angle stereo layers are used in conjunction with the axial layers to provide z coordinate measurements. Stereo layers also provide a highly-efficient fast z-trigger combined with the cathode strips. We determined the stereo angles in each stereo super-layer by maximizing the z-measurement capability while keeping the gain variations along the wire below 10%. Total number of sense wires is 8,400, of which 5,280 is axial and 3,120 is stereo. The cathode strips are divided into eight segments in ϕ direction and 64 segments (8.2 mm pitch) in z to provide z-coordinate information used for the fast trigger.

50% helium-50% ethane gas mixture is used in the Belle CDC. Low-Z gas is chosen in order to minimize multiple-Coulomb scattering contribution to the momentum resolution. The spatial resolution for tracks passing near the middle of the drift space is measured to be approximately 130 μm in r-ϕ direction.

Charged particle tracking is done by Kalman filtering method [39], taking into account the effect of multiple Coulomb scattering, energy loss, and non-uniformity of the magnetic field. The transverse momentum (p_t) resolution as a function of p_t, measured using cosmic-ray tracks during collision runs, is shown in Fig. 3.13. The upper result is for the case we use only CDC information. The lower is for the case we also use SVD hit information in the track parameter fit. The resolution is

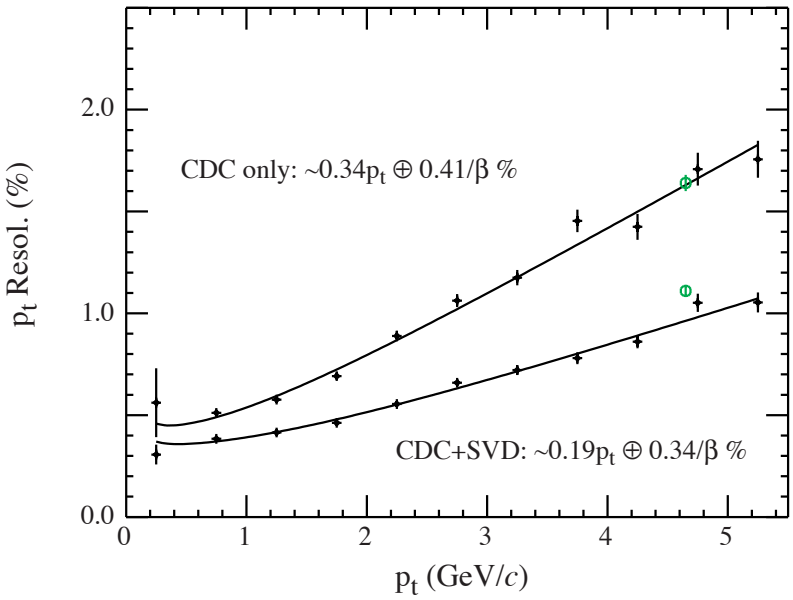

Figure 3.13: Transverse momentum(p_t) resolution of CDC as a function of p_t measured with cosmic rays. Upper is result with CDC only; lower is that with CDC+SVD.

substantially improved by including SVD. The p_t resolution is measured to be

$$\sigma_{p_t}/p_t = (0.34 \oplus 0.19 p_t)\%. \tag{3.6}$$

The mean rate of energy loss (dE/dx) of a charged particle is given by the Bethe-Bloch equation,

$$-\frac{dE}{dx} = 4\pi N_A r_e^2 m_e c^2 \frac{Z}{A}\left(\frac{z}{\beta}\right)^2\left[\ln\left(\frac{2m_e c^2\beta^2\gamma^2}{I}\right) - \beta^2 - \frac{\delta}{2}\right], \tag{3.7}$$

where N_A is Avogadro's number, r_e is the classical electron radius, m_e is the mass of electron. Z and A are the atomic number and mass number of the atoms of the medium, z and v are the charge (in units of e) and velocity of the particle, $\beta = v/c$ and $\gamma = 1/\sqrt{1-\beta^2}$, $I \simeq 16Z^{0.9}$ eV is the mean excitation energy of the medium, and x is the path length in the medium, measured in g cm^{-2}. Equation (3.7) shows that dE/dx is independent of the mass of the particle and depends on β. Therefore we can estimate β from a measurement of dE/dx. The measurement of β can provide a useful method for estimating the rest mass and thus differentiating particle species in conjunction with the momentum measurement .

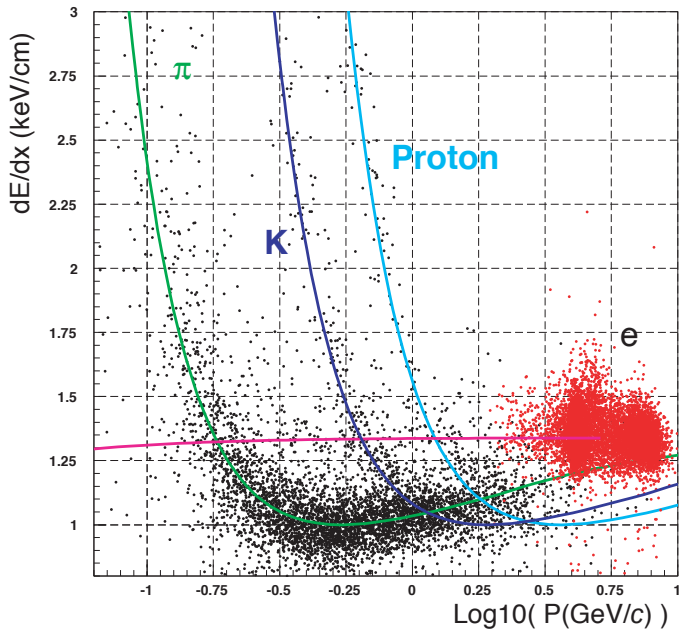

Figure 3.14: Measured dE/dx vs. momentum in collision data. The expected mean energy loss for different species are also shown.

Figure 3.14 shows the measured dE/dx as a function of momentum, together with the expected mean values for different species. Populations of pions, kaons, protons and electrons are clearly seen. The dE/dx resolution is measured to be 7.8% for pions in the momentum range from 0.4 to 0.6 GeV/c, while the resolution for Bhabha and μ-pair events is measured to be about 6%. The dE/dx information provides $\geq 3\sigma$ K/π separation up to 0.8 GeV/c. The dE/dx for kaons and pions have cross-over around 1 GeV/c, but they can provide some discrimination between kaons and pions above 2 GeV/c. It also provides more than 3σ e/π separation for the momentum range from 0.3 GeV/$c$to 3 GeV/c.

3.2.3 Aerogel Čerenkov Counter System (ACC)

The Aerogel Čerenkov Counter System (ACC) [40] provides particle identification information in the momentum range from 1.2 to 3.5 GeV/c. The Belle ACC is threshold type Čerenkov detector, which identifies particle species according to whether Čerenkov light is emitted or not in the material. Čerenkov light is emitted if the velocity of the particle β satisfies

$$n > 1/\beta = \sqrt{1 + (m/p)^2}, \tag{3.8}$$

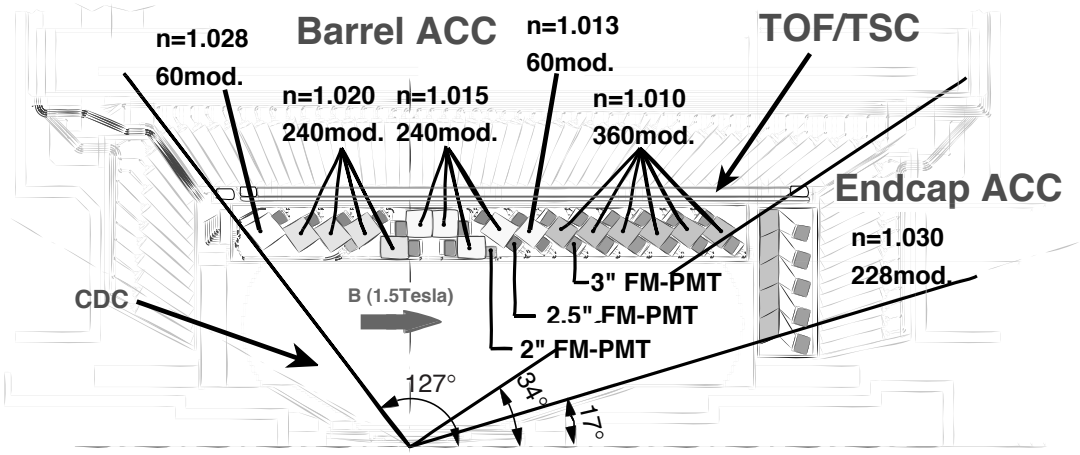

Figure 3.15: Configuration of ACC and TOF system.

where n is the refractive index of the matter, m and p is the mass and momentum of the particle, respectively. Therefore, there is momentum region where pions ($m_{\pi^\pm} \simeq 140$ MeV/c^2) emit Čerenkov light but kaons ($m_{K^\pm} \simeq 500$ MeV/c^2) do not, depending on the refractive index of the matter. For example, pions with momentum 2 GeV/c emit Čerenkov light in the matter if $n > 1.002$, while $n > 1.030$ is necessary for kaons with the same momentum.

Figure 3.15 shows the configuration of the ACC. The ACC consists of 960 counter modules segmented into 60 cells in the ϕ direction for the barrel part and 228 modules arranged in 5 concentric layers for the forward end-cap part of the detector. The possible momentum range of charged particles depends on the polar angle at Belle due to the asymmetric beam energy. In order to obtain good π/K separation for the momentum range from 1.2 to 3.5 GeV/c, the refractive indices of aerogels in the barrel region are selected to be between 1.01 and 1.028, depending on their polar angle. For the endcap ACC, we use aerogel with the refractive index 1.03 in order to cover lower momentum region (which is necessary for flavor tagging), because there is no TOF system in the endcap. Aerogel tiles are stacked in 0.2 mm-thick aluminum boxes. Each aerogel radiator module is viewed by one fine-mesh photomultiplier (FMPMT) in the endcap and two FMPMT's in the barrel. The FMPMT can be operated in a magnetic field of 1.5 T.

Figure 3.16 shows the measured pulse height distributions for barrel ACC for $e^\pm$ tracks in Bhabha events and $K^\pm$ candidates in hadronic events, together with the expectations from Monte Carlo simulation. $K^\pm$ tracks are selected by TOF and dE/dx measurements. Clear separation can be seen from the distributions.

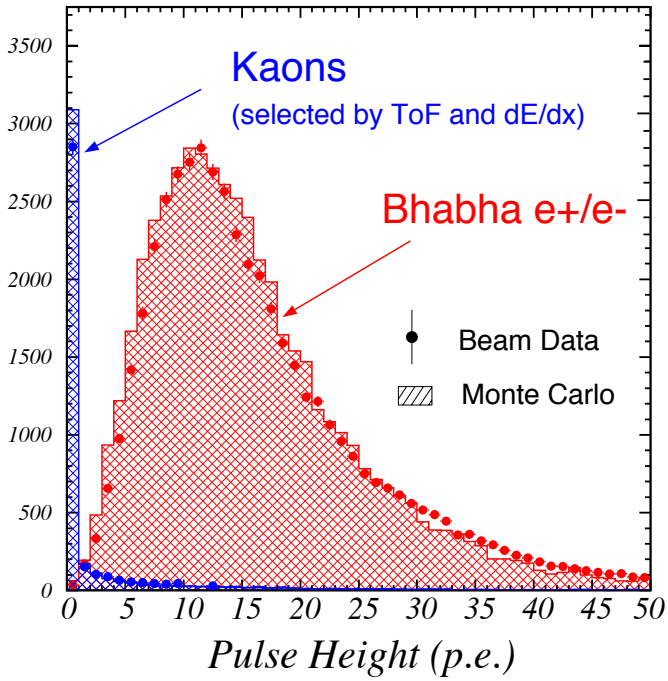

Figure 3.16: Pulse-height spectra in units of photoelectrons observed by ACC for electrons and kaons. The Monte Carlo expectations are superimposed.

3.2.4 Time-of-Flight Counters (TOF)

A time-of fight (TOF) detector system [41] provides particle identification information for particle momenta below 1.2 GeV/c. It also provides fast timing signals for the trigger system.

The TOF system consists of 128 plastic scintillation counters and 64 thin Trigger Scintillation Counters (TSC). Two trapezoidally shaped TOF counters and one TSC counter form a module. Each TOF (TSC) counter is read out by two (one) fine-mesh photomultipliers. In total 64 TOF/TSC modules located at a radius of 1.2 m from the interaction point cover a polar angle range from 33° to 121°. The configuration of the TOF is shown in Fig. 3.15.

The flight time T of particle in length L is expressed as

$$T = \frac{L}{c}\sqrt{1 + c^2(m/p)^2}. \tag{3.9}$$

Because the time-of-flight depends on the mass of particle, we can use it to distinguish particle species. For the case of particles with momentum of 1.2 GeV/c, $T = 4.0$ ns for pions and $T = 4.3$ ns for kaons with $L = 1.2$ m. Thus, time resolution

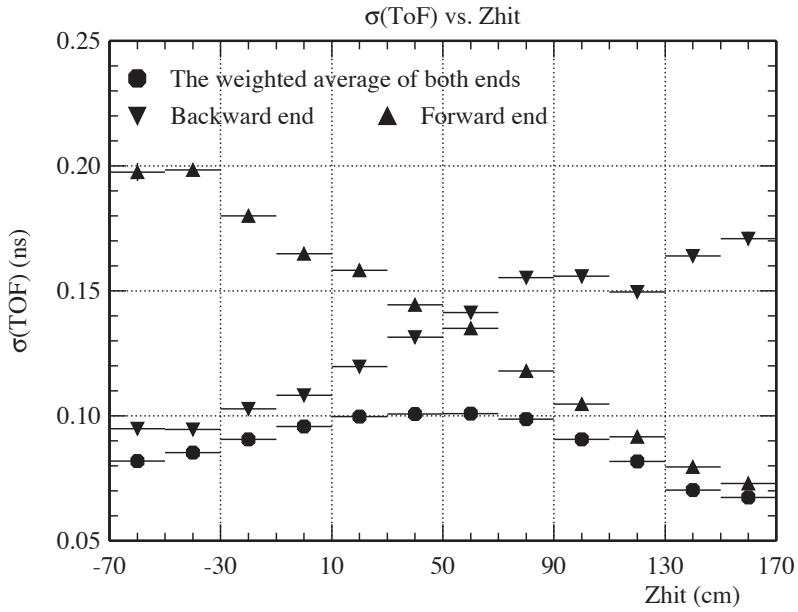

Figure 3.17: Measured time resolution for μ-pair events by TOF.

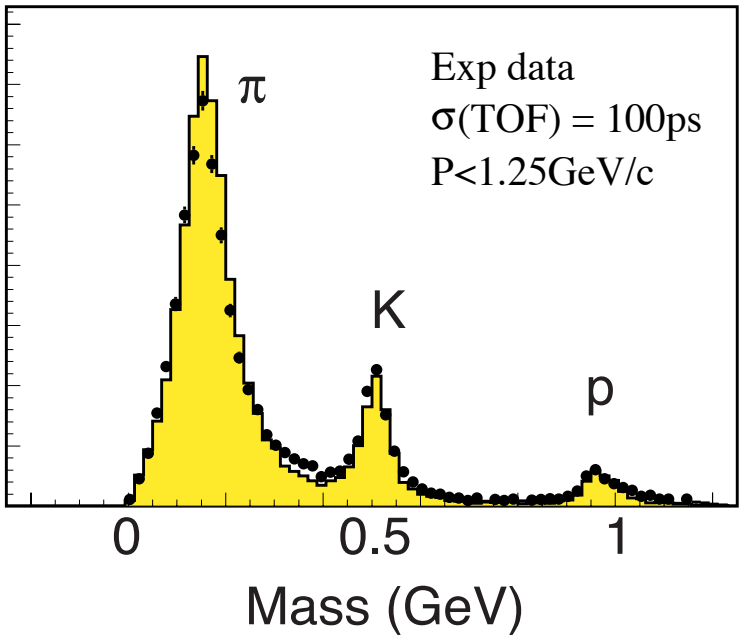

Figure 3.18: Distributions of hadron mass calculated from measured time-of-flight and momentum for particles with $p < 1.25$ GeV/c.

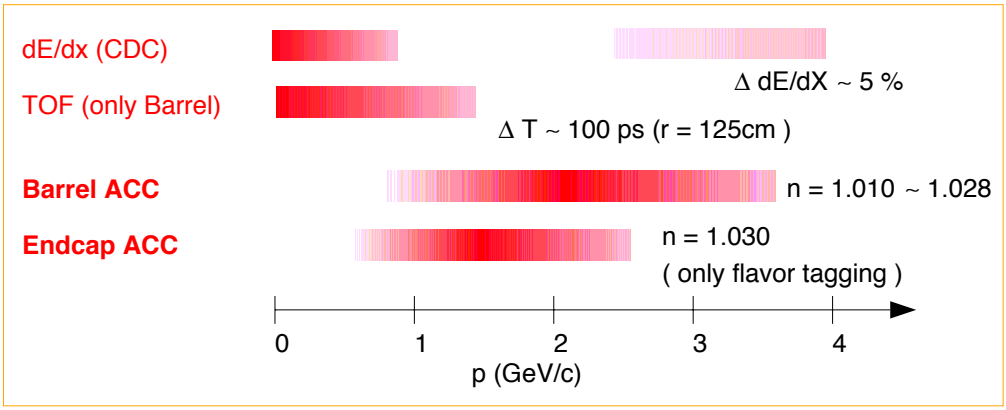

Figure 3.19: Momentum coverage of each detector used for K/π separation.

of 100 ps would provide more than 3 σ separation below 1.2 GeV/c.

The time resolution of Belle TOF system is measured to be about 100 ps with a small z dependence as shown in Fig. 3.17. Figure 3.18 shows the distributions of particle mass calculated from measured time-of-flight and momentum for particles with momentum less than 1.25 GeV/c. We can see clear peaks corresponding to pions, kaons and protons.

K/π separation

Kaon identification (or K/π separation) is very important for flavor tagging in the measurement of CP violation. It is also important in the study of rare decays of B meson like $B \to KK/K\pi/\pi\pi$ to suppress background from other rare decays. K/π separation at Belle [42] is designed to cover whole momentum region necessary for these studies, based on energy loss (dE/dx) measurement in the CDC, time-of-flight measurement by the TOF, and the response of the ACC. We combine the information from these three subdetectors to realize more than 3σ separation between kaons and pions over whole momentum range up to 3.5 GeV/c, which is the kinematic upper limit of particle momenta from B decays in Belle. Momentum coverage of each detector for K/π separation is shown in Fig. 3.19.

For each charged track, we estimate probability with kaon and pion hypothesis for each subdetector. Then a combined likelihood function is calculated in such a way that pion-like particle gives value close to zero and kaon-like particle gives one. We use this combined likelihood to discriminate kaons from pions. The efficiency is measured to be about 85% with charged pion fake rate below 10% for all momenta

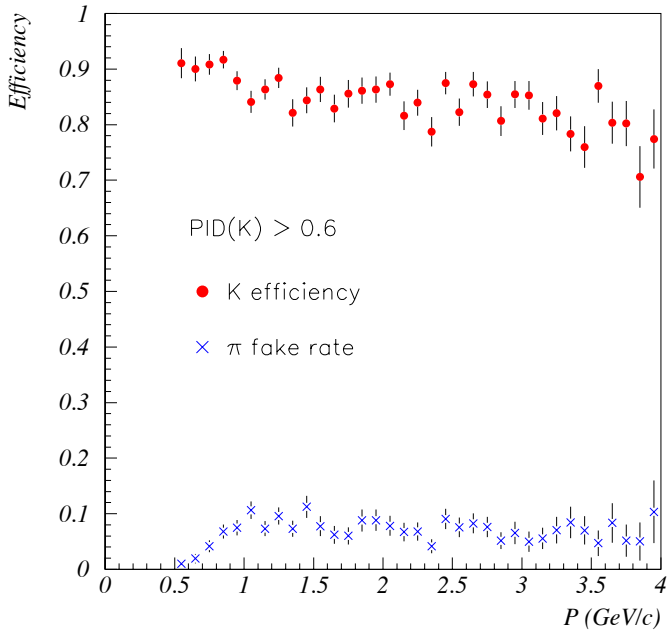

Figure 3.20: Kaon identification efficiency and π fake rate as a function of momentum.

up to 3.5 GeV/c with nominal likelihood requirement. Figure 3.20 shows the kaon identification efficiency and fake rate as a function of momentum, measured with $D^{*+} \to D^0(K\pi) + \pi^+$ decays.

3.2.5 Electromagnetic Calorimeter (ECL)

When a high-energy electron or photon is incident on a thick absorber, it initiates an electromagnetic cascade as pair production and bremsstrahlung which generate more electrons and photons with lower energy. The longitudinal development of the electromagnetic shower scales as the radiation length X_0 of the matter, which is defined as the mean distance over which a high-energy electron loses all but $1/e$ of its energy by bremsstrahlung.

The main purpose of Electromagnetic Calorimeter (ECL) [43] is the detection of photons from B meson decays with high efficiency and good resolutions in energy and position. It also plays a primary role in the electron identification in Belle.

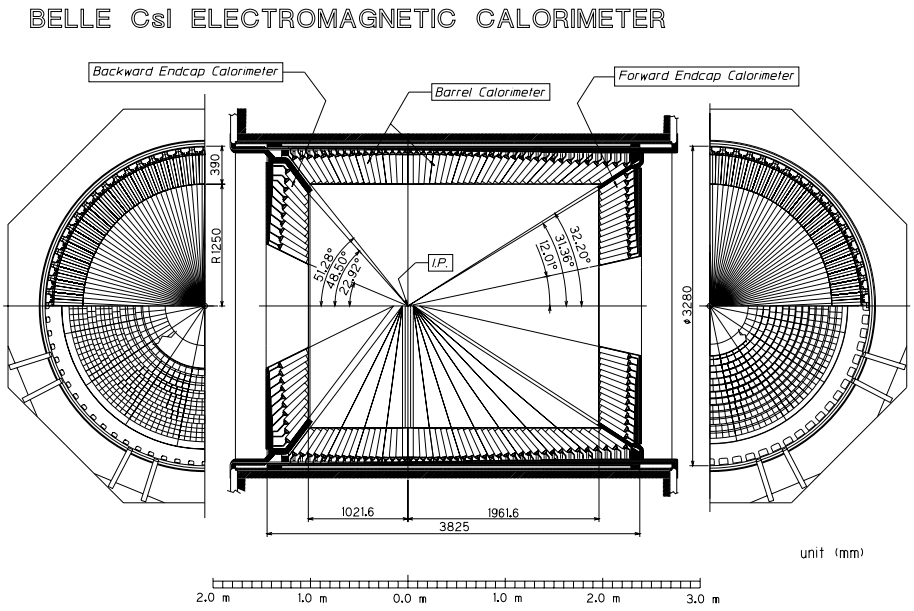

Figure 3.21: Configuration of the ECL.

The ECL consists of 8,736 CsI(Tl) crystals. CsI(Tl) is chosen because of its large photon yield, weak hygroscopicity, mechanical stability and moderate price. An overall configuration of the ECL is shown in Fig. 3.21. Each CsI(Tl) crystal has a tower-like shape with 30 cm length (corresponds to 16.2 X_0) and is assembled to point near the interaction point. The transverse dimensions of crystal vary depending on their polar angle positions. Typical dimensions of a crystal are 55 mm $\times$ 55 mm (at front surface) and 65 mm $\times$ 65 mm (at rear surface) for the barrel part. The barrel part has a 46-fold segmentation in θ and 144-fold segmentation in ϕ. The forward (backward) endcap has a 13- (10-) fold segmentation in θ and ϕ segmentation ranging from 48 to 144 (from 64 to 144). Each crystal is read out by two 2 cm $\times$ 1 cm photodiodes.

The photon energy resolution is measured to be

$$\frac{\sigma_E}{E} = 1.34 \oplus \frac{0.066}{E} \oplus \frac{0.81}{E^{1/4}}\% \tag{3.10}$$

with the beam test before installation into the Belle structure.

Electron identification in Belle [44] primarily relies on a comparison of the charged particle momentum measured by the CDC and the energy deposit in the ECL. While electrons lose all their energies in the ECL crystals with electromagnetic showers, hadrons and muons deposit only part of their energies in the ECL. Therefore the ratio

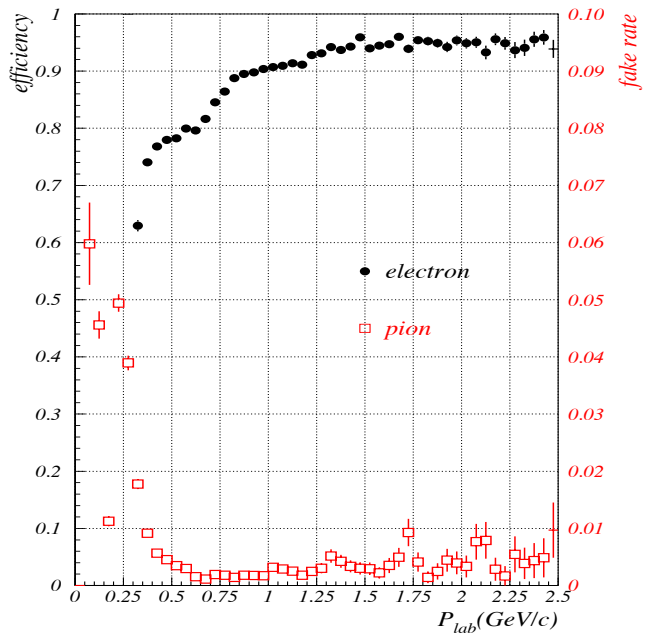

Figure 3.22: Electron identification efficiency (circles) and fake rate for $\pi^{\pm}$ (rectangles) as a function of momentum. Note the different scale for the efficiency and fake rate.

of cluster energy measured by the ECL and charged track momentum measured by the CDC, E/p is close to unity for electrons and tends to be lower for other particles. Because electromagnetic showers evolve faster than hadronic showers, the shapes of showers detected at the ECL are also different between electrons and hadrons. The dE/dx measurement at the CDC is another very effective tool to identify electrons. In particular, it has higher discriminating power in the low momentum region, where E/p is less significant because E/p of electron tends to be apart from unity due to its interaction with materials in front of the ECL. We also use ACC information for hadron rejection in the low momentum region.

The efficiency of electron identification is greater than 90% and a hadron fake rate (the probability to misidentify hadron as electron) is ∼0.3% for $p > 1$ Gev/c. Figure 3.22 shows the electron identification efficiency as a function of momentum.

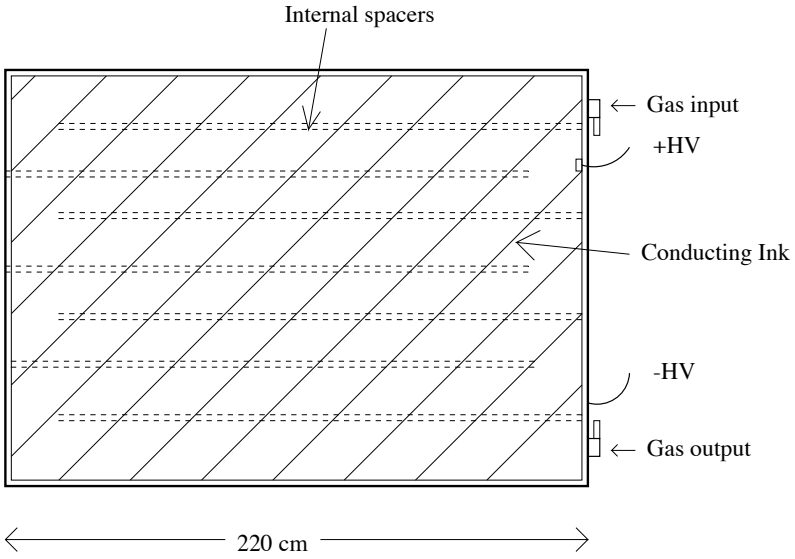

Figure 3.23: Barrel RPC module.

3.2.6 Superconducting Solenoid

A superconducting solenoid provides a magnetic field of 1.5 T parallel to the beam pipe. The superconducting coil consists of a single layer of a niobium-titanium-copper alloy embedded in a high purity aluminum stabilizer. It is wound around the inner surface of an aluminum support cylinder with 3.4 m in diameter and 4.4 m length.

3.2.7 K_L and Muon Detection System (KLM)

As is mentioned in Section 2.4.2, K_L can be identified as hadron shower in the ECL and crude hadron calorimeter. It is not necessary to measure the energy of K_L precisely, but it is required to detect K_L with high efficiency and accurate angular resolution. Also, we need to identify muons with high efficiency and purity over a broad momentum range. The K_L and Muon Detection System (KLM) [45] is designed to satisfy these requirements.

The KLM system consists of alternating layers of charged particle detectors and 4.7 cm thick iron plates. The barrel-shaped region around the interaction point covers an angular range from 45° to 125° in polar angle and the endcaps in forward and backward directions extend this range to 20° and 155°. There are 15 detector layers and 14 iron layers in the octagonal barrel region and 14 detector layers and

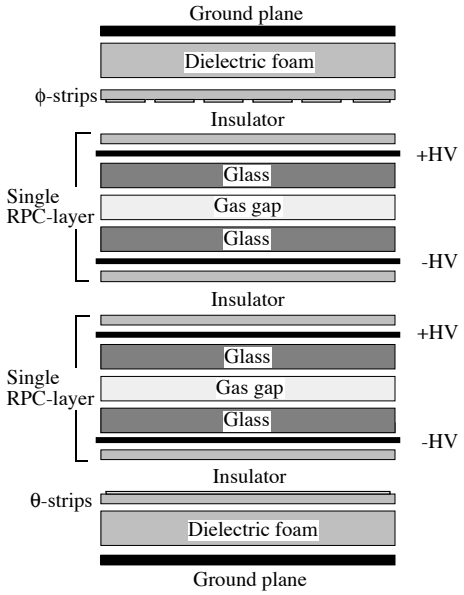

Figure 3.24: Cross-section of a KLM super-layer module.

14 iron layers in each of forward and backward endcaps. The length scale of hadron shower is set by the nuclear interaction length. The iron plates provide a total of 3.9 interaction lengths of material for a particle travelling normal to the detector planes. In addition, the electromagnetic calorimeter (CsI) provides another 0.8 interaction length of material to convert K_L's. The K_L which interacts in the iron or in the ECL produces a cascade shower of ionizing particles, which are detected by the glass-electrode resistive plate counters (RPCs).

RPCs have two parallel plate electrodes with high bulk resistivity ($> 10^{10}$ Ωcm) separated by a gas-filled gap. We have chosen a non-flammable mixture of 62% CH_2FCF_3 (HFC-134a), 30% argon, and 8% butane-silver. The butane-silver is a mixture of approximately 70% n-butane and 30% iso-butane. An ionizing particle traversing the gap initiates a streamer in the gas that results in a local discharge of the plates. This discharge is limited by the high resistivity of the plates and the quenching characteristics of the gas. The discharge induces a signal on external pickup strips, which can be used to record the location and time of the ionization.

Figure 3.23 shows a barrel RPC module. The glass plates are separated by 1.9 mm thick spacers. The spacers are placed every 10 cm so that they channel gas flow through the RPC. Figure 3.24 shows the cross section of a RPC super-layer. Two RPCs are sandwiched between orthogonal pickup strips with ground planes for

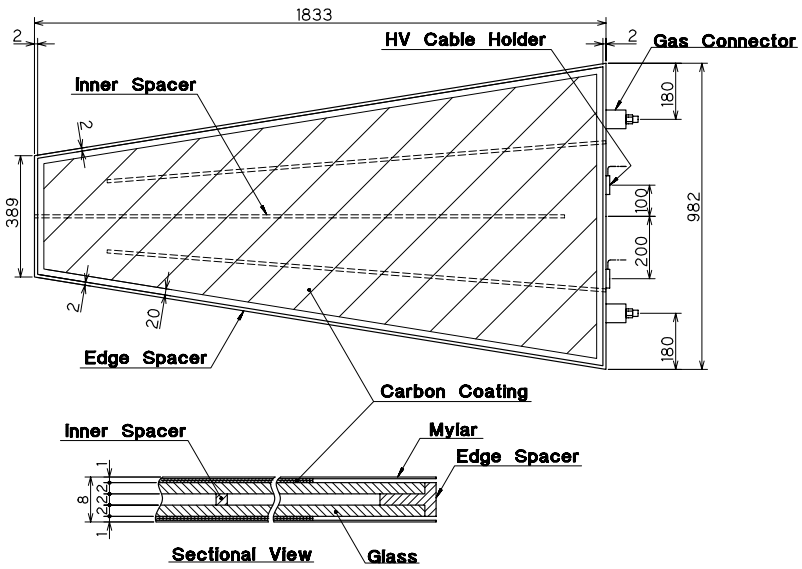

Figure 3.25: Endcap RPC module.

signal reference and proper impedance. Signals from both RPCs are read out by copper strips above and below the pair of RPCs, providing three dimensional spacial information. The thickness of the super-layer module is less than 3.7 cm. The size of barrel RPC module varies from 2.2×1.5 m^2 to 2.2×2.7 m^2. Each barrel module has 48 z strips perpendicular to the beam direction. The inner 7 super-layers have 36 ϕ strips and the outer 8 super-layers have 48 ϕ strips. The pickup strips vary in width from layer to layer but are approximately 50 mm wide in the barrel module.

The endcap module consists of π-shaped RPCs shown in Fig. 3.25. The endcap super-layer module contains 10 RPC modules. Each super-layer module has 96 ϕ and 46 θ pickup strips. The θ strips are 36 mm wide and vary in length from 2 m to 5 m. The ϕ strips are 1.83 m long and vary in width from 19 mm to 47 mm.

The K_L angular resolution and detection efficiency are studied using MC simulation [46]. The K_L candidates are selected from clusters in the KLM which are not associated with charged tracks. If there is a cluster in the ECL within 15° from the KLM cluster, we consider that ECL cluster as associated with the KLM cluster. If an ECL cluster is associated, the direction of the K_L candidate is defined as the direction of the ECL cluster, since the angular resolution of the ECL is better than that of the KLM. Otherwise, the direction of the K_L candidate is defined as the direction of the center of the KLM cluster.

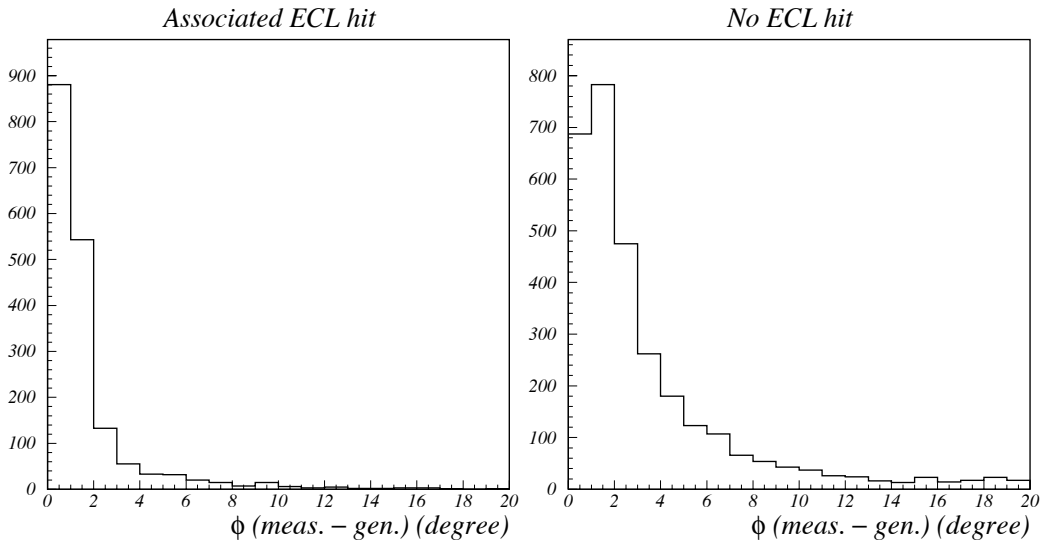

Figure 3.26: Difference between the directions of generated and detected K_L in the single K_L MC. The left (right) plot shows the case with (without) associated ECL cluster.

The angular resolution of K_L is estimated using a *single* K_L *MC*, in which each event contains only single K_L. K_L's are uniformly generated in the momentum range between 0.3 and 4.5 GeV/c, θ range from 25° to 150°, and ϕ range from 0° to 360°. Figure 3.26 shows the difference between the directions of generated and detected K_L. The angular resolutions of the direction of the K_L candidates are estimated to be 1.5° and 3° in FWHM for those with and without ECL cluster, respectively. We find that the angular resolution is independent of the K_L momentum.

Figure 3.27 shows the detection efficiency as functions of momentum in the laboratory frame and polar angle of the K_L in the single K_L MC. The efficiency includes the acceptance of the detector. The detection efficiency increases as the momentum of the K_L increases, but is almost independent of the angle.

Figure 3.28 shows the difference between the direction of K_L candidate measured in the KLM and the missing momentum vector in the real data. The missing momentum is calculated using all other measured particles in the event. The histogram shows a clear peak where the direction of the K_L candidate is consistent with that of the missing momentum of the event. This is the evidence that we are correctly detecting K_L with the KLM. A large deviation of the missing momentum direction from K_L direction is mainly due to undetected neutrinos and particles escaping the detector acceptance.

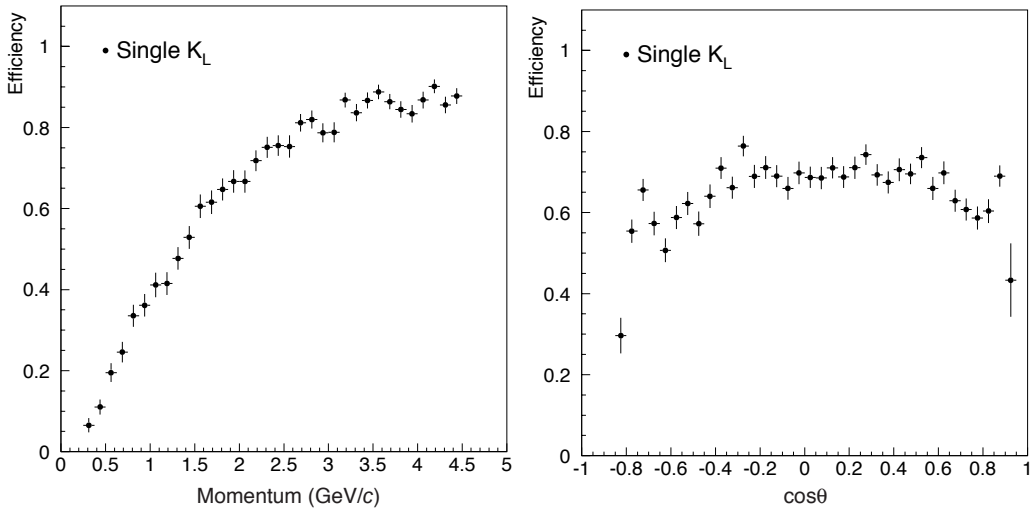

Figure 3.27: K_L detection efficiency as a function of momentum in the laboratory frame (left) and polar angle (right) estimated using single K_L MC.

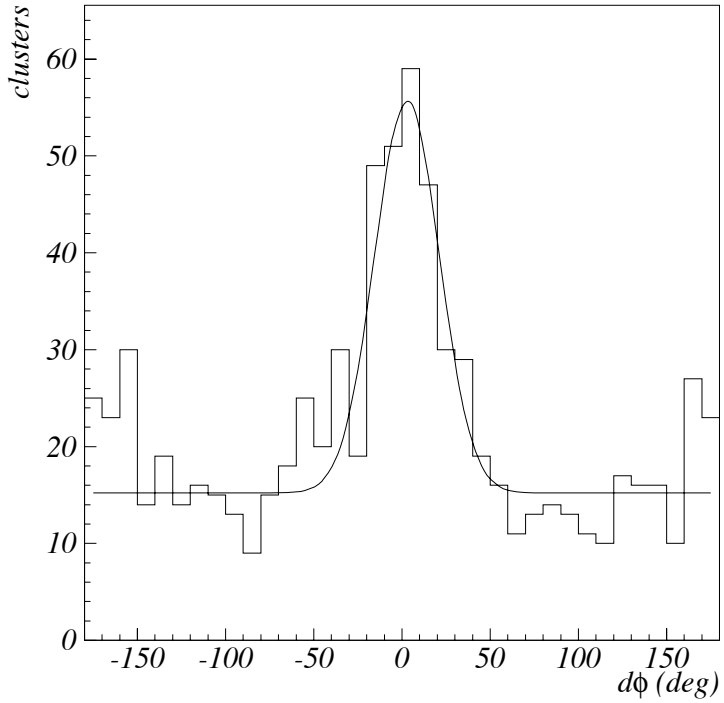

Figure 3.28: Difference between the KLM cluster direction and the missing momentum direction in the real data.

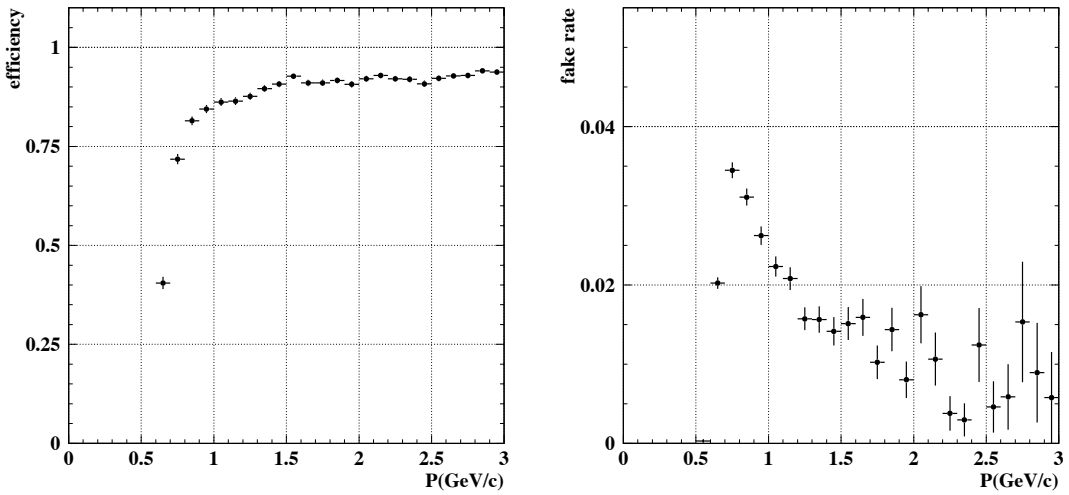

Figure 3.29: Muon identification efficiency (left) and fake rate (right) as a function of momentum.

Muons do not cause hadronic interaction. Because of their much heavier mass than electrons, muons do not cause bremsstrahlung, thus they do not cause electromagnetic showers. Therefore muons penetrate more matters with less interactions compared with other particle species. The multiple layers of RPC and iron allow discrimination between muons and charged hadrons based on their range (distance before stopping in the material) and transverse scattering.

Muon identification is done based on KLM response to the track [47]. The efficiency is measured to be greater than 90% and a hadron fake rate is less than 2% for $p > 1$ GeV/c. Figure 3.29 shows the muon identification efficiency as a function of momentum.

3.2.8 Trigger and Data Acquisition (DAQ) System

The cross section of $B\overline{B}$ events is rather small at Belle, compared to other processes like Bhabha and $\gamma\gamma$ processes. We need to achieve high trigger efficiency for important processes while suppressing the trigger rate for uninteresting events. Due to the high beam current of the KEKB, the trigger suffers from severe beam background. Because the background rate is very sensitive to the accelerator conditions, the trigger system is required to be flexible so that background rates are kept within the tolerance of the data acquisition system.

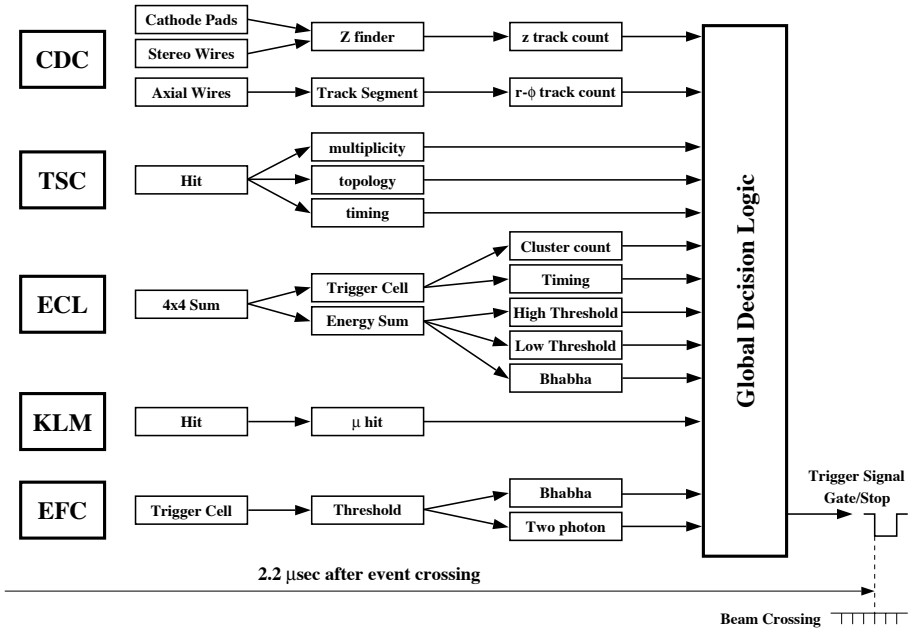

Figure 3.30: Schematic view of the Belle trigger system.

Figure 3.30 shows a schematic view of the Belle Level 1 trigger system. It is composed of the sub-detector trigger system and the central trigger system, called the Global Decision Logic (GDL) [48]. The sub-detector trigger system is based on two categories; track triggers and energy triggers. The CDC and TOF/TSC are used to provide the trigger signals from charged particles. The ECL trigger system provides triggers based on total energy deposit and cluster counting of crystal hits. The KLM trigger provides additional information on muons. The sub-detectors process event signals in parallel and provide trigger information to the GDL. The GDL combines all the information to characterize the event type and makes decision to initiate the data acquisition within 2.2 μsec from the beam crossing. The trigger rate is typically 200 Hz and the deadtime is about 4% in normal operation condition.

A schematic view of the Belle data acquisition(DAQ) system is shown in Fig. 3.31. The data from the sub-detectors should be digitized within 200 μsec in order to keep deadtime less than 10% at 500 Hz trigger rate. The signals from sub-detectors are converted to timing signals by Q-to-T converters except for SVD and KLM. The KLM provides the time-multiplexed information on a single line. All subdetectors except the SVD use TDC readout system which is controlled by VME and FASTBUS system. A typical data size of a hadronic event is measured to be about 30 kB, which corresponds to a maximum data transfer rate of about 15 MB/s.

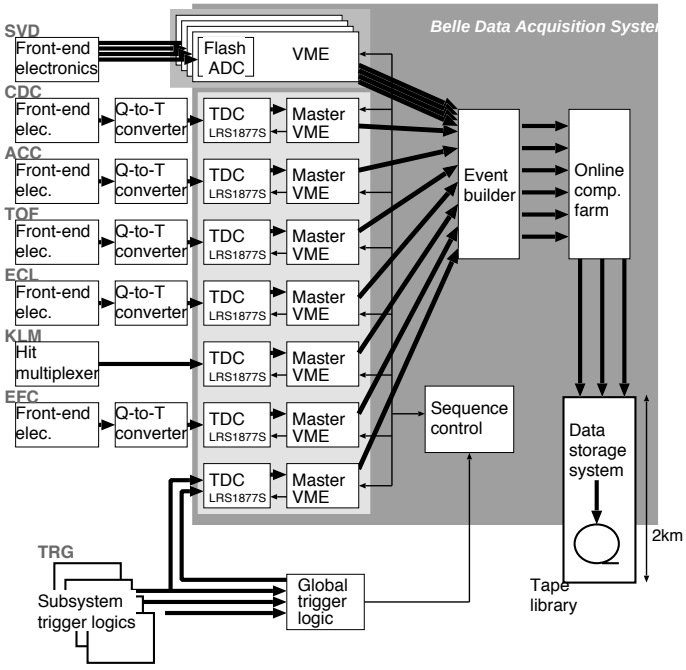

Figure 3.31: Schematic view of the Belle data acquisition (DAQ) system.

Data from each subsystem are combined into a single event record by an Event Builder, which sorts the detector-wise parallel data to the event-wise parallel data. The output from the Event Builder is transferred to an online computer farm. The online computer farm sends data to the mass storage system at the Computer Center, 2 km apart from the Belle detector, via optical fiber. The online computer farm also sends the sampled events to the Data Quality Monitor, by which experiment shifters monitor the detector and DAQ condition.

3.2.9 Offline Software and Computing

All software except for a few HEP specific and non HEP specific free software packages has been developed by the members of the Belle collaboration. In particular, the mechanisms to handle event structure and input and output formatting and to process events in parallel on a large SMP (Symmetric Multiple Processor) compute server have been developed locally using C and C++ programming languages.

The event processing framework, called BASF (Belle AnalysiS Framework), takes users' reconstruction and analysis codes as modules which are linked dynamically at the run time. A module is written as an object of a class of C++. The class,

inherited from the module class of BASF which has virtual functions for events, begins and ends run processing and other utility functions such as initialization, termination and histogram definitions. Modules written in Fortran and C can also be linked using wrapper functions.

The data transfer between modules is managed by PANTHER, an event and I/O management package developed by the Belle collaboration. PANTHER describes the logical structure and inter-relationships of the data using an entity relationship model. In order to store data (structure) in the event structure one writes a description file as an ASCII text file. A PANTHER utility converts the description file into C and C++ header files and source code. The user will include the header files in his/her code and the source code is compiled and linked into the user module to have access to the data structure in the module.

Monte Carlo Simulation

The QQ event generator [49], which was originally developed by CLEO collaboration, is used to simulate the physical process of production and decay of particles. The QQ can handle both $\Upsilon(4S)$ decay and continuum ($u\bar{u}$, $d\bar{d}$, $s\bar{s}$, and $c\bar{c}$) process. The decay of $\Upsilon(4S)$ is simulated according to the decay tables which contain decay modes and their branching fractions. The decay table has been updated by Belle collaboration to include most up-to-date measurements. The continuum generation uses the LUND program (JETSET 7.3 [50]), in which the subsequent hadronization process is based on the Lund string fragmentation model. The output of the event simulator is saved in a table and fed to a detector simulator.

The full detector simulator package has been developed based on GEANT3 [51] library. GEANT is a library developed at CERN to simulate passage of particles through materials. It traces all particles step by step, simulating their reactions with materials. The geometry of the Belle detector is implemented in the full simulator package. The simulated events are analyzed through the same analysis chain as the real data. All MC samples used in this analysis are simulated with this full-simulation package.

Chapter 4

Event Selection

In this chapter, we describe the reconstruction procedures of the B^0 decays into $J/\psi K_L$ and $J/\psi K^{*0}$. We call the reconstructed B meson in the procedure described in this chapter a CP-side B meson or B_{CP} since it is supposed to decay into CP eigenstate. The other B meson is referred to as a tagging-side B meson or $B_{\rm tag}$ since it is used for flavor tagging in our analysis.

4.1 Data Set

The analysis presented in this thesis is based on the data taken between January 2000 and July 2001. The integrated luminosity for each day and the history of total accumulated luminosity are shown in Fig. 4.1. The total integrated luminosity of 31.4 fb^{-1} has been accumulated in this period, where 29.1 fb^{-1} and 2.3 fb^{-1} are taken on Υ(4S) resonance and 50 MeV below, respectively. The total number of $B\overline{B}$ pairs used for this analysis is approximately 3.1×10^7.

4.2 $B\overline{B}$ Event Selection

Although KEKB operates on the $\Upsilon(4S)$ resonance, several other processes also take place with comparable, or even substantially larger cross sections with $B\overline{B}$ production. To reduce the size of data processed, we impose selection criteria which suppress uninteresting events. Such background events generally have lower multiplicity of particles than B decays. Also, the energy deposited to the calorimeter is efficient to discriminate hadronic events and other processes. For example, $\tau^+\tau^-$ events and $\gamma\gamma$ events have smaller energy deposit, while in QED process, such as Bhabha

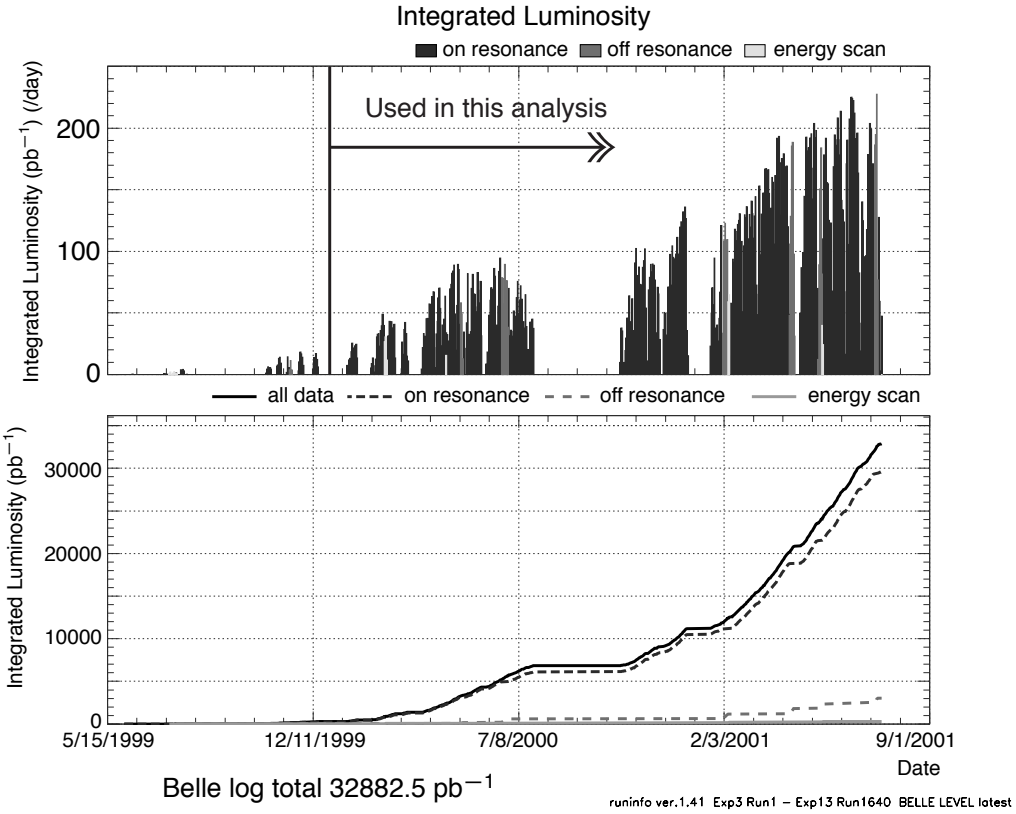

Figure 4.1: History of KEKB luminosity

($e^+e^- \to e^+e^-$ scattering) events, most of energy is deposited in the calorimeter.

There is also a contribution from interactions between electron/positron beams and residual gas molecules in the beam pipe. Furthermore, the beam electrons or positrons can be knocked out of orbit by these beam-gas interactions and hit the beampipe wall, causing additional backgrounds. These types of events can be rejected by requiring the tracks should come from the interaction region, in addition to the requirements of particle multiplicity and energy deposit.

We have developed a set of selection criteria to suppress non-hadronic events [52]. We impose the following criteria to select the hadronic events:

- At least three "good" tracks must exist, where a "good" track is defined by (i) $|r| < 2.0$ cm and $|z| < 4.0$ cm at the closest approach to the beam axis, (ii) transverse momentum > 0.1 GeV/c.

- More than one "good" cluster must be observed in the barrel region of the calorimeter, where a "good" cluster is defined in such a way that it is detected

in the good acceptance region with an energy deposit greater than 0.1 GeV and no track is associated with the cluster.

- The absolute value of the momentum balance in the z-component calculated in the rest frame of $\Upsilon(4S)$ should be less than 50% of the center-of-mass energy:

$$\left|\sum p_z^*\right| < 0.5\sqrt{s}, \tag{4.1}$$

where s is the square of the center-of-mass energy.

- The event vertex, which is reconstructed from the "good" tracks defined above, must be within 1.5 cm and 3.5 cm from the nominal interaction point in r and z direction, respectively.

- The total visible energy, which is computed as a sum of the energy of the "good" tracks assuming pion mass and that of the "good" clusters, in the rest frame of $\Upsilon(4S)$ should exceed 20% of the center-of-mass energy:

$$E_{\rm vis} > 0.2\sqrt{s}, \tag{4.2}$$

- A sum of all cluster energies, after boosted back into the rest frame of $\Upsilon(4S)$, should be between 10% and 80% of the center-of-mass energy:

$$0.10\sqrt{s} < E_{\rm sum} < 0.80\sqrt{s}. \tag{4.3}$$

From Monte Carlo study, these selection criteria allow us to retain more than 99% of $B\overline{B}$ events and 84% of continuum ($e^+e^- \to q\bar{q}, \quad q = u, d, s, c$) events, while rejecting most of non-hadronic events [1].

Events passing the hadronic event selection criteria are further selected by their event shape to suppress continuum events. The k-th Fox-Wolfram moment H_k is defined as [53]

$$H_k = \frac{1}{s}\sum_i^N \sum_j^N [|\vec{p_i}| \cdot |\vec{p_j}| P_k(\cos\phi_{ij})] \tag{4.4}$$

where N is the number of particles, $P_k(x)$ is the Legendre polynomial of order k and ϕ_{ij} is the angle between the three-momenta of the i-th and j-th particles. R_2, defied as $R_2 = H_2/H_0$, is close to 1 if the event is jet-like (likely to be continuum events) and 0 if the event is spherical (likely to be $B\overline{B}$ events).

[1] We have other pre-selection criteria to save processing resources, but they are looser than the criteria described in this chapter and have no effect on the result.

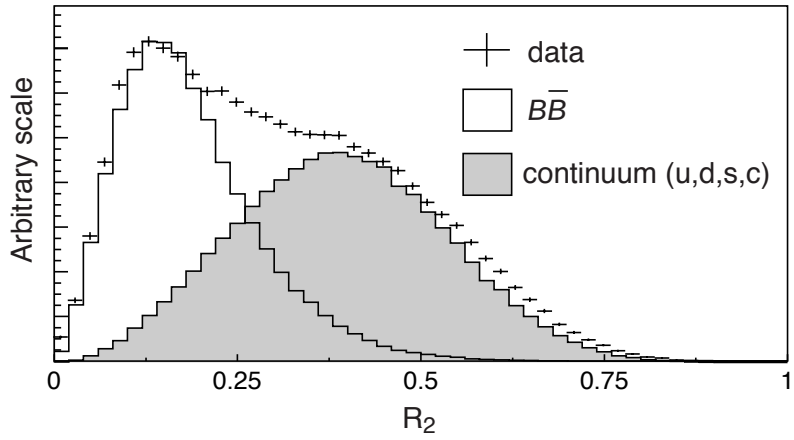

Figure 4.2: R_2 distributions for real data after hadronic event selection (cross), $B\overline{B}$ MC (open histogram) and continuum MC (hatched histogram) events. The normalizations of MC distributions are arbitrary.

Events satisfying $R_2 \leq 0.8$ are used in the $B^0 \to J/\psi K_L$ analysis. This requirement is determined to be loose enough so that we do not throw away the signal events, since we use R_2 as an input of likelihood ratio for event selection as described in later section. For $B^0 \to J/\psi K^{*0}$ analysis, we use events satisfying $R_2 \leq 0.5$ to suppress background from continuum events. Figure 4.2 shows the distributions of R_2 for real data after hadronic event selection, $B\overline{B}$ and continuum MC events. The normalizations of MC distributions are arbitrary. We can see almost all $B\overline{B}$ events pass the R_2 criterion. The efficiency loss due to the requirement $R_2 \leq 0.5$ ($R_2 \leq 0.8$) is estimated to be less than 0.5% (almost zero). The effect of the $B\overline{B}$ selection criteria to the asymmetry measurement is negligible.

4.3 Reconstruction of $B^0 \to J/\psi K_L$

As mentioned in Section 3.2.7, we cannot measure the magnitude of momentum of K_L mesons ($|\mathbf{p}_{K_L}|$) with our detector but only the direction of K_L meson is measured. This makes the analysis technique of $B^0 \to J/\psi K_L$ decays quite different from that of other exclusive modes like $B^0 \to J/\psi K_S$.

The J/ψ is reconstructed through its leptonic decay into a pair of electrons or muons and its four-momentum ($\mathbf{p}_{J/\psi}, E_{J/\psi}$) is completely determined. The K_L direction is measured by its hadronic interaction in ECL and/or KLM. For a $B^0 \to$

$J/\psi K_L$ decay, the K_L momentum must satisfy

$$m_{B^0}{}^2 = \left(E_{J/\psi} + \sqrt{|\mathbf{p}_{K_L}|^2 + m_{K_L}{}^2}\right)^2 - \left|\mathbf{p}_{J/\psi} + \mathbf{p}_{K_L}\right|^2, \tag{4.5}$$

where m_{B^0} and m_{K_L} are the mass of B^0 and K_L, respectively, and both are known. Since we measure the direction of the K_L momentum, the only unknown parameter in Eq. (4.5) is the magnitude of momentum of K_L, $|\mathbf{p}_{K_L}|$. Eq. (4.5) has two solutions for $|\mathbf{p}_{K_L}|$, but only one is positive, hence acceptable. The momentum and the energy of B^0 are then simply the sum of those of the J/ψ and the K_L. The magnitude of B^0 momentum in the center-of-mass system, p_B^*, is obtained by a Lorentz transformation, using the known values of the beam energies. If the assumption is correct, p_B^* must be monochromatic, apart from the spread in the beam energies and the angular resolution in reconstructing the K_L direction. With angular resolution of 30 mr, the contribution to the error of p_B^* from the measurement of K_L direction is comparable to that from the spread in beam energies.

Due to fewer constraints, $B^0 \to J/\psi K_L$ mode suffers from larger background than $B^0 \to J/\psi K_S$ mode. Most of background comes from other $B \to J/\psi X$ decays, where J/ψ is combined with K_L candidates either originating from real K_L, or faked by other charged and neutral particles or noise hits. The largest source of background is $B^0(B^+) \to J/\psi K^{*0}(K^{*+})$, where K^* decays into K_L and π, due to its similar event topology. Since $B^0 \to J/\psi K^{*0}(K^{*0} \to K_L\pi^0)$ has its own CP asymmetry, it must be correctly treated in the analysis procedure. There are also other background modes with their own CP asymmetry, like $J/\psi K_S$, $\psi(2S)K_L$, and $\chi_{c1}K_L$. These backgrounds can dilute the true $\sin 2\phi_1$ significantly, if they are not treated correctly in the analysis procedure. Therefore, suppression and understanding of the background is crucial in the analysis of $B^0 \to J/\psi K_L$ mode.

4.3.1 Reconstruction of J/ψ

J/ψ is reconstructed using dilepton decays, $J/\psi \to \mu^+\mu^-$ and e^+e^-. We use oppositely charged track pairs, where both tracks are positively identified as leptons (electrons or muons). In order to remove either poorly reconstructed events or tracks which do not come from the interaction region, we require the closest approach of the track to the nominal interaction point in the $r(z)$ direction to be less than 1(4) cm for both lepton tracks from J/ψ.

For $J/\psi \to e^+e^-$ channel, we include the four-momentum of every photon detected within 0.05 radians of the original e^+ or e^- direction in the invariant mass calculation to correct final state interaction or real bremsstrahlung in the inner parts of

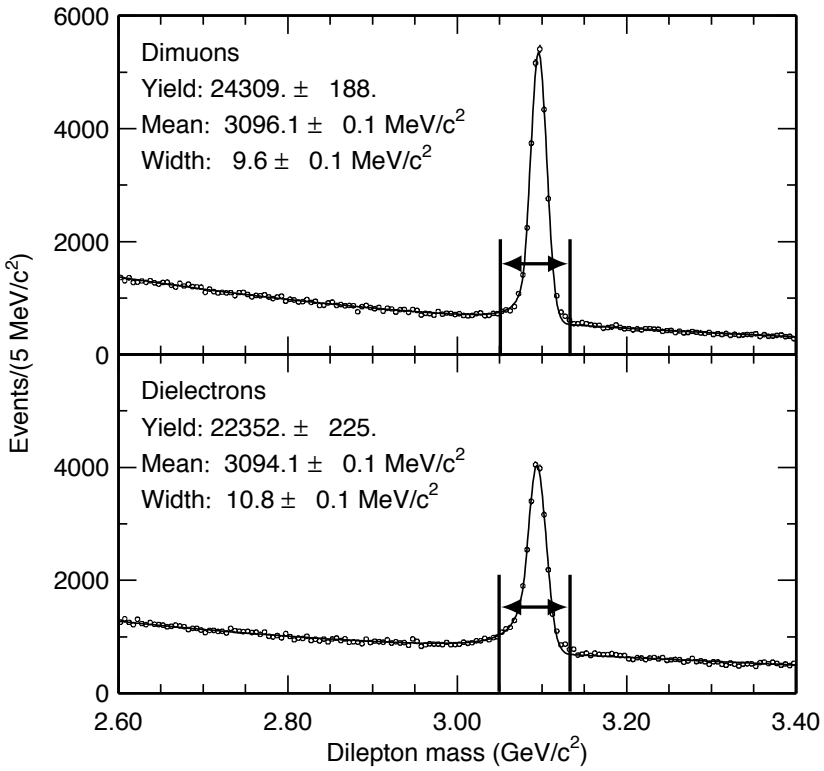

Figure 4.3: Invariant mass distribution of J/ψ candidates for $\mu\mu$ (upper) and ee (lower) modes. The arrows indicate the signal region.

the detector. We accept pairs with invariant mass within $3.05 < M_{\ell\ell} < 3.13$ GeV/c^2. This requirement is somewhat tighter than other exclusive mode reconstruction (*e.g.* $B^0 \to J/\psi K_S$). This is because the purity of final $B^0 \to J/\psi K_L$ sample heavily depends on the amount of background in J/ψ sample. For the lepton tracks of the J/ψ candidates, a kinematic fit with mass and vertex constraint is performed using the kinematic fitting routine developed at Belle [54] to improve the resolution of the J/ψ momentum.

Figure 4.3 shows the invariant mass distribution for $J/\psi \to \mu^+\mu^-$ and e^+e^- channels. Clear peaks of signal can be seen above very low backgrounds in both distributions.

4.3.2 Reconstruction of K_L

K_L candidates are reconstructed as hadronic showers in KLM and/or ECL. Hit information from each subdetector is analyzed to form *clusters*, which are used as

a unit to search for K_L candidates. Main background sources of K_L candidates are interactions of other particles with detector which are misidentified as K_L. This is caused by neutral particles such as γ, or charged particles such as $\pi^{\pm}$ when we fail to associate tracks with clusters or to reconstruct tracks.

We classify K_L candidates depending on whether they have associated clusters in KLM or not. *KLM candidates*, which are defined as K_L candidates with KLM hits, are expected to have fewer background than candidates with only ECL clusters, since ECL clusters contain more background from γ's. K_L candidates with only ECL cluster are called *ECL candidates*.

We select K_L candidates in each category as follows.

KLM candidates

KLM candidates are selected by the following algorithm:

1. Combine nearby KLM hits which are within 5° opening angles to form KLM clusters. Repeat the process until no more hit is found within 5° opening angle from each other.

2. If the KLM cluster is accompanied by an ECL cluster with energy more than 0.16 GeV within 15° opening angle, we consider this ECL cluster is associated with the KLM cluster. The requirement on energy deposit in the ECL is to eliminate background from γ's.

3. Check if the number of RPC superlayers in the KLM cluster is greater than one to reject noise hits in the KLM when there is no associated ECL cluster.

4. Calculate the direction of K_L candidates:

   - If an ECL cluster is associated, the direction of the K_L candidate is defined as the direction of the ECL cluster, since the angular resolution of the ECL is better than that of the KLM.
   - If there is no associated ECL cluster, the direction of the K_L candidate is defined as the direction of the center of the KLM cluster.

5. Check if the KLM cluster is associated with charged tracks. If there is a charged track within 15° cone around the K_L candidate direction, we reject this cluster to suppress background caused by charged particles.

ECL candidates

The most significant background in K_L candidates with only ECL clusters is interaction of γ's with the ECL. We discriminate electromagnetic interactions caused by γ from hadronic interactions caused by K_L based on the difference in characteristics of electromagnetic and hadronic showers, such as the deposited energy and the width of interaction. Another source of background is interactions of charged particles with ECL. We reject them by checking the matching of charged particle tracks and ECL clusters.

ECL candidates are selected using following information:

- The distance between the ECL cluster and the closest charged track position. This information is used to reject clusters originating from charged particles.
- The cluster energy recorded in the ECL.
- The ratio of energy summed in a 3×3 array of CsI crystals surrounding the crystal at the center of the shower to that of a sum of 5×5 array of crystals centered on the same crystal ($E9/E25$). Because electromagnetic shower evolves faster than hadronic shower, the shape of shower detected at ECL is also different. $E9/E25$ tends to be lower for hadronic shower than that for electromagnetic shower.
- The ECL shower width. For the same reason as the difference in $E9/E25$, the shower width tends to be broader for the hadronic shower than the electromagnetic shower.
- The invariant mass of the shower. When an ECL cluster has a non-zero width, it can be subdivided into several subclusters. We calculate invariant mass from four-momentum of each sub-cluster. Clusters originate from K_L tend to have heavier invariant mass due to their hadronic nature than that from γ's.

Instead of requiring discrete criteria for each discriminating variables, we use likelihood value which combines all the information together and, therefore, is expected to have more discriminating power. We calculate the likelihood for the above five discriminant variables for both true K_L clusters and fake clusters based on distributions obtained from MC simulations. Figure 4.4 shows the distributions of discriminant variables. Taking the products of the above five likelihood for true K_L and fake clusters, we form the likelihood ratio $L_{K_L}/(L_{K_L} + L_{fake})$. Upper figures

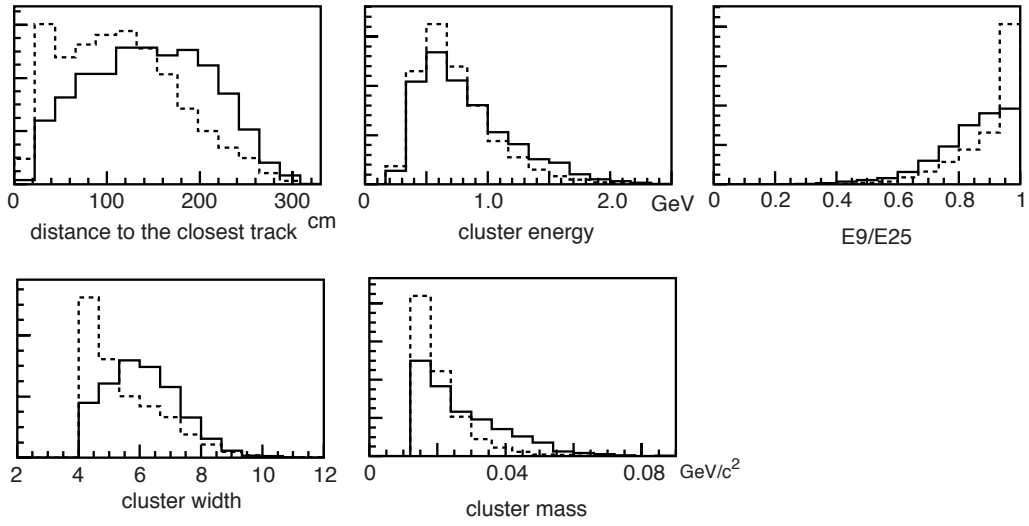

Figure 4.4: Distributions of likelihood variables to select ECL candidates. Solid (dashed) histograms are for signal (background) in MC.

of Fig. 4.5 show the likelihood ratio distributions for signal and background. As is discussed in Section 2.4.2, we can define the Figure of Merit (FOM) as $S/\sqrt{S+B}$, where $S(B)$ denotes the number of signal (background). We look for the optimal requirement on the L to maximize FOM while keeping S/B ratio as high as possible. Lower figures show the S/B ratio and FOM as a function of likelihood requirement. From these distributions, we require the likelihood ratio to be greater than 0.5 for ECL candidates.

4.3.3 Reconstruction of B^0

We treat K_L candidates with and without KLM hits separately in $B^0 \to J/\psi K_L$ reconstruction, because we expect K_L candidates with KLM hits to have higher purity compared to candidates selected with ECL information only and because the background source may be different for these two categories. We therefore first try to reconstruct $B^0 \to J/\psi K_L$ with only KLM candidates, and only if none of the KLM candidates satisfies the selection criteria described below, the ECL candidates are considered.

We impose the following criteria to select the $B^0 \to J/\psi K_L$ candidates:

- The K_L candidate should be within 45° from the expected direction of K_L

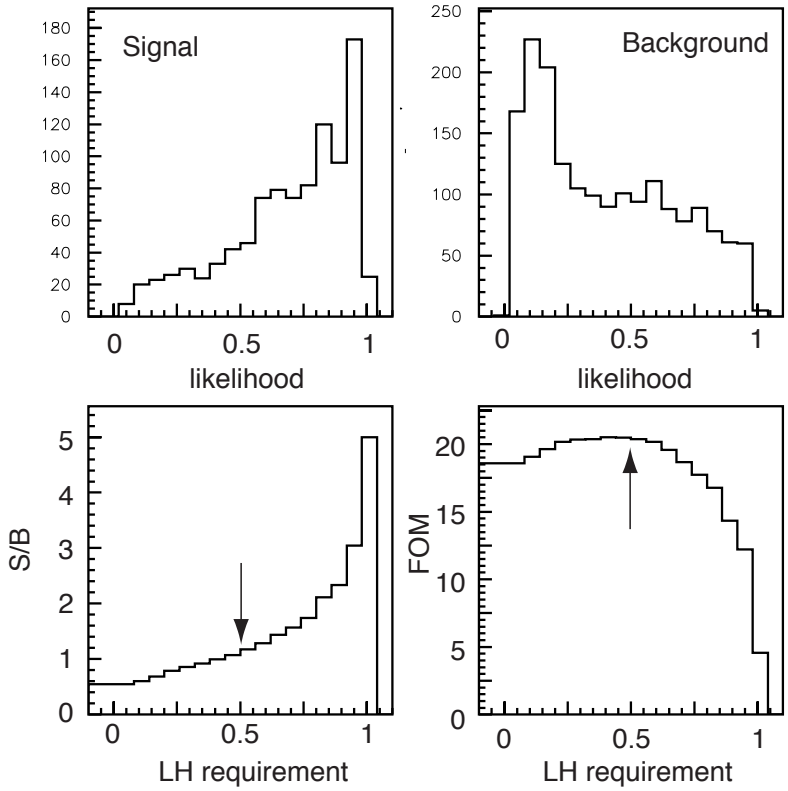

Figure 4.5: (Upper) ECL candidate likelihood ratio distributions for signal (left) and background (right). (Lower) S/B (left) and FOM($=S/\sqrt{S+B}$) (right) vs. likelihood requirement value. The arrows indicate the requirement value we use.

calculated from the J/ψ momentum assuming the B^0 being at rest in the rest frame of $\Upsilon(4S)$. This requirement is effective to reject the background from accidental combination of J/ψ and K_L.

- The direction of K_L candidate should not match the direction of π^0's daughter photons with the momentum of π^0 above 0.8(1.2) GeV/c for KLM(ECL) candidates, where the π^0 is required to have $0.12 < M_{\gamma\gamma} < 0.15$ GeV/c^2.

- The selected J/ψ is not consistent with any of the following modes: $J/\psi K^+$, $J/\psi K_S$, $J/\psi K^{*+}$ ($K^{*+} \to K^+\pi^0, K_S\pi^+$), and $J/\psi K^{*0}$ ($K^{*0} \to K^+\pi^-, K_S\pi^0$), where $|\Delta E|$ <50 MeV and 5.27< $M_{\rm bc}$ <5.29 GeV/c^2 are required. [2]

The last requirement is effective not only to reduce the number of background events from these modes, but also to reduce the systematic bias in the vertex recon-

[2]The definition of ΔE and $M_{\rm bc}$ will be given in the next section.

struction, as is explained below. It is possible in the background events of $J/\psi K_L$ that charged particles from B_{CP} are considered as from $B_{\rm tag}$ in the reconstruction, which may cause significant bias in the vertex reconstruction.

Let us consider the case of $J/\psi K^+$ event as an example. If this event has a fake K_L cluster and is reconstructed as $J/\psi K_L$, K^+ is not used in the reconstruction of B_{CP} and therefore used in the tagging-side B reconstruction. This leads to a shorter reconstructed Δz than the true value, since K^+ (with high-momentum, hence with good position resolution) pulls the tagging-side vertex towards CP-side vertex. Table 4.1 shows the r.m.s. of Δz distribution in MC sample for signal and some of background modes with and without vetoing exclusive modes. The signal mode ($J/\psi K_L$) is added as a reference since it does not have bias due to the above effect. Without excluding reconstructed exclusive modes, the r.m.s. tends to be smaller due to contamination from charged tracks from B_{CP}. After applying the veto, the effect is clearly reduced.

Table 4.1: r.m.s. of Δz distribution for signal and major background modes with and without vetoing exclusive modes.

Mode	$J/\psi K_L$	$J/\psi K^+$	$J/\psi (K^+\pi^+)$	$J/\psi (K^+\pi^0)$
w/o veto (μm)	309.1	246.4	244.9	265.2
w/ veto (μm)	309.1	304.5	266.3	278.5

In order to suppress the background further, we use a likelihood ratio based on the following variables:

- The momentum of J/ψ in the center-of-mass system (CMS), $p^*_{J/\psi}$. Due to its two-body kinematics, J/ψ from $B^0 \to J/\psi K_L$ decay has higher momentum than that from backgrounds. ($p^*_{J/\psi}$ is not monochromatic due to the motion of B in the CMS.)

- Fox-Wolfram moment ratio R_2.

- Angle between K_L direction and the closest charged track with momentum larger than 0.7 GeV/c. This variable is effective to reject fake K_L clusters made by charged particles. The momentum requirement is determined based on MC study.

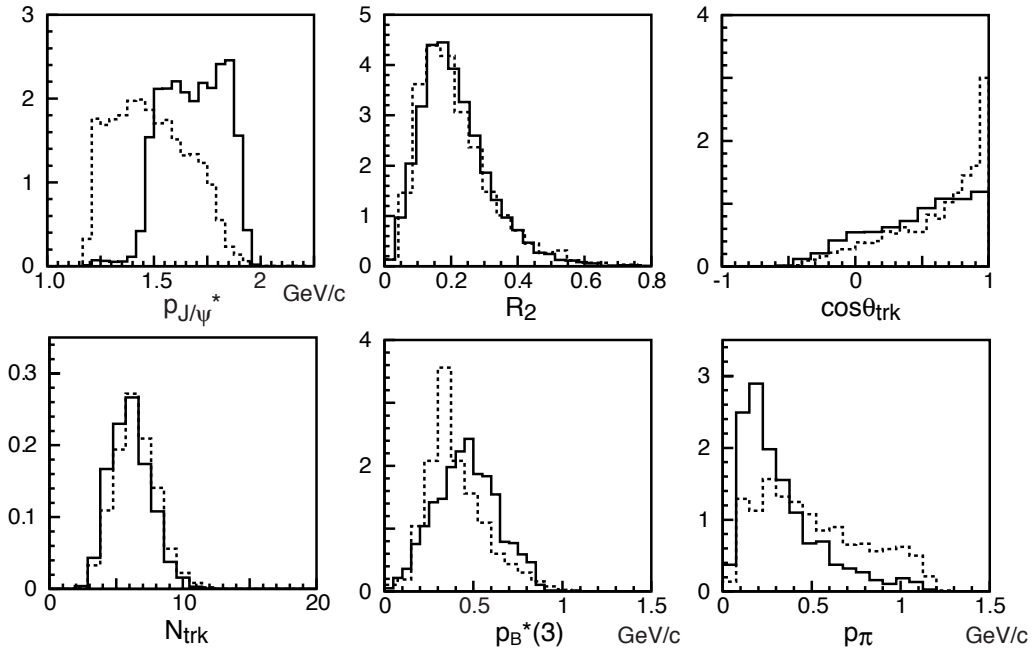

Figure 4.6: Distributions of $B^0 \to J/\psi K_L$ likelihood variables. Solid(dashed) histograms are for signal(background) MC.

- Number of charged tracks with p_t >0.1 GeV/c, $|dr|$ <2 cm, and $|dz|$ <4 cm, where dr and dz represent the distance of closest approach to the interaction point. Because $B^0 \to J/\psi K_L$ event has only two charged tracks in the CP side, the average multiplicity of charged tracks in the signal is lower than other decays.

- The polar angle of the reconstructed B meson momentum in the CMS, $\cos\theta_B^*$. The distribution for the signal should be proportional to $\cos^2\theta_B^*$ because B meson pair is produced in a p-wave state, while that for the background is expected to be flat.

- In order to suppress background from $B^+ \to J/\psi K^{*+}$, which is the largest single source of background, we use following variables. With each charged track, we calculate the CMS momentum of B assuming three-body decay $B^+ \to J/\psi K_L \pi^+$ ($p_B^*(3)$) and invariant mass of K_L and the charged track ($M_{K_L\pi}$). If the $M_{K_L\pi}$ closest to the nominal K^* mass in the event satisfies $0.85 < M_{K_L\pi} < 0.93$ GeV/c^2, we use $p_B^*(3)$ and the momentum of charged tracks (p_π) as discriminant variables.

We use one million J/ψ inclusive MC events, which corresponds to about 370 fb^{-1} of $B \to J/\psi X$ decays, to determine the distribution of each variable. Figure 4.6 shows the distributions of likelihood variables for signal and background events.

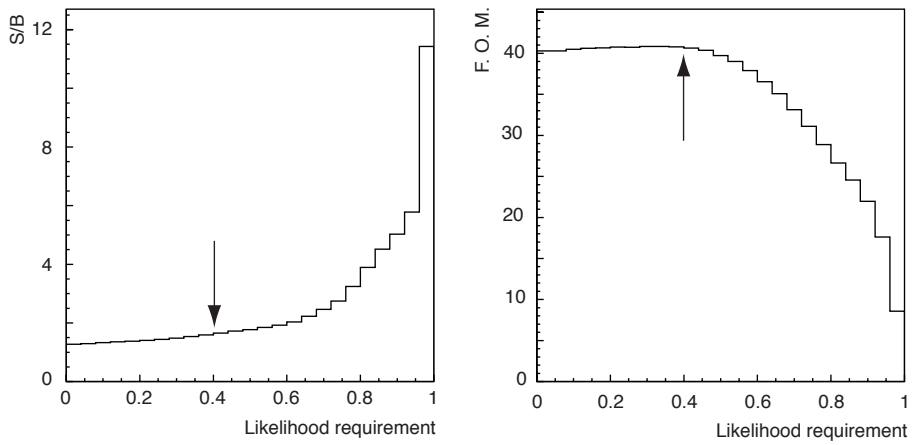

Figure 4.7: S/B ratio (left) and figure of merit (right) as a function of requirement value on the likelihood ratio L withJ/ψ inclusive MC. The arrows indicate the requirement value we use.

The combined likelihood is calculated by taking product of the above likelihood variables for signal and background. We take a likelihood ratio $L = L_{sig}/(L_{sig} + L_{bg})$ as a discriminant variable. We look for the optimal requirement on the L to maximize FOM ($=S/\sqrt{S+B}$) while keeping S/B ratio as high as possible. Figure 4.7 shows signal-to-background ratio (S/B) and FOM with J/ψ inclusive MC as a function of the requirement on the likelihood ratio L (L is required to be greater than these values). The FOM is rather flat up to $L = 0.4$, and decreases rapidly above 0.4, while S/B ratio keeps increasing. Based on the FOM distribution, we decide to require L greater than 0.4. Figure 4.8 shows the likelihood ratio distributions for the data and MC. The distributions for data and MC are in agreement.

The effectiveness of this likelihood approach is studied with MC. We compared the S/B ratio and the FOM between the likelihood-based selection described above and the square cut selection, where requirements are imposed on each variable instead of combined likelihood. The requirement for each variable is also optimized to maximize the FOM based on J/ψ inclusive MC. The likelihood approach enables us to improve S/B ratio from 1.42 to 1.66, while keeping the FOM almost the same level.

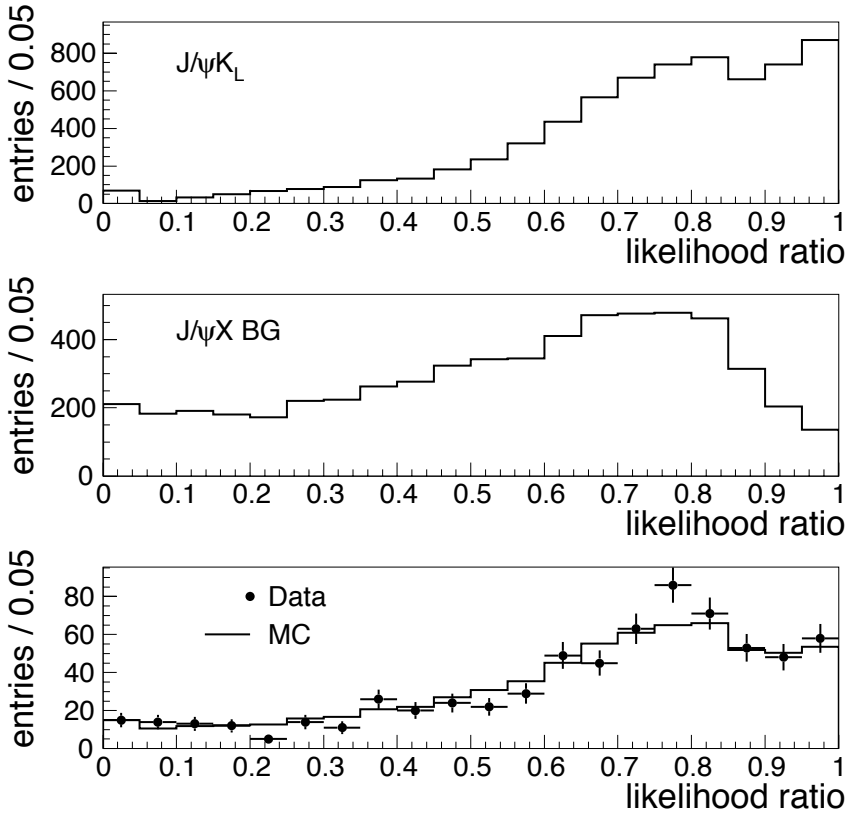

Figure 4.8: Distributions of $B^0 \to J/\psi\, K_L$ likelihood ratio. $J/\psi\, K_L$ MC (top), $J/\psi\, X$ background MC (middle), and the data together with MC events overlaid (bottom).

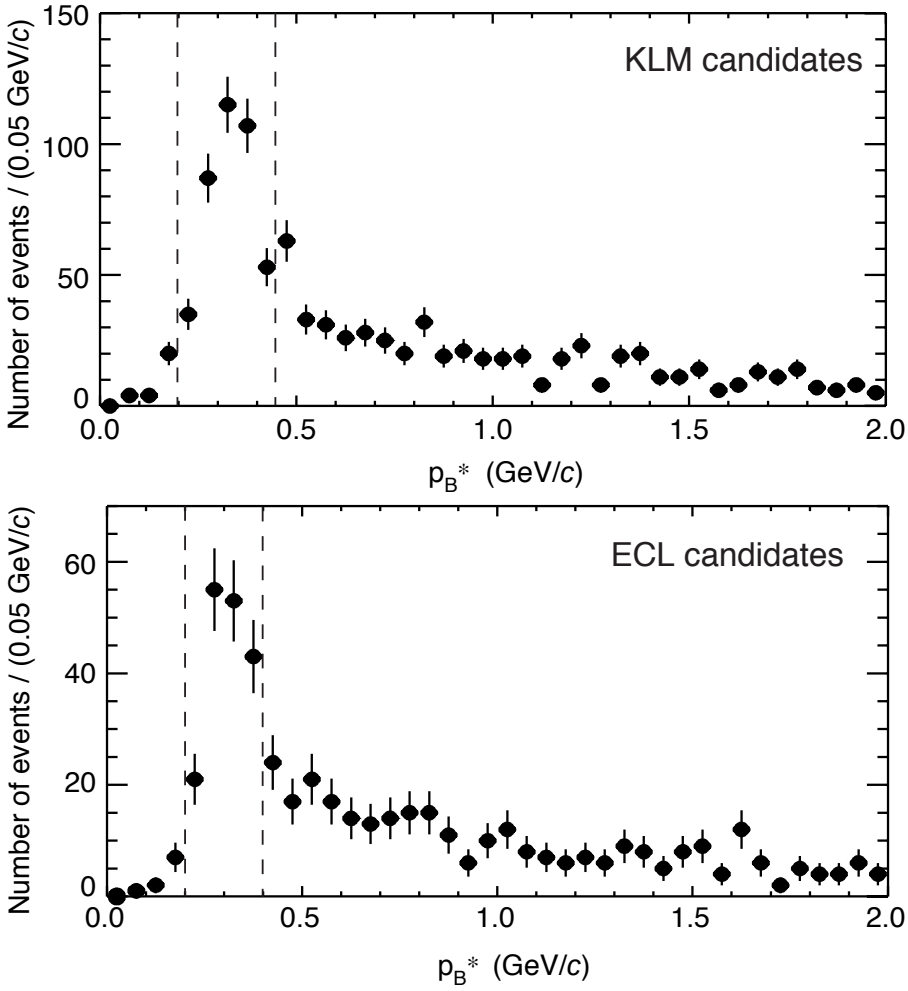

Figure 4.9: p_B^* distribution of $B^0 \to J/\psi K_L$ candidates in KLM candidates (top) and ECL candidates (bottom). The vertical lines show the signal region.

Finally, we require the CMS momentum of B^0, p_B^*, to be between 0.2 and 0.45 (0.40) GeV/c for KLM (ECL) candidates, where p_B^* is calculated from the K_L direction, the J/ψ momentum, and the mass of each particle assuming two-body decay of the B^0. The difference of higher-side requirement is due to the better angular resolution in the ECL than in the KLM and the difference of background composition. We optimize these requirements to maximize the FOM based on the MC study.

The reconstruction efficiency of $B^0 \to J/\psi K_L$ event is estimated to be 36.5% in MC, but we have significant defect in the K_L detection in the real data as explained later. Figure 4.9 shows the p_B^* distributions of $B^0 \to J/\psi K_L$ candidates with $0.0 < p_B^* < 2.0$ GeV/$c$in the real data. The distributions for KLM candidates and ECL

candidates are shown separately. We can see clear peaks around expected value of 330 MeV/c. 569 candidates remain after the selection, out of which 397 are KLM candidates and 172 are ECL candidates. An event display of $B^0 \to J/\psi\, K_L$ is shown in Fig. 4.10. The signal yield is estimated to be 346.3 ± 28.8 events, corresponding to the purity of 60.9%. The detail of the estimation procedure of the signal and the background fractions will be given in later section.

4.4 Reconstruction of $B^0 \to J/\psi\, K^{*0}$

In this section, we describe the reconstruction procedure of $B^0 \to J/\psi\, K^{*0}(K^{*0} \to K_S\pi^0)$ events. For this mode, we can reconstruct both J/ψ and K^{*0} with their four-momenta completely determined.

4.4.1 Reconstruction of J/ψ

The reconstruction procedure of J/ψ is almost the same as that in $J/\psi\, K_L$ analysis. The only difference is the invariant mass interval used to select J/ψ candidates: $2.95 < M_{ee} < 3.15$ GeV/c^2 and $3.05 < M_{\mu\mu} < 3.15$ GeV/c^2 are used for this mode. The asymmetric invariant mass interval for di-electron mode is chosen to accommodate the effect of the radiative tail.

4.4.2 Reconstruction of K^{*0}

K^{*0} candidates are reconstructed in the decay mode $K^{*0} \to K_S\pi^0$.

K_S candidates are reconstructed from oppositely-charged track pairs. Since the lifetime of K_S is about 90 psec, it travels order of cm before its decay ($c\tau \simeq 2.7$ cm). We select the oppositely-charged track pairs with the distance of the closest approach to each other is less than 300 μm. We require the reconstructed decay vertex is at least 1 mm apart from the interaction point. The ϕ coordinate of the vertex point and the ϕ direction of the momentum vector of K_S candidate must agree within 0.15 radian. This requirement is effective to reject combinatorial background. We also require the invariant mass of two pion system, $M_{\pi\pi}$ to be in the range $0.47 < M_{\pi\pi} < 0.52$ GeV/c^2. We fit the momenta of the pion pair with vertex constraint to improve the momentum resolution of K_S.

π^0 candidates are reconstructed from two clusters in the ECL with energy greater than 40 MeV. A photon pair with invariant mass in the range $0.125 < M_{\gamma\gamma} <$

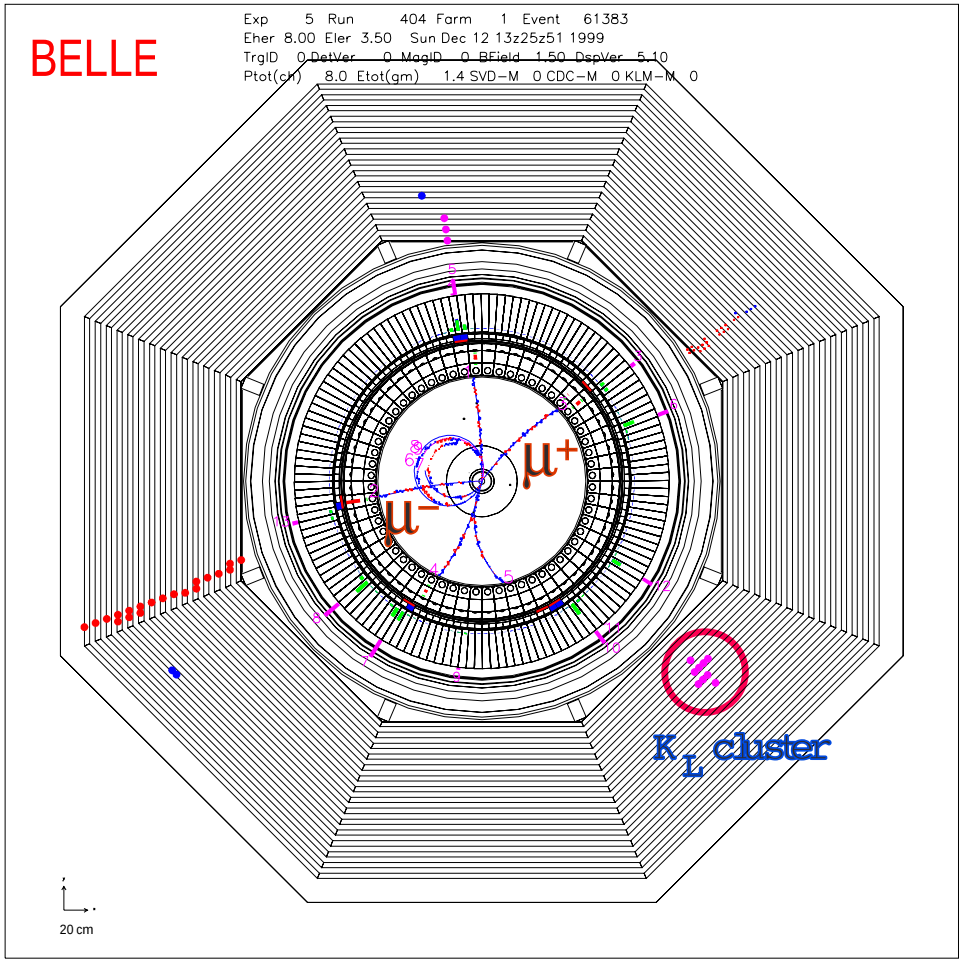

Figure 4.10: A display of typical $B^0 \to J/\psi K_L$ event.

0.145 GeV/c^2 is identified as a π^0. We fit the momenta of γ pairs with mass constraint to the nominal π^0 mass to improve the momentum resolution of the π^0.

To eliminate the low-momentum π^0 background, we require $\cos\theta_{K^*} < 0.8$, where θ_{K^*} is the angle of the K_S with respect to the K^* direction in the K^* rest frame. The distribution of θ_{K^*} makes peak at $\cos\theta_{K^*} = 1$ for low-momentum π^0 background, while it is flat for the signal.

Finally, we require the invariant mass of K_S-π^0 pairs to be within 75 MeV/c^2 of the nominal $K^*(892)$ mass [18].

4.4.3 Reconstruction of B^0

For $B^0 \to J/\psi K^{*0}$ reconstruction, we use following two variables to discriminate the signal and the background: the energy difference (ΔE) and the beam constrained mass ($M_{\rm bc}$).

The energy difference is defined as

$$\Delta E = E_B^* - E_{\rm beam}^*, \tag{4.6}$$

where E_B^* is the energy of reconstructed B meson in the center-of-mass system (CMS) and $E_{\rm beam}^*$ is the beam energy in the CMS. The resolution in ΔE is about 10 MeV, although the ΔE distribution for $B^0 \to J/\psi K^{*0}$ mode has long tail in the lower side due to the shower leakage from the ECL. This good energy resolution allows us to select the final states with little contamination from decays involving extra particles.

The beam constrained mass is defined as

$$M_{\rm bc} = \sqrt{{E_{\rm beam}^*}^2 - {p_B^*}^2}, \tag{4.7}$$

where p_B^* is the B candidate momentum in the CMS. By using $M_{\rm bc}$, one can take advantage of much better resolution in $E_{\rm beam}^*$ than the total energy from the individual reconstructed energies. The typical resolution in $M_{\rm bc}$ is 3 MeV/c^2.

Candidate B mesons are selected by requiring $5.27 < M_{\rm bc} < 5.29$ GeV/c^2 and $-50 < \Delta E < 30$ MeV. The asymmetric range of ΔE is chosen to take into account the shower leakage from the ECL.

The scatter plot of $M_{\rm bc}$ and ΔE is shown in Fig. 4.11, together with the projections onto each axis. We find 41 candidates in the signal box after the selection.

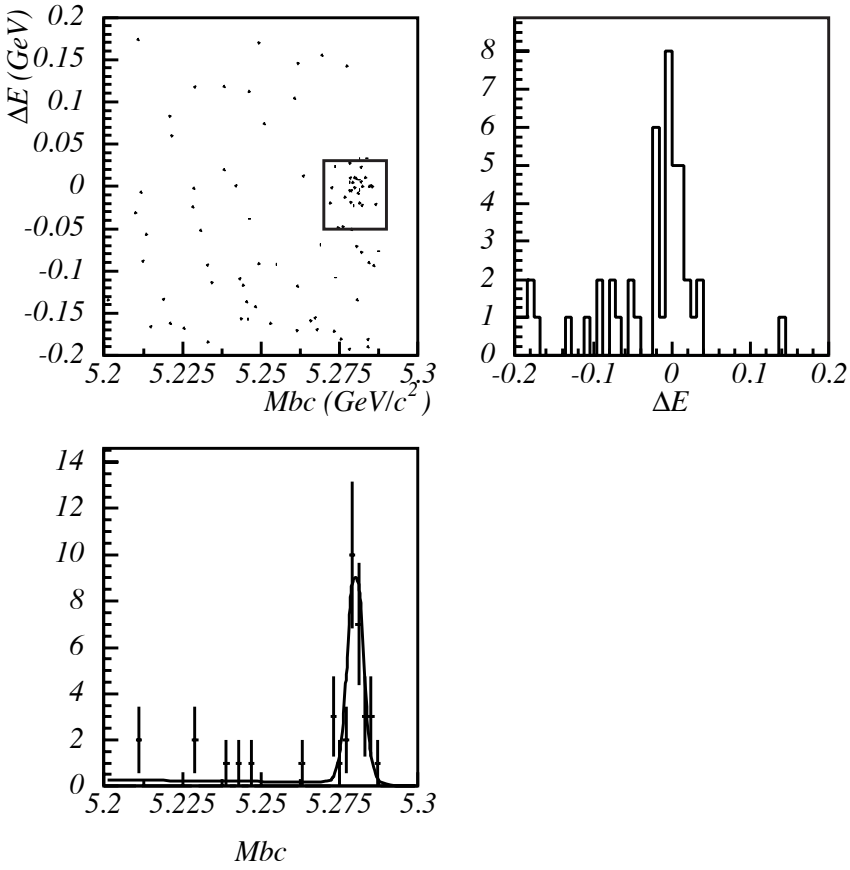

Figure 4.11: Scatter plot of ΔE *vs.* $M_{\rm bc}$ for $B^0 \to J/\psi K^{*0}(K^{*0} \to K_S\pi^0)$. The box represents the signal region.

Chapter 5

Flavor Tagging and Proper-Time Difference Reconstruction

In order to observe *CP* asymmetry, we need to determine the flavor and measure the proper-time difference of B meson decays. We also need to estimate the fraction of incorrect flavor determination and precision of vertex position measurement. In this chapter, we present the algorithms we use to obtain necessary inputs for $\sin 2\phi_1$ measurement.

5.1 Flavor Tagging

We need to determine the flavor of accompanying (or "tagging-side") B meson in order to measure the *CP* asymmetry. This procedure is called the *flavor tagging*.

Flavor tagging is realized by using flavor information carried by daughter particles of tagging-side B mesons such as:

- high momentum leptons from $b \to c\ell^-\nu$;
- lower momentum leptons from cascade decay, $b \to c \to s\ell^+\nu$;
- charged kaons and Λ from $b \to c \to s$. K^+ and Λ are likely to come from B^0, while K^- and $\overline{\Lambda}$ are from $\overline{B}^0$;
- high momentum pions from $B^0 \to D^{(*)-} + (\pi^+, \rho^+, a_1^+, \text{etc.})$;
- low-momentum pions from $D^{*-} \to \overline{D}^0\pi^-$ (which are called slow-pions).

Among these, the high-momentum lepton gives the simplest and cleanest information about the flavor of B meson. The information from the charge of kaon is also useful because of its large inclusive branching fraction ($B^0 \to K^+X = 78 \pm 8\%$ [18]). Contributions from other tagging information are less significant, but they are also important to maximize the performance of flavor tagging.

As described in Section 2.4.2, the performance of flavor tagging is determined by two parameters. The parameter ϵ is the probability that the b flavor can be identified, and w is the probability of wrong flavor assignment (so-called *wrong tagging fraction*). With realistic flavor tagging, we have correctly-tagged events with a fraction of $(1 - w)$ and wrongly-tagged events, which have opposite sign of asymmetry, with w. Then, the observed asymmetry becomes $(1 - 2w)\mathcal{A}$, where $\mathcal{A}$ is the true asymmetry. Since the statistical error of the measured asymmetry is proportional to $\epsilon^{-1/2}$, the statistical significance of the observed asymmetry with a certain number of events is proportional to the square-root of $\epsilon_{\rm eff} \equiv \epsilon(1-2w)^2$, which we call the *effective tagging efficiency*. We need to maximize $\epsilon_{\rm eff}$ to obtain maximum sensitivity. In following sections, we describe the detail of the flavor tagging method we have developed and its performance.

5.1.1 Flavor tagging method

In order to maximize the effective tagging efficiency, we need to combine all discriminants useful for flavor tagging in an efficient way. Especially, it is important to take into account the correlations between the discriminants. For example, it is possible that there are both high-momentum lepton and kaon in an event. By taking into account the correlation of the charge of these particles, we can assign more proper wrong tagging fraction for the event than neglecting the correlation (and just taking the information from the lepton, which has less wrong tagging fraction, for example). Thus we can increase the effective tagging efficiency.

However, it is practically difficult to consider all the correlations between all the information in an event, because the number of the particles in an event can be large. We have developed a flavor tagging method with two stages to avoid such a difficulty. In the first stage, we categorize tracks into four classes; lepton-like tracks, slow-pion-like tracks, Λ-like tracks, and kaon-like tracks. For each category, we prepare a look-up table that contains a list of Q, which is a measure of the flavor information carried by the track. The output value Q depends on the values of input discriminants such as the track momentum and the quality of particle identification.

The look-up tables are created with a large statistics MC sample. The output value Q for a certain cell of a look-up table is defined as

$$Q \equiv \frac{N(B^0) - N(\overline{B}^0)}{N(B^0) + N(\overline{B}^0)} \tag{5.1}$$

where $N(B^0)$ and $N(\overline{B}^0)$ are the number of B^0 and $\overline{B}^0$ which fall in the cell in MC sample. In this definition, the sign of Q, $q \equiv sign(Q)$ indicates whether the tagged B is B^0-like (when $q = +1$) or $\overline{B}^0$-like ($q = -1$). The absolute value of Q, $r \equiv |Q|$ is a measure of correctness of the flavor assignment, ranging from zero to one. If r is zero, it means there is no flavor discrimination, while r equal to one means perfect flavor assignment. If the statistics is large enough, the definition (5.1) become

$$Q \simeq P(B^0) - P(\overline{B}^0) = q \cdot (1 - 2w) \tag{5.2}$$

where $P(B^0)$ and $P(\overline{B}^0)$ are the probability that B^0 and $\overline{B}^0$ fall into this cell, respectively, and w is the wrong tag fraction for this cell.

In the second stage, the results from the four track categories are combined to assign the final output Q for the event. Again look-up table is prepared to provide Q.

In this way, we can summarize the information of tracks into outputs from four track categories and keep the sizes of look-up tables reasonably small. Figure 5.1 shows a schematic view of the flavor tagging method used in this analysis. We give the detailed description of each stages in the following sections.

5.1.1.1 Track level flavor tagging

The first (track-level) stage consists of four track classes; lepton track class, slow-pion track class, Λ track class, and kaon track class. Each track which is not used in CP eigenstate reconstruction is examined and passed to one of track classes. We use only well-reconstructed tracks satisfying $|dr|$ <2 cm and $|dz|$ <10 cm for flavor tagging, where dr and dz are distances from the nominal interaction point in r-ϕ plane and z direction, respectively. We reject a track if it forms K_S or photon conversion (an invariant mass being less than 100 MeV/c^2) with any opposite charged track.

lepton track class

The lepton track class consists of two sub-categories, electron-like track class and muon-like track class. If the momentum in the CMS momentum, p^*, of a track is

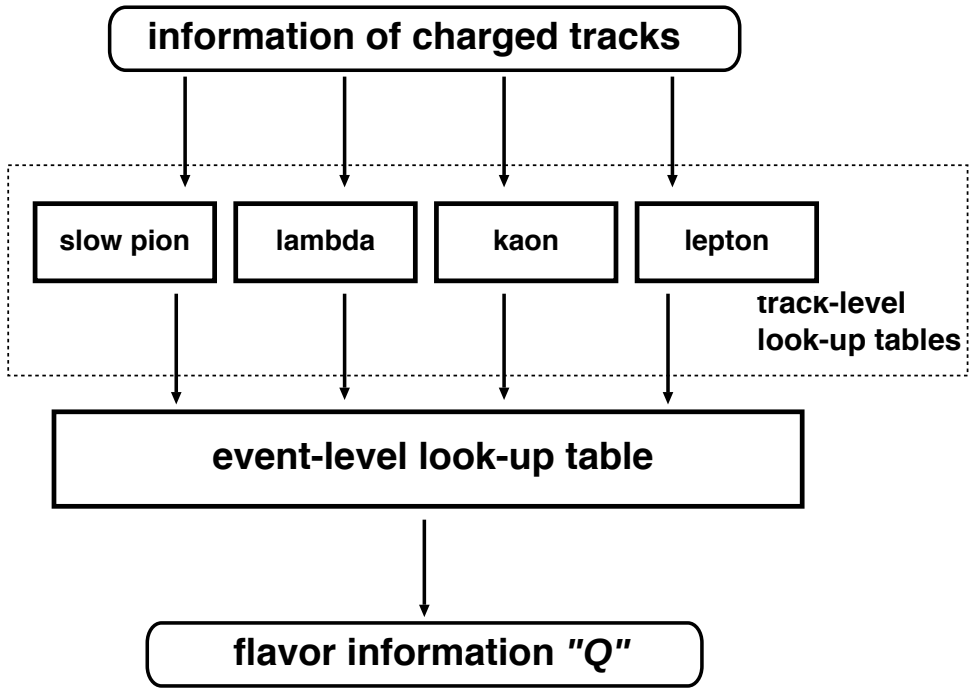

Figure 5.1: Schematic view of the flavor tagging.

larger than 0.4 GeV/c and identified as an electron (electron likelihood is greater than 0.8), it is passed to the electron-like track class. If p^* is larger than 0.8 GeV/c and identified as a muon (muon likelihood is greater than 0.95), it is passed to the muon-like track class.

The lepton track class takes care of high and middle momentum leptons and some fast pions. Because fast pions from *e.g.* $B^0 \to D^{(*)-} + \pi^+$ has the same charge correlation as high-momentum leptons from semi-leptonic decays, they are also treated in this class. Following variables are used to determine Q_i for each track:

- the CMS momentum (p^*);
- the polar angle in laboratory frame ($\theta_{\rm lab}$);
- the recoil mass $M_{\rm recoil}$, which is an invariant mass of the tag-side tracks except the "lepton" track;
- magnitude of missing momentum in the CMS, $p^*_{\rm miss}$;
- the lepton identification likelihood value ($L_{\rm lep}$);

- the charge of the track.

The p^* and $\theta_{\rm lab}$ are used to take into account the momentum and angular dependence of lepton identification performance. The p^*, $M_{\rm recoil}$ and $p^*_{\rm miss}$ are used to discriminate between the leptons from primary B decays and those from secondary D decays. The number of bins in this class is eleven for p^*, six for $\theta_{\rm lab}$, ten for $M_{\rm recoil}$, six for $p^*_{\rm miss}$, four for $L_{\rm lep}$, and two for the charge ($11 \times 6 \times 10 \times 6 \times 4 \times 2 = 31680$ in total). The output of the lepton-like track class, Q_l, is the Q_i of the track with highest $|Q_i|$ value among the tracks passed to this class.

slow-pion track class

If p^* of a track is less than 0.25 GeV/c and it is not identified as a kaon, that track is passed to the slow-pion track class. This category is intended to utilize the charge of low momentum pions from $D^{*\pm}$. The main background is low momentum pions from other decays and electrons from the photon conversion and $\pi^0 \to e^+e^-\gamma$.

The direction of slow pion follows that of D^*, due to its very low momentum in the rest frame of D^* (39 MeV/c). It then tends to follow the direction of thrust axis.[1] The cosine of the angle between the thrust axis and the track direction ($\cos\theta_{\rm thr}$), and the momentum and polar angle of π in the laboratory frame ($p_{\rm lab}$) are used to discriminate the slow pions from D^* decay and the others. $p_{\rm lab}$ is chosen to account for the angular dependence of low-momentum track finding efficiency. Electrons can be effectively separated from pions using dE/dx measured with CDC.

The discriminant variables for this class are:

- the momentum in the laboratory frame ($p_{\rm lab}$);
- the polar angle in laboratory frame ($\theta_{\rm lab}$);
- the electron likelihood value based on dE/dx ($L_{\rm eid}$);
- cosine of the angle between the thrust axis and the track direction ($\cos\theta_{\rm thr}$);
- the charge of the track.

The number of bins for each variable is ten for $p_{\rm lab}$, ten for $\theta_{\rm lab}$, five for $L_{\rm eid}$, and seven for $\cos\theta_{\rm thr}$ ($10 \times 10 \times 5 \times 7 \times 2 = 7000$). The output of the slow-pion track class, Q_π is the Q_i of the track with highest $|Q_i|$ value among the tracks passed to this class.

[1] The thrust axis is defied as the axis which maximizes the sum of all momenta projected onto it.

Λ and kaon track classes

If a track forms Λ with other track, it is passed to the Λ tack class. This class examines if the tracks are consistent with the decay $\Lambda \to p\pi^-$. The discriminant variables are:

- an invariant mass of the reconstructed Λ;
- an angle difference between the Λ momentum vector and the direction of Λ vertex from the IP;
- a distance in z-direction of two tracks at the Λ vertex;
- proton identification likelihood value;
- the flavor of Λ (Λ or $\overline{\Lambda}$).

The number of bins in this class is two for each variable.

If a track is not categorized into any of the above and is not positively identified as a proton, it is passed to the kaon track class.

The discriminant variables for this class are:

- number of K_S candidates in the event (N_{K_S});
- the momentum CMS (p^*);
- the polar angle in laboratory frame ($\theta_{\rm lab}$);
- the kaon identification likelihood value ($L_{\rm K}$);
- the charge of the track.

If there are K_S in a event, correlation of the kaon charge and the B-flavor is weaker than the case without K_S's. The next two variables are to take care of momentum and angular dependence of kaon identification performance. The number of bins for each variable is two for N_{K_S} (zero or not), 21 for p^*, 18 for $\theta_{\rm lab}$, and 13 for $L_{\rm K}$ ($21 \times 18 \times 13 \times 2 = 9828$ in total).

The output of the kaon and Λ track classes, $Q_{K/\Lambda}$, is given by the product of Q for all tracks assigned to these classes:

$$Q_{K/\Lambda} = \frac{\prod_i (1+Q_i) - \prod_i (1-Q_i)}{\prod_i (1+Q_i) + \prod_i (1-Q_i)}, \tag{5.3}$$

where the subscript i runs over all the tracks fed into the kaon and Λ track classes. The above $Q_{K/\Lambda}$ is designed to use the information from the sum of the strangeness. t provides better flavor-tagging performance than choosing the best candidate.

5.1.1.2 Event-level flavor tagging

The second (event-level) stage gathers the information from all track classes and outputs the final Q value for an event. The Q value for an event is determined from a look-up table, which is generated using an MC sample independent of that used to make the look-up tables for the track-level classes. The event-level look-up table takes three discriminant variables: the best "Q_i" values from lepton and slow-pion classes (Q_l and Q_π), and the product output from kaon and Λ classes ($Q_{K/\Lambda}$). The number of bins for each variable is 25 for Q_l, 19 for Q_π, and 35 for $Q_{K/\Lambda}$ (16625 in total). This look-up table can take into account the correlation among the outputs from different track classes.

5.1.1.3 Construction of look-up tables

We use ten million $B^0\overline{B}^0$ MC events to generate look-up tables. Four million events are used to make track-level look-up tables and the rest are used for event-level look-up tables.

For the track-level likelihood, we use a technique to compensate the limited statistics of the reference MC sample. When we calculate likelihood for each bin, we use information of neighbor bins with weights depending on the "distance" between the bins:

$$L_{ij\ldots} = \sum_l \sum_m \ldots \left\{ N_{lm\ldots} \exp\left(\frac{1}{2} \left[(i-l)^2 + (j-m)^2 + \ldots \right] \right) \right\}, \qquad (5.4)$$

where $ij\ldots$ are indices of a bin, $lm\ldots$ are indices of its neighbor bin, and $N_{lm\ldots}$ is the number of entries in the neighbor bin with sample MC.

For the event-level likelihood, we do not use this technique in order to avoid possible bias in the output value. We make separate tables for $q = +1$ and $q = -1$ cases, taking into account the possible charge dependence of input variables.

5.1.2 Performance

An MC study indicates the tagging efficiency, $\epsilon_{\rm tag}$, defined as a probability that we can obtain non-zero Q value for an event, is 99.6%. We use the $r(= |Q|)$ value, which is a measure of the wrong tagging fraction, just to categorize candidate events to six groups: $0.0 < r \leq 0.25$, $0.25 < r \leq 0.5$, $0.5 < r \leq 0.625$, $0.625 < r \leq 0.75$, $0.75 < r \leq 0.875$, and $0.875 < r \leq 1.0$. The total effective tagging efficiency is

Table 5.1: Summary of measured wrong-tag fractions (w_l) and event fractions (f_l) for each r interval. The errors include both statistical and systematic uncertainties.

l	r	f_l	w_l
1	$0.000 - 0.250$	0.405	$0.465^{+0.010}_{-0.009}$
2	$0.250 - 0.500$	0.149	$0.352^{+0.015}_{-0.014}$
3	$0.500 - 0.625$	0.081	$0.243^{+0.021}_{-0.030}$
4	$0.625 - 0.750$	0.099	$0.176^{+0.022}_{-0.017}$
5	$0.750 - 0.875$	0.123	$0.110^{+0.022}_{-0.014}$
6	$0.875 - 1.000$	0.140	$0.041^{+0.011}_{-0.010}$

defined as the sum of $\epsilon(1-2w)^2$ values in all six r intervals. An MC study shows it is 29.6%.

We determine the wrong tagging fraction (w_l) for each r interval from real data, where l is the region number. In this way the analysis is insulated from possible systematic differences between the MC simulation and the data due to imperfections in the modeling of the detector response, decay branching fractions, etc.

We use exclusively reconstructed, self-tagged $B^0 \to D^{*-}\ell^+\nu$, $D^-\pi^+$, $D^{*-}\pi^+$, and $D^{*-}\rho^+$ decays to determine w_l for the six r intervals. The procedure of wrong-tagging fraction measurement is summarized in Appendix C.

Table 5.1 shows the measured wrong tagging fractions (w_l) and event fractions (f_l) for each r interval ($l = 1, 6$). Because we find no significant difference in wrong-tagging fractions between $q = +1$ and $q = -1$ cases, we combine them and use same values for $q = +1$ and $q = -1$ samples. The overall and effective efficiency in the real data are found to be 99.7% and $(27.0 \pm 0.8(\text{stat})^{+0.6}_{-0.9}(\text{syst}))\%$, respectively.

5.2 Proper-Time Difference Reconstruction

We need to measure the proper-time difference of two B meson decays, Δt, to observe CP asymmetry. Since the momentum of B in the $\Upsilon(4S)$ rest frame is small, the decay distance along the beam direction ($\Delta z \equiv z_{CP} - z_{\text{tag}}$) can be used to extract $\sin 2\phi_1$, where z_{CP} and z_{tag} are the z coordinates of decay vertices of CP-side and tagging-side B mesons, respectively.

The impact parameter resolution of tracks without associated SVD hit is worse

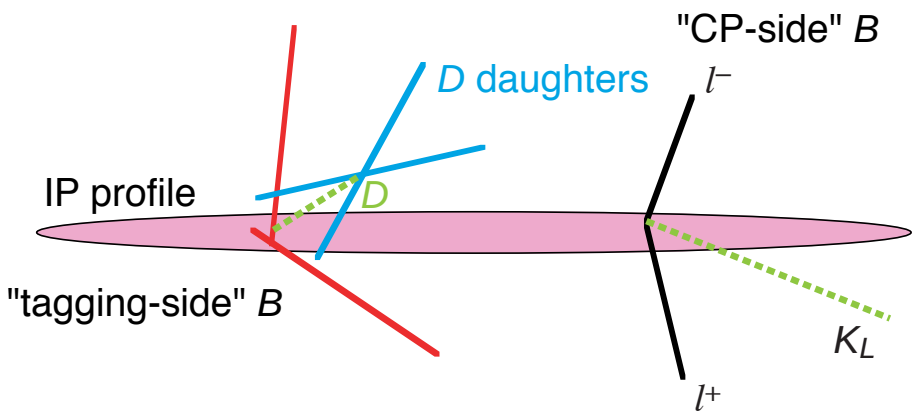

Figure 5.2: Schematic drawing of vertex reconstruction.

than tracks with SVD hits. Especially, the position resolution of the CDC in z direction ($\simeq 3$ mm) is not sufficient for the precise measurement of z vertex position. For reconstruction of the decay vertices of B_{CP} and $B_{\rm tag}$, we use only tracks with sufficient number of associated SVD hits, *i.e.* with both r-ϕ and z hits in at least one layer and with two or more z hits in total. The requirement of multiple hits in z direction rejects tracks with mis-associated SVD hits and ensures good impact parameter resolution. We use a constraint to the interaction point (IP) to improve vertex resolution further.

Since the decay length of J/ψ is negligibly small, we reconstruct the vertex positions for B_{CP} using leptons from J/ψ. For tagging side B mesons, significant fraction of events have charmed hadrons in the decay daughters, which have lifetimes of order ps and therefore deteriorate the vertex resolution. For this reason, the vertex reconstruction of $B_{\rm tag}$ is not straightforward as that of B_{CP} and determined based on a detailed study. Figure 5.2 shows a schematic view of vertex reconstruction. We describe the reconstruction procedure of each vertex position after the description of IP profile calculation.

5.2.1 Interaction point (IP) profile

We use the interaction point constraint to improve the position resolution of the vertex reconstruction. The average size and rotation angle of the interaction point are calculated offline for every accelerator fill from the primary vertex position distribution of hadronic events. The size of the interaction region in the horizontal direction (x) is calculated using the primary vertex position distribution; the effect

of the primary vertex resolution ($\simeq 70$ μm) is deconvoluted in this calculation. The vertex resolution is obtained using the primary vertex position distribution in the vertical direction (y), since the beam size in this direction is known to be less than 10 μm and the σ of the y direction is a good representation of the vertex resolution. The size of the interaction region in the y direction is determined from the special measurement by the KEKB accelerator group. The size of the interaction region is typically 100 μm in x, 3–5 μm in y, and 3–4 mm in z. Figure 5.3 shows the run dependence of the size of the interaction region. Uncertainty of $\sim$20 μm is added in quadrature to the interaction region size to account for the uncertainty in the B decay position due to the transverse motion of the B meson.

The nominal position of the interaction point is also calculated offline. Since we know IP position could move during an accelerator fill, we calculate the position of IP for every 60,000 events using hadronic events. Figure 5.4 shows the calculated IP position for a typical run.

5.2.2 Vertex reconstruction of B_{CP}

The vertex positions for B_{CP} are reconstructed using leptons from J/ψ. We use leptons only if there are sufficient number of SVD hits associated. We require the vertex position to be consistent with IP profile smeared in the r-ϕ plane by the B meson decay length. If only one lepton from J/ψ satisfies the SVD-hit requirement, we reconstruct CP-side vertex from that track and IP profile. In order to reject events with mis-reconstructed tracks, we require the reduced χ^2 (χ^2/n, n = number of degrees of freedom) of the vertex fit be less than 15. The efficiency of B_{CP} vertex reconstruction is measured to be 94.9% using $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}(K^{*0} \to K^+\pi^-)$ events in real data. The resolution for CP side vertex is estimated to be typically 75 μm (rms) using the MC.

5.2.3 Vertex reconstruction of $B_{\rm tag}$

The algorithm for the tagging-side vertex reconstruction is carefully chosen to minimize the effect of long-lived particles and poorly-reconstructed tracks. The effect of the secondary charm vertex moves the decay vertex point of the tagging-side B toward charm flight direction. It also degrade the vertex resolution. For this reason, the resolution of Δz measurement is dominated by that of tagging-side vertex. We use tracks which are not used for CP-side vertex reconstruction and satisfy the following requirements:

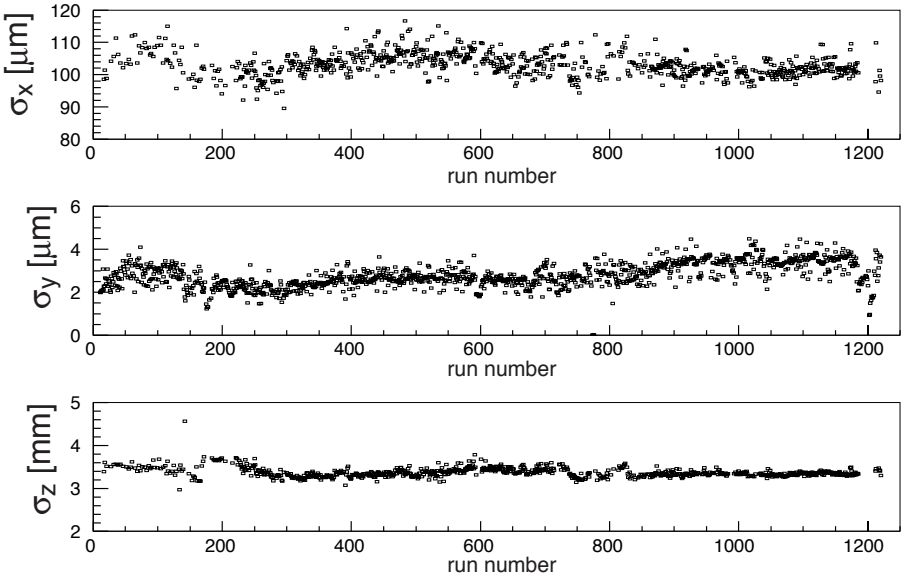

Figure 5.3: Run dependence of the size of interaction region.

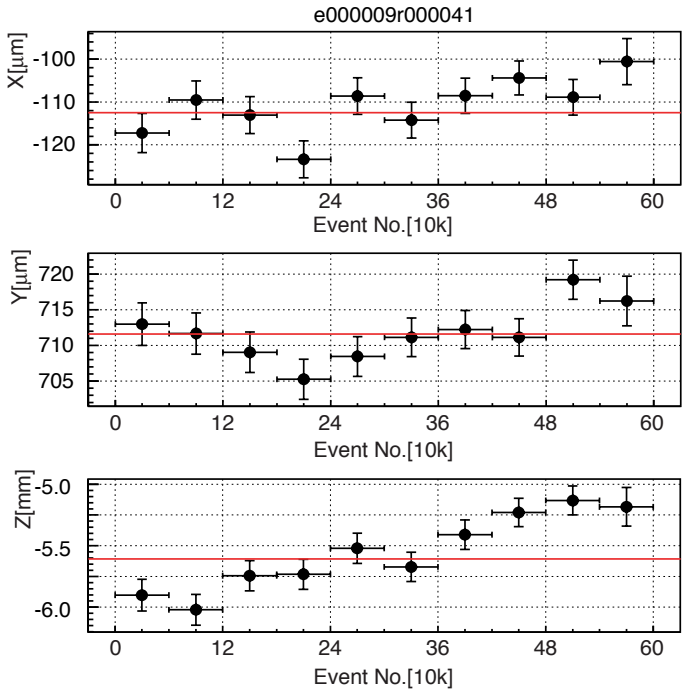

Figure 5.4: Calculated IP position for a typical run. The vertical bar at each point shows the error of each position measurement for every 60,000 events. The horizontal line shows the average IP position in the run.

- The track must be associated with SVD-hits (the same requirement as *CP*-side).
- The expected track error in z direction (σ_{dz}) must be less than 0.5 mm to eliminate poorly-reconstructed tracks.
- Tracks which form K_S candidates with other track are rejected since they do not originate from the B primary vertex.
- The K_S daughter tracks are further reduced by eliminating tracks with $\delta z >$ 1.8 mm, where δz is the z distance between the *CP*-side vertex position and the track position at closest approach to the *CP*-side vertex. The requirement is loose enough not to compromise the reconstruction efficiency while effectively rejecting K_S and poorly-reconstructed tracks.
- Tracks with impact parameter to the *CP*-side vertex position in r-ϕ plane (δr) greater than 0.5 mm are rejected to further reduce the K_S daughter.

If the χ^2/n of the vertex fit with IP constraint is worse than 20, the track that gives the largest contribution to the χ^2 is removed. In case such a track is a lepton with p_l^* >1.1 GeV/c, however, we keep the lepton and remove the track with the second worst χ^2 contribution, since high momentum lepton is very likely to originate from the B primary vertex. This procedure is iterated until the reduced χ^2 requirement is satisfied or only one track is left. The mean shift and resolution of the decay point for the $B_{\rm tag}$ are approximately 20 μm and 150 μm (rms), respectively. The $B_{\rm tag}$ vertex reconstruction efficiency is measured to be 92.6% using $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}(K^{*0} \to K^+\pi^-)$ events in real data.

We estimate the vertex resolution with signal MC sample. The distributions of the difference between reconstructed and generated vertex positions in z-direction ($z({\rm rec}) - z({\rm gen})$) for *CP*- and tagging-side vertex (z_{CP} and $z_{\rm tag}$) and $\Delta z = z_{CP} - z_{\rm tag}$ are shown in Fig. 5.5.

Because B meson pair is produced nearly at rest in the $\Upsilon(4S)$ center of mass system, the proper decay time difference of the two B decays can be defined, in good approximation,

$$\Delta t \simeq \frac{\Delta z}{(\beta\gamma)_{\Upsilon(4S)} c} \tag{5.5}$$

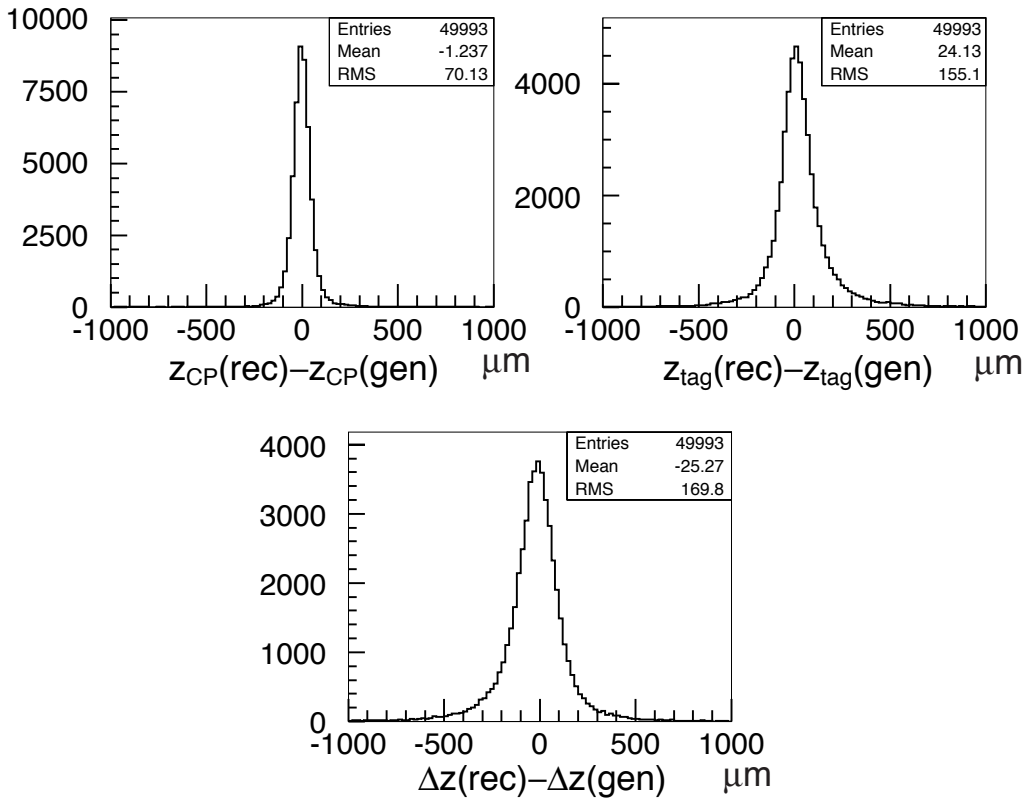

Figure 5.5: z(rec) − z(gen) distributions for z_{CP}(top-left), $z_{\rm tag}$(top-right), and Δz(bottom) with MC.

where $(\beta\gamma)_{\Upsilon(4S)} = 0.425$ is Lorentz factor of $\Upsilon(4S)$, and c is the speed of light. The effect of non-zero momentum of B meson in the $\Upsilon(4S)$ CMS ($p^*_B = 330$ MeV/c) will be discussed later.

After flavor tagging and vertex reconstruction, 523 and 36 candidate events remain for $B^0 \to J/\psi K_L$ and $B^0 \to J/\psi K^{*0}$ modes, respectively. These candidate events are used to extract $\sin 2\phi_1$ using an unbinned maximum likelihood method, as we discuss in the next chapter.

Chapter 6

Measurement of $\sin 2\phi_1$

We have described how we reconstruct B^0 decays into CP eigenstates, determine the flavor of associated B meson, and measure the proper-time difference of two B mesons from $\Upsilon(4S)$ decays. The last step is to determine what value of $\sin 2\phi_1$ is most consistent with the data. In order to make maximal use of the available statistics, we use a method of unbinned maximum likelihood fit.

In this chapter, we first briefly introduce the maximum likelihood fit method. Then, we describe the detailed procedure of the fit and present the result. We first discuss $J/\psi K_L$ mode. The fit procedure and the result of $J/\psi K^{*0}$ mode are presented in the last section.

6.1 Maximum Likelihood Fit

6.1.1 Method of maximum likelihood

Suppose a set of independently measured quantities x_i came from a known *probability density function* (PDF) $P(x;\boldsymbol{\alpha})$; *i.e.* the probability of observing x between x and $x+dx$ is written as $P(x;\boldsymbol{\alpha})dx$, where $\boldsymbol{\alpha}$ is an unknown set of parameters. The problem is to estimate the most probable values of $\boldsymbol{\alpha}$ from a set of actual observations. We define *likelihood function* $\mathcal{L}$ as

$$\mathcal{L}(\boldsymbol{\alpha}) = \prod_{i=1}^{n} P(x_i;\boldsymbol{\alpha}), \tag{6.1}$$

where n is the number of measurements performed. This likelihood function can be considered as a joint probability density of getting a particular experimental result,

$x_1, \ldots, x_n$, assuming $P(x; \boldsymbol{\alpha})$ is the true *normalized* distribution function;

$$\int P(x; \boldsymbol{\alpha}) dx = 1. \tag{6.2}$$

Then, the most probable values of $\boldsymbol{\alpha}$, $\hat{\boldsymbol{\alpha}}$, are the ones which maximize $\mathcal{L}$. It can be proved that in the limit of large n, $\hat{\boldsymbol{\alpha}}$ approaches the true value of $\boldsymbol{\alpha}$, and there is no other method of estimation that is more accurate [55].

It is usually easier to work with the logarithm of $\mathcal{L}$, $\ln \mathcal{L}$, than with $\mathcal{L}$ itself, since $\mathcal{L}$ can easily become too small to be handled with computers, while both are maximized with the same set of $\boldsymbol{\alpha}$. Using the logarithm has the additional advantage that the quantity $-2 \ln \mathcal{L}$ behaves like a χ^2 in the limit of large n. Then the s-standard-deviation statistical errors are determined from the contour given by the $\boldsymbol{\alpha}'$ such that

$$-2 \ln \mathcal{L}(\boldsymbol{\alpha}') = -2 \ln \mathcal{L}_{\max} + s^2, \tag{6.3}$$

where $\mathcal{L}_{\max}$ is the value of $\mathcal{L}$ at the maximum point. Therefore we can estimate the statistical error by looking for $\boldsymbol{\alpha}'$ which gives $-2 \ln \mathcal{L}$ larger than $-2 \ln \mathcal{L}_{\max}$ by one. We can use general numerical minimization package such as the CERN program MINUIT [56] to find the values of $\hat{\boldsymbol{\alpha}}$ and to estimate their errors.

6.1.2 Application to this analysis

In our case, the unknown parameter to be determined from the fit is $\sin 2\phi_1$. Although we know the distribution predicted by the theory, we cannot simply use it as the PDF for each event, because what we can observe is the distribution diluted and smeared with the finite experimental accuracy. Therefore, we have to construct the PDF which properly takes into account the response of the detector.

We may use an identical set of parameters (such as wrong tagging fraction) for all the events to describe the effect of the finite detector resolution. In this case, the parameters represent averaged detector performance for all kind of events. However, in the data there should exist events with very clean information (*e.g.* unambiguous flavor tagging, small vertex error etc.) while there also be events with poor information. The approach with averaged detector response cannot make use of such information. We take into account this information by calculating parameters to represent the detector response based on the event-by-event information. With this approach, we can increase the statistical significance of the measurement.

There are three major sources that affect the measurement of *CP* asymmetry. The first thing we need to consider is the imperfection of flavor tagging. We can

take it into account by replacing $\sin 2\phi_1$ with $(1-2w)\sin 2\phi_1$ in the PDF, where w is the wrong tagging fraction for an event. We use w_l, the measured wrong tagging fraction for corresponding r assigned by flavor tagging, as the wrong tagging fraction for the event. We use six r intervals as listed in Table 5.1. We may use finer interval of r in order to make more use of tagging quality information, but it would be very difficult in that case to control the systematic uncertainties related to the flavor tagging with limited statistics of the data.

The second item is the effect of the finite vertex resolution. By introducing the resolution function, $R(\Delta t)$, which represents the uncertainty of the vertex position differences measured by the SVD and effects of non-primary tracks such as charm daughters, we can define the PDF P as:

$$P(\Delta t) = \int_{-\infty}^{+\infty} \mathcal{P}(\Delta t') \cdot R(\Delta t - \Delta t') d(\Delta t'), \tag{6.4}$$

where Δt is the measured proper-time distribution for the event and $\mathcal{P}(\Delta t)$ is the theoretical distribution of Δt. We use an event-by-event resolution function $R^i(\Delta t)$ calculated based on the errors of the vertex positions assigned by the kinematic fitting program and the quality of vertex reconstruction.

The last effect we consider is the existence of background in the data. We can treat it by defining the overall PDF P as

$$P = f_{\rm sig} P_{\rm sig} + (1 - f_{\rm sig}) P_{\rm bkg}, \tag{6.5}$$

where $f_{\rm sig}$ is the probability that the event is signal, $P_{\rm sig}$ is the PDF for the signal and $P_{\rm bkg}$ is the PDF for the background. The signal probability $f_{\rm sig}$ is calculated event-by-event based on kinematical variables. For the case of $B^0 \to J/\psi K_L$, we use p_B^* to calculate the signal probability. For $B^0 \to J/\psi K^{*0}$, we use ΔE and $M_{\rm bc}$ to calculate the signal probability. The background PDF, $P_{\rm bkg}$ represents the distribution of the background. We need to choose appropriate representations of the background components not to bias the fit result.

In the following sections, we describe the construction of the PDF we use. We first discuss the signal PDF and the resolution function, and then explain the background PDF.

6.2 Signal PDF

The theoretical distribution expected for the signal events in $B^0 \to J/\psi K_L$ decay is given by

$$\mathcal{P}_{\rm sig}(\Delta t, q, w_l, \xi_f; \sin 2\phi_1) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}} A_{cp}(\Delta t, q, w_l, \xi_f; \sin 2\phi_1), \tag{6.6}$$

where

$$A_{cp}(\Delta t, q, w_l, \xi_f; \sin 2\phi_1) = 1 - \xi_f q(1-2w_l)\sin 2\phi_1 \sin(\Delta m_d \Delta t) \tag{6.7}$$

and $\xi_f = +1$ for $J/\psi K_L$ channel, q is the sign of flavor tagging output Q and $q = +1$ (-1) if the tagged B is B^0 $(\overline{B}^0)$, and w_l is the wrong-tagging fraction for $r = |Q|$ of the event.

We need to take into account the finite proper time resolution by convoluting $\mathcal{P}_{\rm sig}$ with the resolution function $R_{\rm sig}$. Then, the signal PDF $P^i_{\rm sig}$ for i-th event is defined as:

$$P_{\rm sig}(\Delta t_i; \sin 2\phi_1) = \int_{-\infty}^{+\infty} \mathcal{P}_{\rm sig}(\Delta t', q, w_l, \xi_f; \sin 2\phi_1) \cdot R_{\rm sig}(\Delta t_i - \Delta t') d(\Delta t'), \tag{6.8}$$

where Δt_i is the measured Δt for i-th event.

6.2.1 Resolution function

The resolution function $R_{\rm sig}(\Delta t_i - \Delta t')$ is parametrized as a sum of two Gaussian functions: a *main* component due to the SVD vertex resolution, charmed meson lifetimes, and the effect of the motion of the B mesons in the $\Upsilon(4S)$ CMS, and a *tail* component caused by poorly reconstructed tracks.

$R_{\rm sig}(\Delta t_i - \Delta t')$ is expressed as

$$\begin{aligned} R_{\rm sig}(\Delta t_i - \Delta t') &= (1 - f_{\rm tail}) \cdot G(\Delta t_i - \Delta t'; \mu_{\Delta t}, \sigma_{\Delta t}) \\ &+ f_{\rm tail} \cdot G(\Delta t_i - \Delta t'; \mu^{\rm tail}_{\Delta t}, \sigma^{\rm tail}_{\Delta t}), \end{aligned} \tag{6.9}$$

where $f_{\rm tail}$ is the fraction of the tail part of the resolution function and G is the Gaussian function:

$$G(t; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right]. \tag{6.10}$$

$\sigma_{\Delta t}$, $\sigma^{\rm tail}_{\Delta t}$, $\mu_{\Delta t}$ and $\mu^{\rm tail}_{\Delta t}$ are the proper-time difference resolutions and the mean shift of the proper-time difference for the main part and the tail part of the resolution function, respectively. $\mu_{\Delta t}$ and $\mu^{\rm tail}_{\Delta t}$ originate from the mean shift of the Δz_i

measurement due to contamination from the charm daughters in the vertex reconstruction of the tagging-side B meson as we discuss below. $f_{\rm tail}$ is determined to be 0.03 ± 0.02 by the lifetime analysis of the $B \to J/\psi K$ ($K = K^{\pm}, K_S$) sample using the same resolution function. A detailed description of the lifetime analysis for $B \to J/\psi K$ samples can be found in [57].

smearing due to kinematic approximation

The proper-time difference resolution $\sigma_{\Delta t}$ (and $\sigma_{\Delta t}^{\rm tail}$) is calculated as a convolution of the event by event Δz_i resolution $\sigma_{\Delta z}$ and the error $\sigma_{\rm K}$ due to the kinematic approximation ($\Delta t_i \approx \frac{\Delta z_i}{c(\beta\gamma)_\Upsilon}$, $(\beta\gamma)_\Upsilon = \frac{p_z(\Upsilon)}{m(\Upsilon)c}$):

$$\begin{aligned} \sigma_{\Delta t} &= \sqrt{(\frac{\sigma_{\Delta z}}{c(\beta\gamma)_\Upsilon})^2 + \sigma_{\rm K}^2}, \\ \sigma_{\Delta t}^{\rm tail} &= \sqrt{(\frac{\sigma_{\Delta z}^{\rm tail}}{c(\beta\gamma)_\Upsilon})^2 + (\sigma_{\rm K}^{\rm tail})^2}. \end{aligned} \tag{6.11}$$

The $\sigma_{\rm K}$ and $\sigma_{\rm K}^{\rm tail}$ values are determined to be $\sigma_{\rm K} = 0.287 \pm 0.004$ ps and $\sigma_{\rm K}^{\rm tail} = 0.32 \pm 0.19$ ps using the MC simulation since these parameters are independent of the detector performance.

detector resolution and smearing due to charm daughters

The Δz_i resolution $\sigma_{\Delta z}$ ($\sigma_{\Delta z}^{\rm tail}$) is calculated from the vertex resolutions of the two B meson decays (CP-side and tagging side), σ_z^{CP} and $\sigma_z^{\rm tag}$:

$$\begin{aligned} \sigma_{\Delta z}^2 &= S_{\rm det}^2(\sigma_z^{CP})^2 + (S_{\rm det}^2 + S_{\rm charm}^2)(\sigma_z^{\rm tag})^2, \\ (\sigma_{\Delta z}^{\rm tail})^2 &= (S_{\rm det}^{\rm tail})^2(\sigma_z^{CP})^2 + \{(S_{\rm det}^{\rm tail})^2 + (S_{\rm charm}^{\rm tail})^2\}(\sigma_z^{\rm tag})^2, \end{aligned} \tag{6.12}$$

where $S_{\rm det}$ and $S_{\rm det}^{\rm tail}$ are the global scaling factors to account for the systematic uncertainties in the vertex resolutions σ_z^{CP} and $\sigma_z^{\rm tag}$ computed from the track helix errors in the vertex-fit, and $S_{\rm charm}$ and $S_{\rm charm}^{\rm tail}$ are scaling factors to account for the degradation of the vertex resolution of the tagging side B meson due to contamination from the charm daughters.

The $S_{\rm charm}$ and $S_{\rm charm}^{\rm tail}$ values are determined to be $S_{\rm charm} = 0.59 \pm 0.01$ and $S_{\rm charm}^{\rm tail} = 2.16 \pm 0.10$ using the MC simulation. The $S_{\rm det}$ and $S_{\rm det}^{\rm tail}$ values must be determined from the data as these parameters depend on the detector performance. $S_{\rm det}$ is determined using $D^0 \to K^+\pi^-$ sample because we can use a large amount of clean sample in this mode. Production point of the D^0 is obtained,

with the IP constraint, from the primary tracks in the hemisphere that contains the D^0 candidate. Distance between the D^0 decay vertex and production vertex in z direction is fit with the same resolution function and known D^0 lifetime to obtain the $S_{\rm det}$. The $S_{\rm det}$ value is measured to be $S_{\rm det} = 0.88 \pm 0.01$ from the data and $S_{\rm det} = 1.05 \pm 0.01$ from the MC simulation using the D^0 sample. Since we find $S_{\rm det} = 1.035 \pm 0.003$ for $B \to J/\psi K$ MC sample, we fix the $S_{\rm det}$ value to be $S_{\rm det} = (0.88 \pm 0.01) \cdot (1.035 \pm 0.003)/(1.05 \pm 0.01) = 0.86 \pm 0.01$ for the data. $S_{\rm det}^{\rm tail}$ is determined to be 3.51 ± 0.88 by the lifetime analysis of the $B \to J/\psi K$ sample.

effect of poorly reconstructed vertices

A small fraction of events have a large reduced χ^2 (χ^2/n, n = number of degrees of freedom) in the kinematical vertex fit. We have found that the vertex error computed from track helix errors in the kinematical fit underestimates the vertex resolution and the vertex with larger χ^2 has worse resolution. In order to take into account this effect, we introduce the effective vertex resolutions $\tilde{\sigma}_z^{CP}$ and $\tilde{\sigma}_z^{\rm tag}$ when χ^2/n is greater than three:

$$\begin{aligned}(\tilde{\sigma}_z^{CP})^2 &\equiv [1 + \alpha_{CP}\{(\chi^2/n)_{CP} - 3\}](\sigma_z^{CP})^2 : \ (\chi^2/n)_{CP} > 3, \\ (\tilde{\sigma}_z^{\rm tag})^2 &\equiv [1 + \alpha_{\rm tag}\{(\chi^2/n)_{tag} - 3\}](\sigma_z^{\rm tag})^2 : \ (\chi^2/n)_{tag} > 3, \end{aligned} \tag{6.13}$$

where $(\chi^2/n)_{CP}$ and $(\chi^2/n)_{tag}$ are reduced χ^2 of vertex-fits for the CP - and tagging-side B decay vertices, respectively. We find $\alpha_{CP} = 1.02 \pm 0.03$ and $\alpha_{\rm tag} = 1.64 \pm 0.05$ from the MC simulation.

mean shift due to charm daughters

As is mentioned above, the $\mu_{\Delta t}$ (and $\mu_{\Delta t}^{\rm tail}$) originates from the mean shift of the Δz_i measurements $\mu_{\Delta z}$ (and $\mu_{\Delta z}^{\rm tail}$):

$$\begin{aligned}\mu_{\Delta t} &\equiv \frac{\mu_{\Delta z}}{c(\beta\gamma)_\Upsilon}, \\ \mu_{\Delta t}^{\rm tail} &\equiv \frac{\mu_{\Delta z}^{\rm tail}}{c(\beta\gamma)_\Upsilon}. \end{aligned} \tag{6.14}$$

The mean shifts of the Δz_i, $\mu_{\Delta z}$ and $\mu_{\Delta z}^{\rm tail}$, are caused by the contamination from the charm daughters in the vertex reconstruction of the tagging side B meson and are correlated with the $\sigma_z^{\rm tag}$:

$$\begin{aligned}\mu_{\Delta z}(\sigma_z^{\rm tag}) &= \mu_0 + \alpha_\mu \sigma_z^{\rm tag}, \\ \mu_{\Delta z}^{\rm tail}(\sigma_z^{\rm tag}) &= \mu_0^{\rm tail} + \alpha_\mu^{\rm tail} \sigma_z^{\rm tag}. \end{aligned} \tag{6.15}$$

Table 6.1: Parameters of resolution function.

Fit parameters	Fit values
$f_{\rm tail}$	0.03 ± 0.02
$\sigma_{\rm K}$ (ps)	0.287 ± 0.004
$\sigma_{\rm K}^{\rm tail}$ (ps)	0.32 ± 0.19
$S_{\rm det}$	0.86 ± 0.01
$S_{\rm det}^{\rm tail}$	3.51 ± 0.88
$S_{\rm charm}$	0.59 ± 0.01
$S_{\rm charm}^{\rm tail}$	2.16 ± 0.10
α_{CP}	1.02 ± 0.03
$\alpha_{\rm tag}$	1.64 ± 0.05
μ_0 (μm)	-21.4 ± 3.7
α_μ	-0.10 ± 0.01
$\mu_0^{\rm tail}$ (μm)	151 ± 128
$\alpha_\mu^{\rm tail}$	-1.42 ± 0.17

The α_μ and $\alpha_\mu^{\rm tail}$ values are determined using the MC simulation. The μ_0 and $\mu_{\Delta z}^{\rm tail}$ are determined from real data. The resulting values are $\mu_0 = -21.4 \pm 3.7$ μm, $\alpha_\mu = -0.10 \pm 0.01$, $\mu_0^{\rm tail} = 151 \pm 128$ μm, and $\alpha_\mu^{\rm tail} = -1.42 \pm 0.17$

The parameters of the resolution function are summarized in Table 6.1.

Figure 6.1(left) shows the $\Delta t_{\rm rec} - \Delta t_{\rm gen}$ distribution for the MC signal events and the resolution function, where $\Delta t_{\rm rec}$ and $\Delta t_{\rm gen}$ is the reconstructed and true proper-time differences, respectively. The resolution function is obtained by summing an event-by-event resolution function. The distribution is well-represented by the resolution function. The resolution function obtained from $J/\psi K_L$ data is shown in Fig 6.1(right). The typical values of parameters obtained from real data are: $\mu_{\rm main}$= -0.24 ps, $\mu_{\rm tail}$= 0.18 ps, $\sigma_{\rm main}$= 1.49 ps, $\sigma_{\rm tail}$= 3.85 ps.

As a validation of our understanding of the resolution function, we measure the lifetimes of B^0 and $B^\pm$ mesons with the same data set. We apply the same vertex reconstruction algorithm for fully reconstructed decays as the *CP*-eigenstate decays and the rest of tracks as the tag side. An exponential PDF convoluted with the same Δt resolution function as the *CP* eigenstates fit are used. Combining

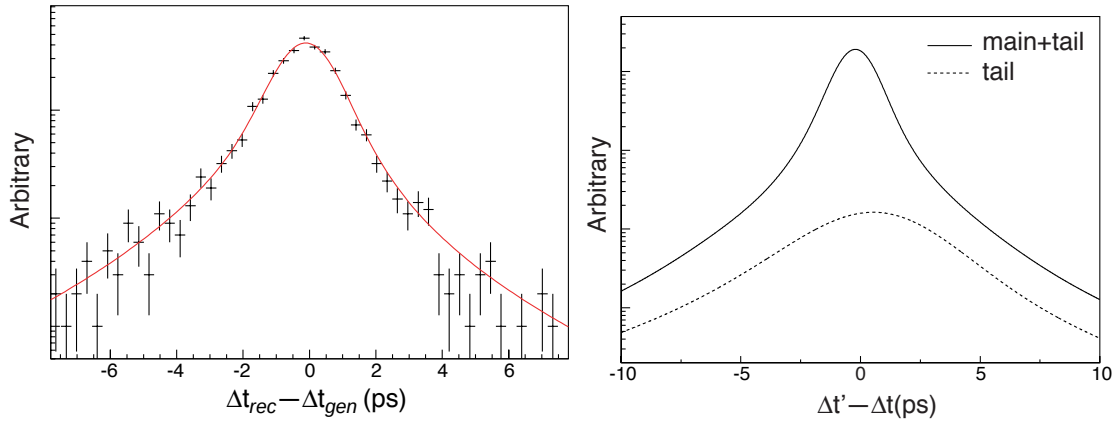

Figure 6.1: The $\Delta t_{\rm rec} - \Delta t_{\rm gen}$ distributions for the MC signal events together with resolution function (left) and average resolution function from $J/\psi K_L$ data (right).

$B^0 \to J/\psi K^{*0}(K^+\pi^-)$, $D^{(*)-}\pi^+$, $D^-\rho^+$ and $D^{*-}\ell^+\nu$ decay modes,[1] the B^0 lifetime is measured to be $\tau_{B^0} = (1.547 \pm 0.021)$ ps. Figure 6.2 shows the Δt distribution of $B^0 \to D^{*-}\ell^+\nu$ events together with fit curve. Similarly, combining $B^\pm \to J/\psi K^\pm$ and $D^0\pi^\pm$, we obtain $\tau_{B^\pm} = (1.641 \pm 0.033)$ ps. The errors are statistical only. These results are consistent with the world average values, $\tau_{B^0} = (1.548 \pm 0.032)$ ps and $\tau_{B^\pm} = (1.653 \pm 0.028)$ ps. These results verify the validity of the resolution function we use.

6.3 Background PDF

The PDF for the background is written in a way similar to the signal,

$$P_{\rm bkg}(\Delta t_i) = \sum_j P^j_{\rm bkg}(\Delta t_i) = \sum_j \int_{-\infty}^{+\infty} \left\{ f^j_{\rm bkg} \cdot \mathcal{P}^j_{\rm bkg}(\Delta t') \cdot R^j_{\rm bkg}(\Delta t_i - \Delta t') \right\} d(\Delta t'), \tag{6.16}$$

where $f^j_{\rm bkg}$ is the probability that the event is the background in j-th category. j runs over all the background components. The $R_{\rm bkg}$ is not necessarily the exact model of the vertex resolution for the background in which the J/ψ candidate is formed with a random combination of charged particles (we call this type of background "combinatorial background" hereafter). It is essential that $P^i_{\rm bkg}(\Delta t)$ represents the proper-time distribution as a whole with sufficient precision.

[1] The selection procedure of these candidates is given in Appendix C.

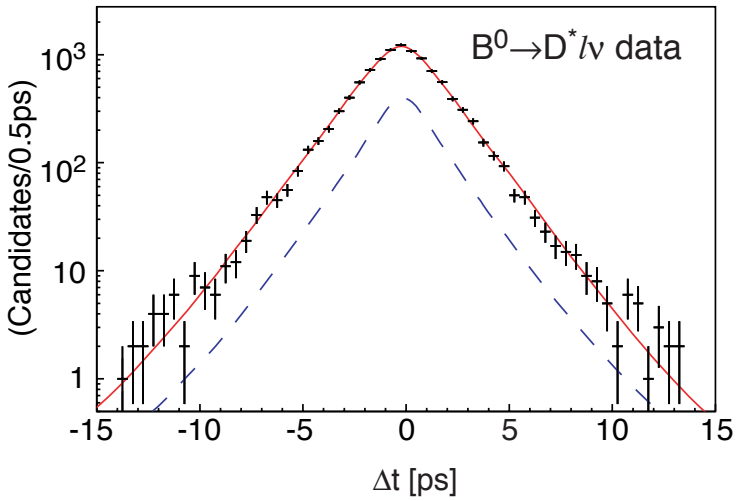

Figure 6.2: The Δt distribution for $B^0 \to D^{*-}\ell^+\nu$ events together with fit curves. Solid (dashed) line represents the distribution of signal plus background (background only).

The background in the $J/\psi K_L$ mode is dominated by $B \to J/\psi X$ decays including the decays into CP eigenstates that need to be treated differently from non-CP eigenstates. Taking into account the difference of Δz distribution, the $\mathcal{P}_{\rm bkg}$ for $J/\psi K_L$ mode is determined separately for each of the following background components.

1. *CP* eigenstates There exist some CP eigenstates in the background. They can be categorized in three sub-classes:

(a) $J/\psi K^{*0}(K_L\pi^0)$, which is a mixture of CP even and odd state;

(b) $\xi_f = +1$ (CP even) CP modes ($\psi(2S)K_L$, $\chi_{c1}K_L$ and $J/\psi\pi^0$);

(c) $\xi_f = -1$ (CP odd) CP modes ($J/\psi K_S$).

For these classes of background, we use the signal PDF given in Equation (6.6) with the appropriate ξ_f values as $\mathcal{P}_{\rm bkg}$. For $J/\psi K^{*0}(K_L\pi^0)$ mode, which is a mixture of $\xi_f = -1$ ($\sim 81\%$) and $\xi_f = +1$ ($\sim 19\%$) states, we use the net CP eigenvalue of $\xi_{\psi K^{*0}} = -0.62$ based on the measurement at Belle [21].

For background from CP eigenstate, we use the signal resolution function $R_{\rm sig}$ as $R_{\rm bkg}$ because for these backgrounds, both z_{CP} and $z_{\rm tag}$ are reconstructed by proper combinations of tracks.

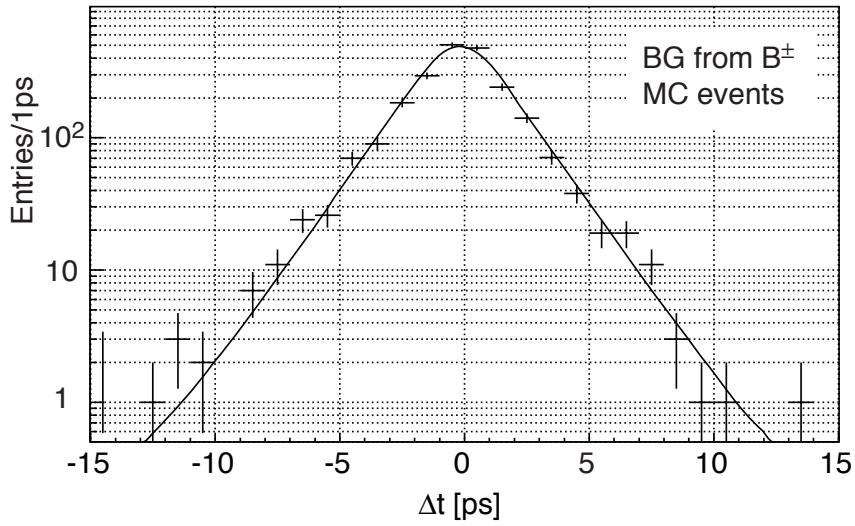

Figure 6.3: Δt distributions of $B^{\pm}$ background events in MC. The result of fit is superimposed.

2. other $B \to J/\psi X$ **decays** This category includes $B^0 \to J/\psi X$ decays that are not *CP* eigenstates and $B^{\pm} \to J/\psi X$. We use exponential functions as $\mathcal{P}_{\rm bkg}$ for these background classes. As we mentioned in Section 4.3.3, we have found by an MC study that the effective lifetime for background from $B^{\pm}$, $\tau_{B^{\pm}}^{\rm bkg}$, is shorter than the lifetime of $B^{\pm}$ due to a contamination from charged tracks from B_{CP} into the tagging side vertex. $\tau_{B^{\pm}}^{\rm bkg}$ is determined from the MC events to be 1.49 ± 0.04 ps. Figure 6.3 shows the Δt distribution of $B^{\pm}$ background events in an MC sample with the fit result. For background from B^0, we have found, also from an MC study, that the effective lifetime is consistent with B^0 lifetime and we thus use B^0 lifetime in the fit. We use the signal resolution function $R_{\rm sig}$ as $R_{\rm bkg}$ for this background category.

3. combinatorial background The Δt distribution of the combinatorial background is expressed as a sum of the finite lifetime component with a fraction $f_{\tau{\rm Cmb}}$ and the effective lifetime $\tau_{\rm Cmb}^{\rm bkg}$, and the prompt component with the fraction of $(1 - f_{\tau{\rm Cmb}})$.

For combinatorial background, we also use a sum of two Gaussian functions as $R_{\rm bkg}$, but with parameters different from $R_{\rm sig}$:

$$\begin{aligned} R_{\rm bkg}(\Delta t' - \Delta t_i) &= (1 - f_{\rm tail}^{\rm bkg}) \cdot G(\Delta t' - \Delta t_i; \mu_{\Delta t}^{\rm bkg}, \sigma_{\Delta t}^{\rm bkg}) \\ &+ f_{\rm tail}^{\rm bkg} \cdot G(\Delta t' - \Delta t_i; \mu_{\Delta t}^{\rm tail,bkg}, \sigma_{\Delta t}^{\rm tail,bkg}), \end{aligned} \tag{6.17}$$

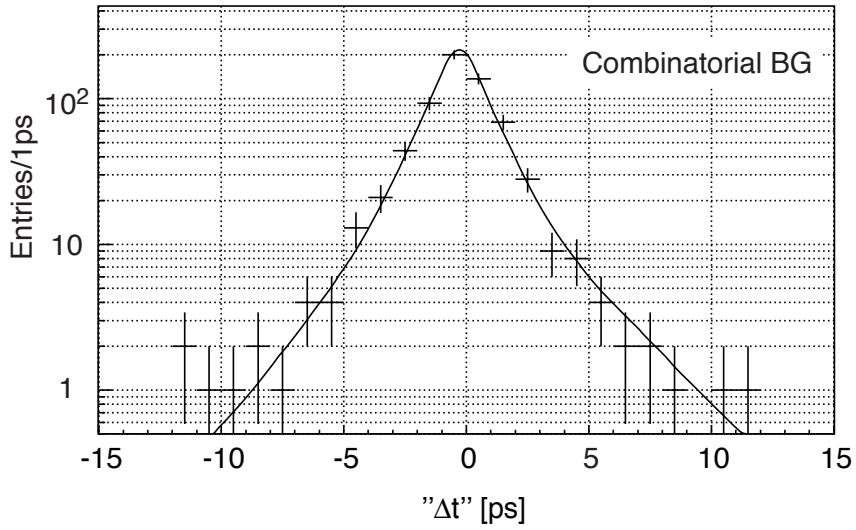

Figure 6.4: Δt distributions of $e\mu$ events used to estimate the parameters for the combinatorial background. The result of the fit is superimposed.

where $f_{\rm tail}^{\rm bkg}$, $\mu_{\Delta t}^{\rm bkg}$, and $\mu_{\Delta t}^{\rm tail,bkg}$ are constants determined from data. We use different $\mu_{\Delta t}^{\rm bkg}$ and $\mu_{\Delta t}^{\rm tail,bkg}$ for the finite lifetime component and for the prompt component, since they are considered to originate from different type of events.

The parameters $\sigma_{\Delta t}^{\rm bkg}$ and $\sigma_{\Delta t}^{\rm tail,bkg}$ are calculated event by event as

$$\begin{aligned} \sigma_{\Delta t}^{\rm bkg} &= \frac{\sigma_{\Delta z}^{\rm bkg}}{c(\beta\gamma)_\Upsilon}, \\ \sigma_{\Delta t}^{\rm tail,bkg} &= \frac{\sigma_{\Delta z}^{\rm tail,bkg}}{c(\beta\gamma)_\Upsilon}, \end{aligned} \tag{6.18}$$

$$\begin{aligned} (\sigma_{\Delta z}^{\rm bkg})^2 &= (S_{\rm det}^{\rm bkg})^2[(\tilde{\sigma}_z^{CP})^2 + (\tilde{\sigma}_z^{\rm tag})^2], \\ (\sigma_{\Delta z}^{\rm tail,bkg})^2 &= (S_{\rm tail}^{\rm bkg})^2[(\tilde{\sigma}_z^{CP})^2 + (\tilde{\sigma}_z^{\rm tag})^2], \end{aligned} \tag{6.19}$$

where $\tilde{\sigma}_z^{CP}$ and $\tilde{\sigma}_z^{\rm tag}$ are calculated as in Equation (6.13).

The parameters $f_{\tau{\rm Cmb}}$, $\tau_{\rm Cmb}^{\rm bkg}$, and the background shape parameters are obtained from a fit to the data using events reconstructed from $e^\pm\mu^\mp$ combination satisfying the selection criteria instead of correct e^+e^- or $\mu^+\mu^-$ combination. Figure 6.4 shows the "Δt" distribution of events used to estimate the background parameters of combinatorial background. The result of the fit is superimposed. The parameters used in the fit are summarized in Table 6.2.

Table 6.2: Background shape parameters for the combinatorial background.

Fit parameters		Fit values
$f_{\tau\mathrm{Cmb}}$		0.26 ± 0.08
$\tau^{\mathrm{bkg}}_{\mathrm{Cmb}}$ (ps)		1.03 ± 0.12
$S^{\mathrm{bkg}}_{\mathrm{det}}$		0.88 ± 0.07
$S^{\mathrm{bkg}}_{\mathrm{tail}}$		3.94 ± 1.14
$f^{\mathrm{bkg}}_{\mathrm{tail}}$		0.05 ± 0.02
lifetime	$\mu^{\mathrm{bkg}}_{\Delta t}$(ps)	-0.33 ± 0.10
component	$\mu^{\mathrm{tail,bkg}}_{\Delta t}$(ps)	1.86 ± 1.28
prompt	$\mu^{\mathrm{bkg}}_{\Delta t}$(ps)	-0.22 ± 0.15
component	$\mu^{\mathrm{tail,bkg}}_{\Delta t}$(ps)	-5.00 ± 1.10

In summary, we obtain the background PDF:

$$
\begin{aligned}
P_{\mathrm{bkg}}(\Delta t) \equiv\ & f^{\mathrm{bkg}}_{\psi K^{*0}} \cdot \left[\frac{e^{-|\Delta t'|/\tau_{B^0}}}{2\tau_{B^0}} A_{cp}(\Delta t', q, w_l, \xi_{\psi K^{*0}}; \sin 2\phi_1) \otimes R_{\mathrm{sig}}(\Delta t - \Delta t') \right] \\
+\ & f^{\mathrm{bkg}}_{CP_{\mathrm{even}}} \cdot \left[\frac{e^{-|\Delta t'|/\tau_{B^0}}}{2\tau_{B^0}} A_{cp}(\Delta t', q, w_l, +1; \sin 2\phi_1) \otimes R_{\mathrm{sig}}(\Delta t - \Delta t') \right] \\
+\ & f^{\mathrm{bkg}}_{CP_{\mathrm{odd}}} \cdot \left[\frac{e^{-|\Delta t'|/\tau_{B^0}}}{2\tau_{B^0}} A_{cp}(\Delta t', q, w_l, -1; \sin 2\phi_1)] \otimes R_{\mathrm{sig}}(\Delta t - \Delta t') \right] \\
+\ & f^{\mathrm{bkg}}_{B^0} \cdot \left[\frac{e^{-|\Delta t'|/\tau_{B^0}}}{2\tau_{B^0}} \otimes R_{\mathrm{sig}}(\Delta t - \Delta t') \right] \\
+\ & f^{\mathrm{bkg}}_{B^\pm} \cdot \left[\frac{e^{-|\Delta t'|/\tau^{\mathrm{bkg}}_{B^\pm}}}{2\tau^{\mathrm{bkg}}_{B^\pm}} \otimes R_{\mathrm{sig}}(\Delta t - \Delta t') \right] \\
+\ & f^{\mathrm{bkg}}_{\mathrm{Cmb}} \left[f_{\tau\mathrm{Cmb}} \cdot \frac{e^{-|\Delta t'|/\tau^{\mathrm{bkg}}_{\mathrm{Cmb}}}}{2\tau^{\mathrm{bkg}}_{\mathrm{Cmb}}} \otimes R^{\tau}_{\mathrm{bkg}}(\Delta t - \Delta t') \right. \\
& \left. +(1 - f_{\tau\mathrm{Cmb}}) \cdot \delta(\Delta t') \otimes R^{\delta}_{\mathrm{bkg}}(\Delta t - \Delta t') \right]
\end{aligned}
\tag{6.20}
$$

where the symbol $\otimes$ stands for a convolution of functions,

$$
f(\Delta t') \otimes g(\Delta t - \Delta t') \equiv \int_{-\infty}^{+\infty} f(\Delta t') \cdot g(\Delta t - \Delta t') d(\Delta t'). \tag{6.21}
$$

$f^{\mathrm{bkg}}_{\psi K^{*0}}$, $f^{\mathrm{bkg}}_{CP_{\mathrm{even}}}$, $f^{\mathrm{bkg}}_{CP_{\mathrm{odd}}}$, $f^{\mathrm{bkg}}_{B^0}$, $f^{\mathrm{bkg}}_{B^\pm}$, and $f^{\mathrm{bkg}}_{\mathrm{Cmb}}$ are the probabilities of background components from $J/\psi K^{*0}(K_L\pi^0)$, $\xi_f = -1$ CP-modes, $\xi_f = +1$ CP-modes, the rest

of $B^0 \to J/\psi X$, $B^\pm \to J/\psi X$, and combinatorial J/ψ background among all the background, respectively ($f^{\rm bkg}_{\psi K^{*0}} + f^{\rm bkg}_{CP_{\rm even}} + f^{\rm bkg}_{CP_{\rm odd}} + f^{\rm bkg}_{B^0} + f^{\rm bkg}_{B^\pm} + f^{\rm bkg}_{\rm Cmb} = 1$). The probabilities of each background component are functions of p^*_B and determined from real data as described later.

6.4 Determination of Signal and Background Probabilities

The understanding of background component is very important in this analysis. We need to estimate the probability that an event is the signal or one of the background categories we have defined in the previous section. We first determine the signal probability by fitting the p^*_B (the magnitude of the momentum of B^0 in the CMS) distribution of the data. Then, the probability of each background category is determined based on the Monte Calro simulation. We treat the KLM candidates and ECL candidates separately, because they are considered to have different background composition.

6.4.1 Estimation of signal probability

We fit the p^*_B distribution with a sum of four components to obtain the signal probability, $f_{\rm sig}$:

1) signal;

2) $B \to J/\psi X$ background which contains true K_L;

3) $B \to J/\psi X$ background which does not contain true K_L;

4) J/ψ combinatorial background.

The classification of category 2) and 3) is to take into account known difference of K_L detection efficiency between MC and data. The shapes of the first three components are determined from 1 million J/ψ inclusive MC events and represented by histograms. The J/ψ combinatorial background is evaluated using events with $e^\pm\mu^\mp$ pair which pass the event selection criteria. The p^*_B distribution of this component is modeled with a second order polynomial function. Figures 6.5 and 6.6 show the p^*_B distributions for each component. In the fit, the offset in p^*_B for the signal is allowed to float taking into account a possible discrepancy of the signal shape.

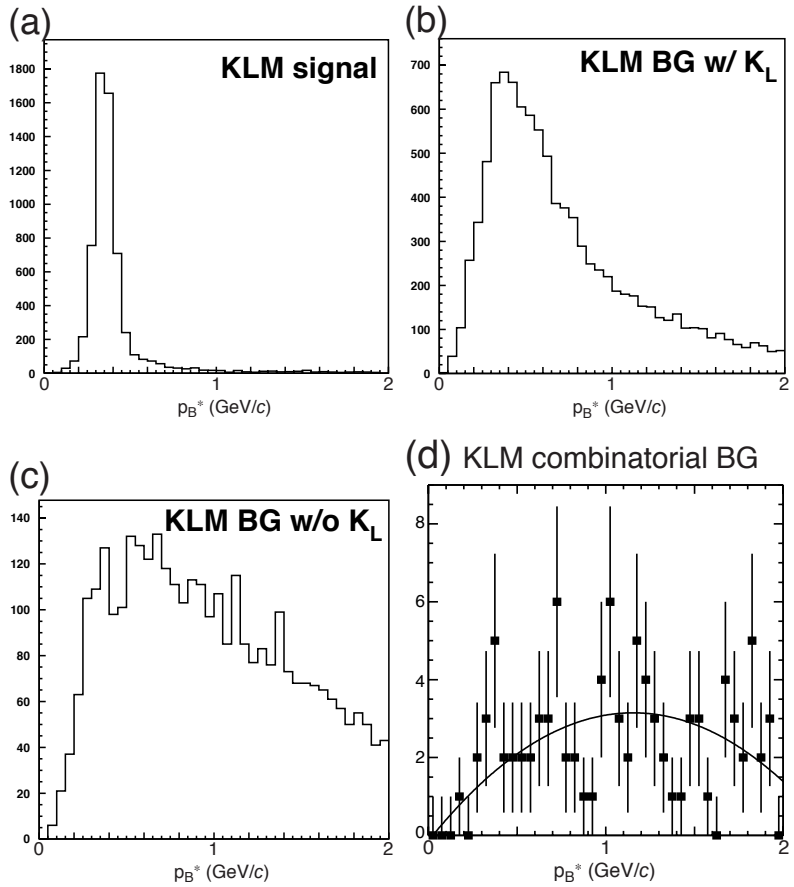

Figure 6.5: p_B^* distributions for KLM candidates: a) signal, b) $B \to J/\psi X$ background with K_L, c) $B \to J/\psi X$ background without K_L, and d) J/ψ combinatorial background. a)–c) are from MC and d) is from data with $e\mu$ pair. Fit with second-order polynomial is superimposed in d).

Figure 6.7 shows the p_B^* distribution with the fit result. The distributions for KLM and ECL candidates are shown separately. The points represent data, the solid line is a sum of the signal and background, and the hatched histograms show the background components. By integrating each component obtained by the fit in the signal region $0.2 \leq p_B^* \leq 0.45$ GeV/c ($0.2 \leq p_B^* \leq 0.40$ GeV/c in the case of ECL candidates), we estimate the signal yield as 346.3±28.8 events.

There is rather large discrepancy in the K_L detection efficiency between MC and real data. From the above fit, the efficiency in the real data is estimated to be about 55% of that in MC for both signal and background (both with and without K_L). The reason is considered to be incorrect modeling of hadronic interaction in the

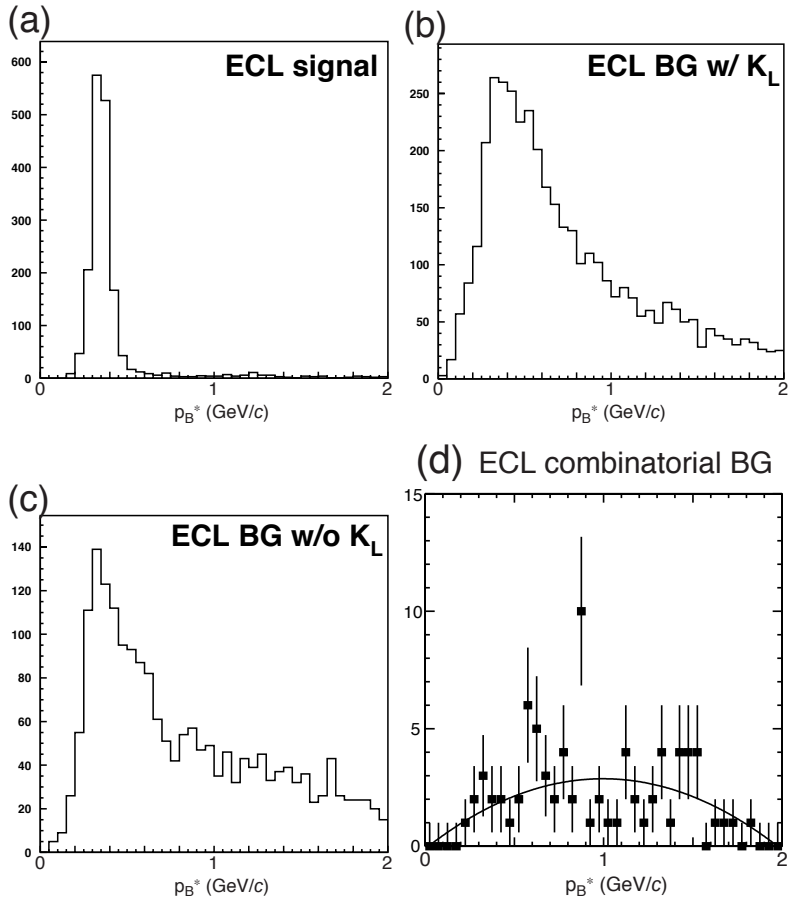

Figure 6.6: p_B^* distributions for ECL candidates: a) signal, b) $B \to J/\psi X$ background with K_L, c) $B \to J/\psi X$ background without K_L, and d) J/ψ combinatorial background. a)–c) are from MC and d) is from data with $e\mu$ pair. Fit with second-order polynomial is superimposed in d).

MC. The study to solve this discrepancy is now under way, but the difference in the detection efficiency does not lead to any bias in the measurement of *CP* asymmetry since flavor tagging and proper-time difference measurement do not correlate with the efficiency. The possible difference in the background component is accounted for by the estimation procedure described here and in the next section.

6.4.2 Estimation of background component probabilities

Now we need to estimate the probabilities of background components which are categorized by their Δt distributions as:

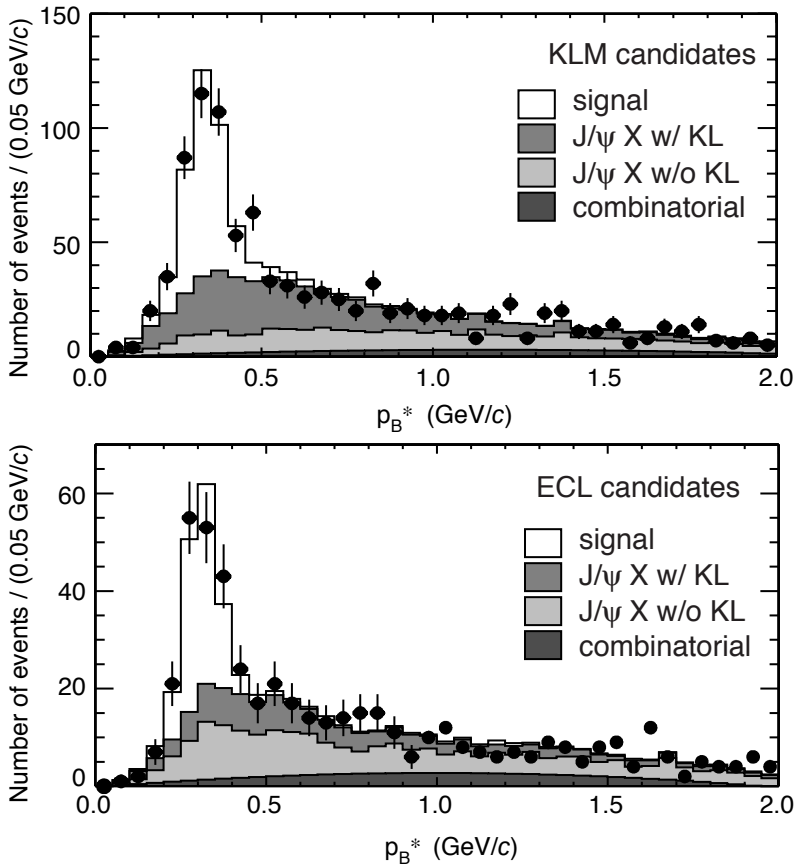

Figure 6.7: The p_B^* distribution with the fit result for KLM candidates (top) and ECL candidates (bottom). The points represent data. The solid line is a sum of the signal and background. The hatched histograms show the distributions of the background components.

1. $J/\psi K^{*0}$,
2. $CP = +1$ modes ($B^0 \to \psi(2S)K_L$, $\chi_{c1}K_L$, $J/\psi\pi^0$),
3. $CP = -1$ modes ($B^0 \to J/\psi K_S$),
4. other $B^0 \to J/\psi X$,
5. $B^+ \to J/\psi X$, and
6. combinatorial J/ψ background.

The probability of the combinatorial J/ψ background, $f_{\rm Cmb}^{\rm bkg}$ is already determined by the fit described in the last section. The probabilities of other components within

total background, $f^{\rm bkg}_{\psi K^{*0}}$, $f^{\rm bkg}_{CP_{\rm even}}$, $f^{\rm bkg}_{CP_{\rm odd}}$, $f^{\rm bkg}_{B^0}$, and $f^{\rm bkg}_{B^\pm}$ are estimated from MC simulation. The p^*_B distribution of each component is estimated with J/ψ inclusive MC and modeled by functions of p^*_B. Then, these functions are used to calculate the probabilities of background components within the total background.

The "$B \to J/\psi X$ background with K_L" category includes contributions from $J/\psi K^{*0}$ ($K^{*0} \to K_L \pi^0$), $\psi(2S)/\chi_{c1} + K_L$, other $B^0 \to J/\psi X$, and $B^\pm \to J/\psi X$. Its fraction in the total background, f_{K_L}, is a function of p^*_B and determined from the fit to the p^*_B distribution of the data, as described in the previous section. The p^*_B distribution of each component in MC, $f_{J/\psi K^{*0}}$, f_{ccK_L}, $f^{K_L}_{B^0}$, $f^{K_L}_{B^\pm}$ is modeled with a function of p^*_B. Since we find f_{ccK_L} is negligibly small in ECL candidates compared to other components, we use this category for only KLM candidates.

Following three functions are used for the modeling of background distributions:

- Reversed Crystal-Ball function,

$$F_{\rm RCB}(\Delta t) = \begin{cases} A\exp(-\frac{(\Delta t-\mu)^2}{2\sigma^2}) & \text{for } \Delta t < \mu + \alpha\sigma \\ A\frac{\exp(-\alpha^2/2)}{(1+\frac{(\Delta t-\mu)\alpha}{n\sigma}-\frac{\alpha^2}{n})^n} & \text{for } \Delta t > \mu + \alpha\sigma \end{cases} \tag{6.22}$$

- Threshold function,

$$F_{\rm Th}(\Delta t) = A(\Delta t - \mu)^n e^{(c_1(\Delta t-\mu)+c_2(\Delta t-\mu)^2)} \tag{6.23}$$

- Gaussian function

$$G(\Delta t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\Delta t-\mu)^2}{2\sigma^2}}. \tag{6.24}$$

We select the function which gives maximum confidence level with a binned maximum likelihood fit to the p^*_B distribution. Figures 6.8 and 6.9 show the p^*_B distributions of background components for "$B \to J/\psi X$ background with K_L" category in MC. The results of the fit are superimposed.

Then, we calculate the probability of each component from these functions. For example, the probability of $J/\psi K^{*0}$ background within the total background, $f^{\rm bkg}_{\psi K^{*0}}$ is calculated as (probability of "BG with K_L" in total BG) $\times$ (probability of $J/\psi K^{*0}$ in "BG with K_L"):

$$f^{\rm bkg}_{\psi K^{*0}} = f_{K_L} \times \frac{f_{J/\psi K^{*0}}}{f_{J/\psi K^{*0}} + f_{ccK_L} + f^{K_L}_{B^0} + f^{K_L}_{B^\pm}}. \tag{6.25}$$

Note that all the probabilities are functions of p^*_B. Probabilities of other components are calculated in the same way.

Similarly, the "background without K_L" contains $J/\psi\,\pi^0$, $J/\psi\,K_S$, other $B^0 \to J/\psi\,X$, and $B^\pm \to J/\psi\,X$. We fit the p_B^* distributions of each component in MC with smooth functions and calculate the probabilities in the same way as "background with K_L". Since we find contribution from $B^0 \to J/\psi\,\pi^0$ is negligibly small in KLM candidates, we use this category for only ECL candidates. Figures 6.10 and 6.11 show the p_B^* distributions of background components for "$B \to J/\psi\,X$ background without true K_L" category in MC. The fit results are superimposed.

Figure 6.12 shows the p_B^* distribution for the sum of KLM and ECL candidates with the breakdown of background component. Table 6.3 shows the fraction of each component integrated over the signal p_B^* region. In the fit to extract $\sin 2\phi_1$, we use probabilities calculated event by event depending of the value of p_B^*.

Table 6.3: Yield of signal and each background component in $J/\psi\,K_L$ mode. The numbers shown are estimated yield in the signal p_B^* region.

	Component	KLM	ECL	sum
Signal	$J/\psi\,K_L$	244.8	101.5	346.3
Background	$J/\psi\,K^{*0}(K^{*0} \to K_L\,\pi^0)$	24.2	5.5	29.7
	$CP = +1$	14.4	18.2	32.6
	$CP = -1$	8.9	13.7	22.6
	others from $B^0 \to J/\psi\,X$	18.6	20.4	39.0
	others from $B^\pm \to J/\psi\,X$	88.4	21.0	109.4
	Combinatorial J/ψ	6.6	5.4	12.0
Total yield		397	172	569

6.5 Result of Fit

Figure 6.13 shows Δt distribution for the events in which the flavor of $B_{\rm tag}$ being tagged as B^0 ($q = +1$, solid points) and $\overline{B}^0$ ($q = -1$, open points). Clear difference between two distributions is seen, indicating a large CP violation in the neutral B meson system. In order to determine the standard model CP violating parameter $\sin 2\phi_1$, we use an unbinned maximum likelihood method.

We have defined the PDF

$$P(\Delta t; \sin 2\phi_1) = f_{\rm sig} P_{\rm sig}(\Delta t) + (1 - f_{\rm sig}) P_{\rm bkg}(\Delta t) \tag{6.26}$$

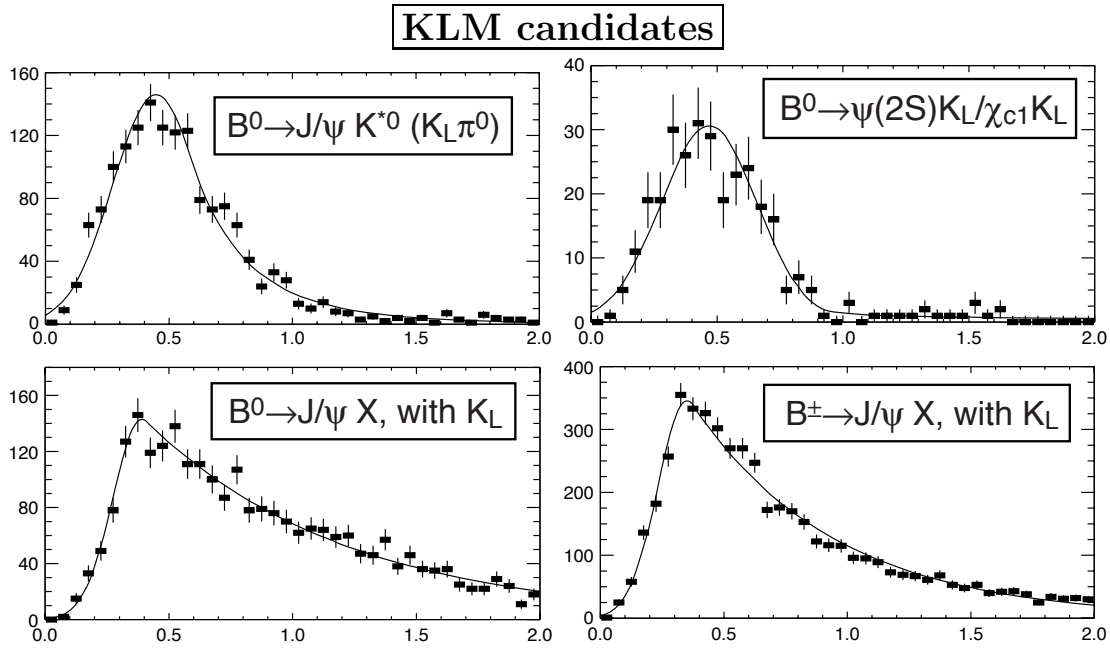

Figure 6.8: p_B^* distributions of background with true K_L in KLM candidates. The results of fit are superimposed.

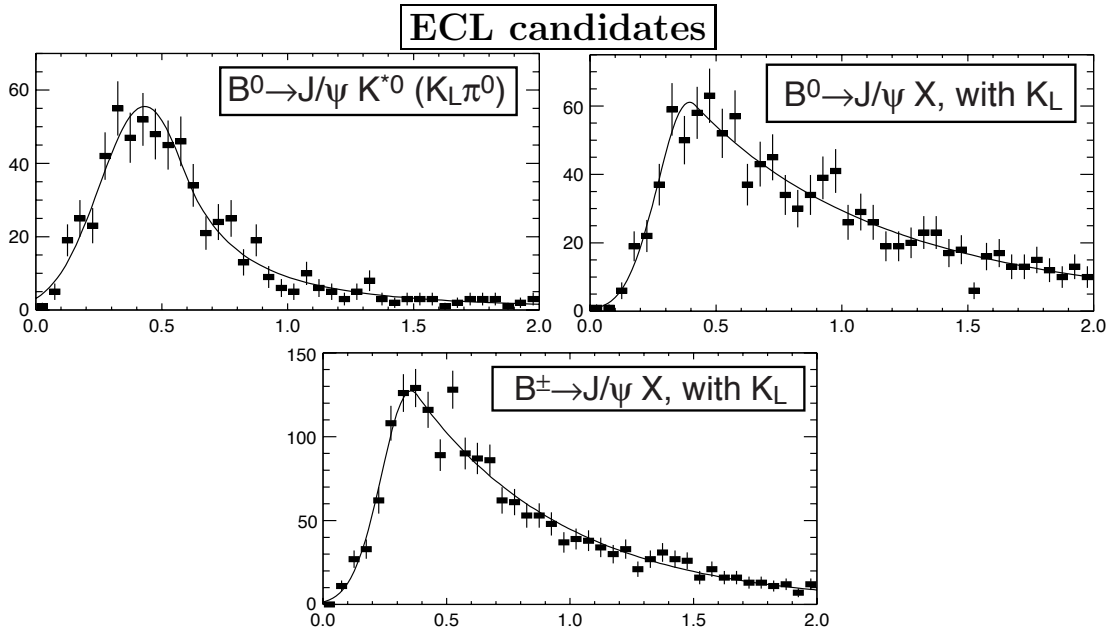

Figure 6.9: p_B^* distributions of background with true K_L in ECL candidates. The results of fit are superimposed.

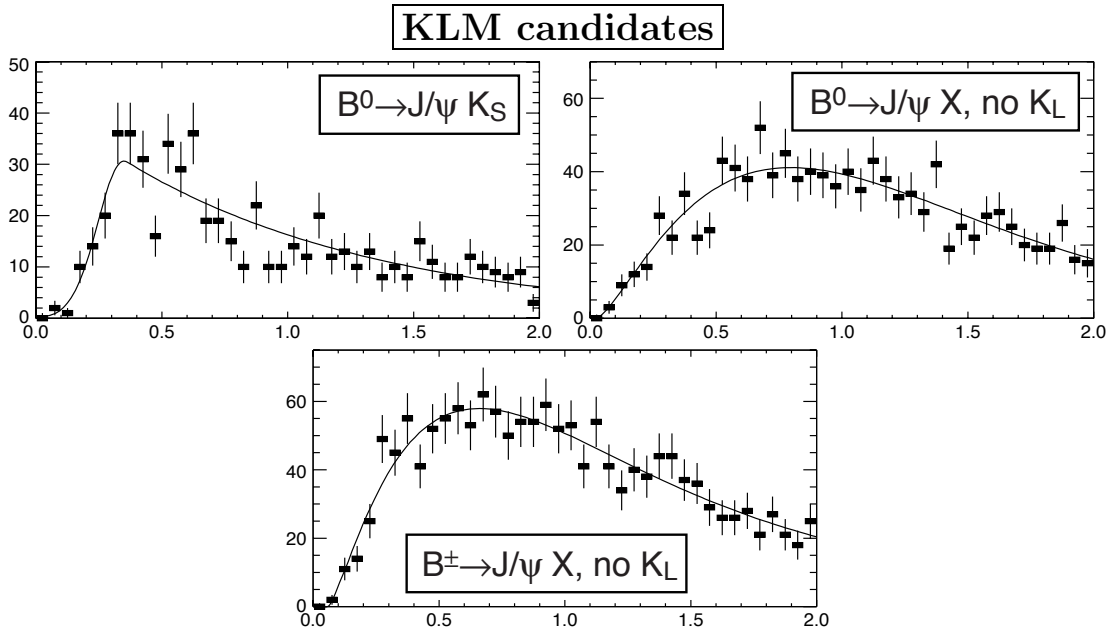

Figure 6.10: p_B^* distributions of background without true K_L in KLM candidates. The results of fit are superimposed.

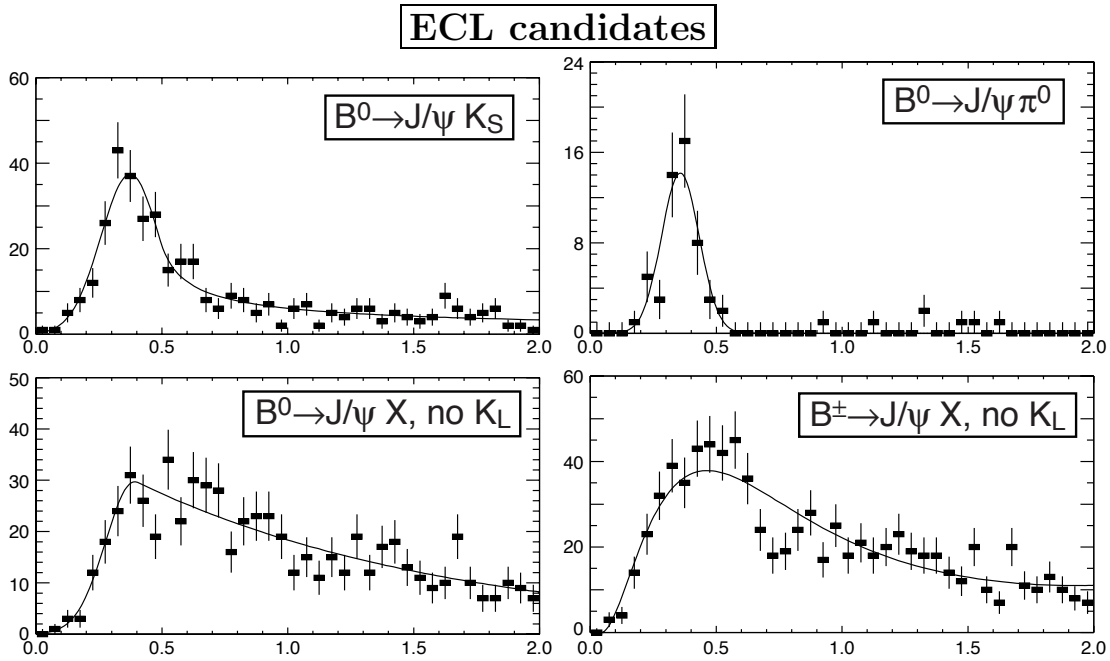

Figure 6.11: p_B^* distributions of background without true K_L in ECL candidates. The results of fit are superimposed.

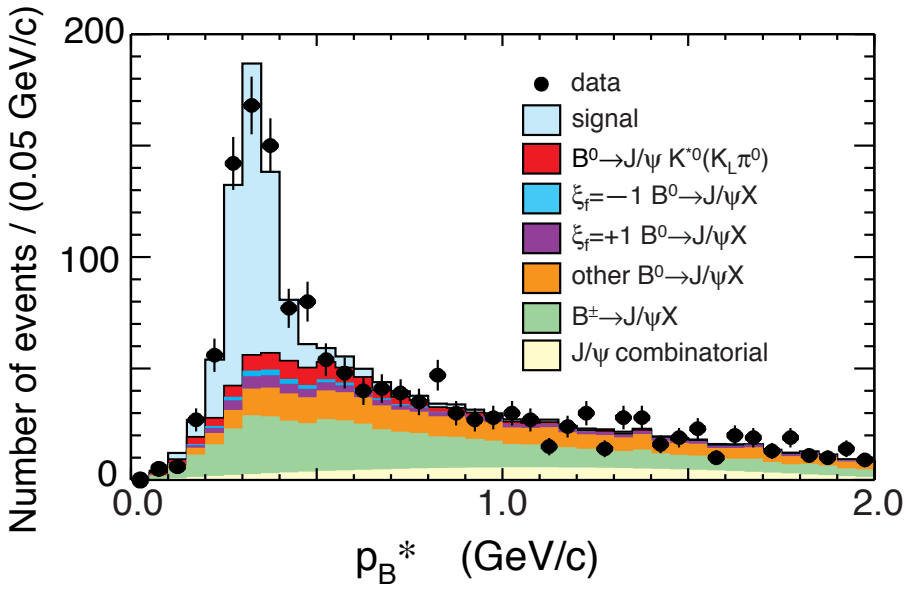

Figure 6.12: The p_B^* distribution for $B^0 \to J/\psi\, K_L$ with the breakdown of background components. The points represent data. The KLM and ECL candidates are combined in this plot.

where $P^i_{\rm sig}$ and $P^i_{\rm bkg}$ are given in Equations (6.8) and (6.20) and $f_{\rm sig}$ is the signal probability calculated event-by-event using p^*_B. We determine the most probable value for $\sin 2\phi_1$ by maximizing the likelihood function

$$\mathcal{L} = \prod_i P(\Delta t_i; \sin 2\phi_1), \tag{6.27}$$

where the product is over all candidate events.

After flavor tagging and vertex reconstruction, 523 candidate events remain for $B^0 \to J/\psi\, K_L$ decays. Using these events, we find the most probable value for $\sin 2\phi_1$ by scanning over $\sin 2\phi_1$ to minimize the $-2\ln(\mathcal{L}) = -2\sum[\ln(\mathcal{L}_i)]$ using MINUIT package [56]. We fix the lifetime and mixing parameter of B^0 meson to the world average values [18], $\tau_{B^0} = 1.548$ ps and $\Delta m_d = 0.472 \times 10^{12}\ \hbar{\rm s}^{-1}$. The $\sin 2\phi_1$ is the only free parameter in the fit.

We obtain

$$\sin 2\phi_1 = 1.31^{+0.19}_{-0.23}(stat.) \tag{6.28}$$

where the error shown is statistical. The obtained PDF distributions are also shown in Fig. 6.13. The distributions are normalized so that the area becomes one for each of $q = \pm 1$ samples.

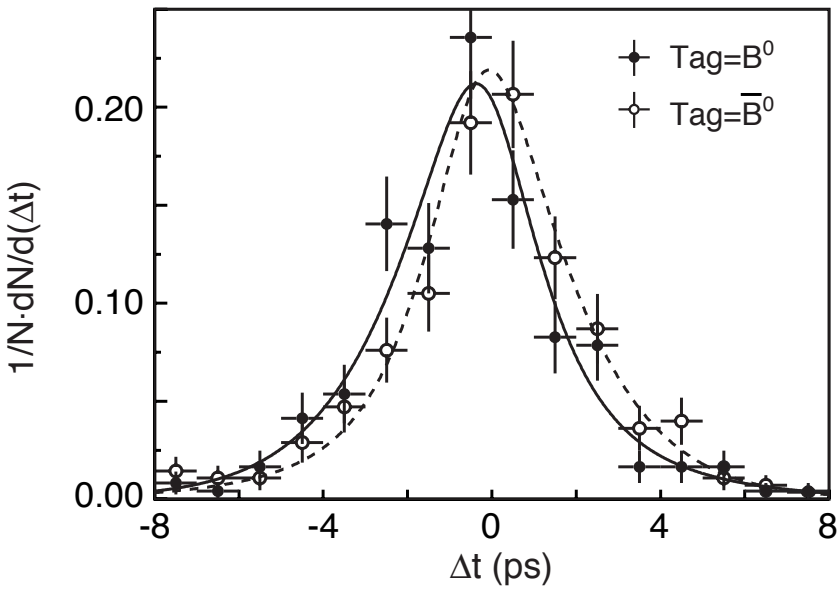

Figure 6.13: Δt distributions for the events with $B_{\rm tag}$ is tagged as B^0 (solid points) and $\overline{B}^0$ (open points). The results of the global fit (with $\sin 2\phi_1 = 1.31$) are shown as solid and dashed curves, respectively.

Figure 6.14 shows the $-2\ln(\mathcal{L}/\mathcal{L}_{\rm max})$ as a function of $\sin 2\phi_1$. Figure 6.15 shows the asymmetry $\sin 2\phi_1 \sin(\Delta m_d \Delta t)$ obtained in each Δt bins for $B^0 \to J/\psi\, K_L$ mode. The unbinned maximum likelihood fit is performed in the same way for each Δt bin separately. The obtained $\sin 2\phi_1$ and its error are multiplied by the $\sin(\Delta m_d \Delta t)$ value for the average Δt in each bin. Points are plotted at the average Δt in each bin. A curve of $\sin 2\phi_1 \sin(\Delta m_d \Delta t)$ for $\sin 2\phi_1 = 1.31$ is also shown in the plot. The agreement between data and the curve in all Δt bins is seen.

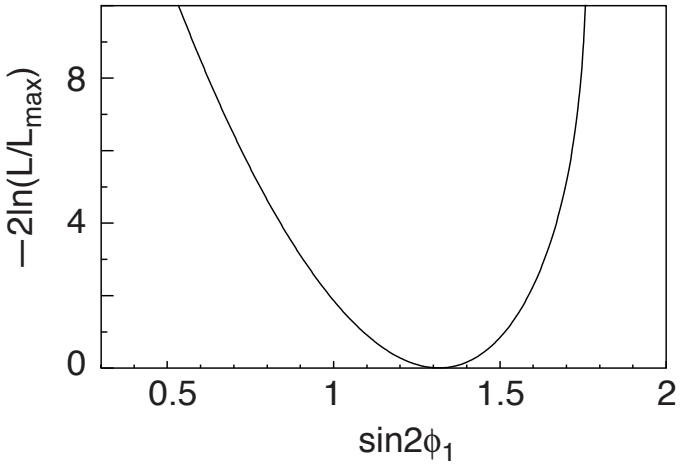

Figure 6.14: Values of $-2\ln(\mathcal{L}/\mathcal{L}_{max})$ *vs.* $\sin 2\phi_1$.

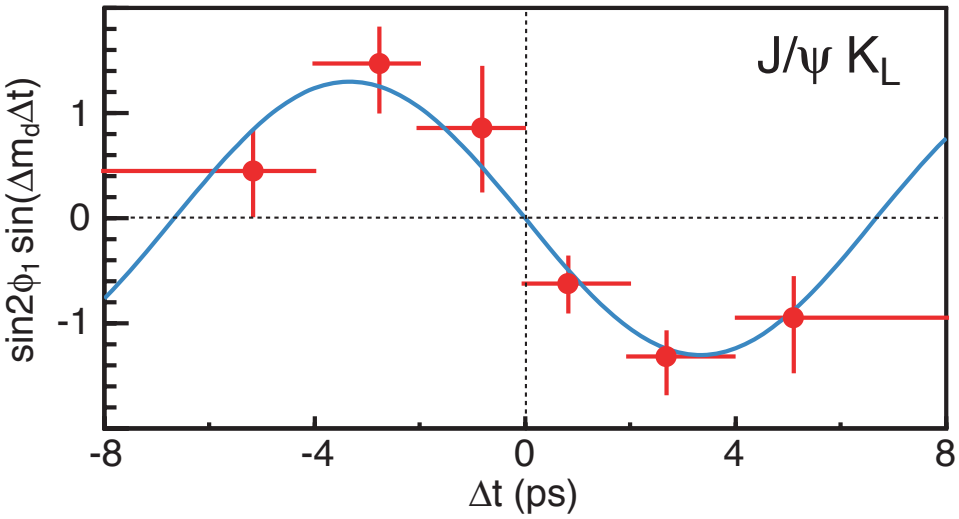

Figure 6.15: The asymmetry obtained from separate fit to each Δt bin. The curve is the result of the global fit, $\sin 2\phi_1 = 1.31$.

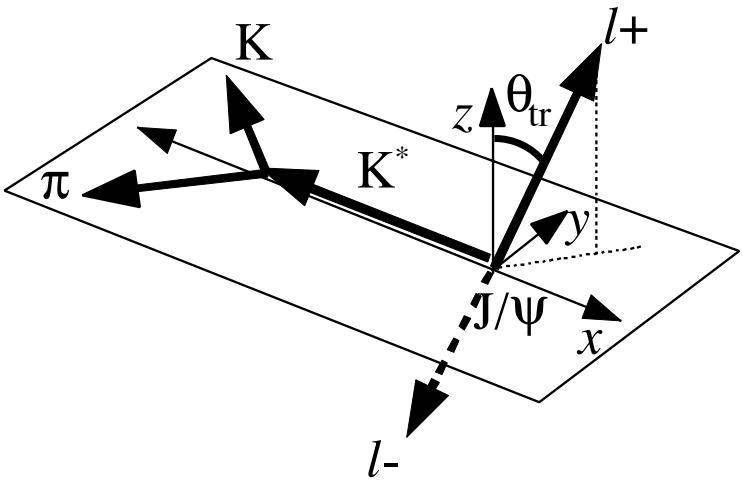

Figure 6.16: Definition of the transversity angles.

6.6 Measurement of $\sin 2\phi_1$ with $B^0 \to J/\psi K^{*0}$ Decays

Since the $J/\psi K^{*0}$ state is a mixture of *CP* even and odd states, we need to use the decay angle information of the decay daughters to separate the *CP* even and odd components. Although we can use the net *CP* eigenvalue calculated from the fraction of *CP* odd component measured by the analysis of angular distribution of decay products, it is statistically beneficial to use the event-by-event angular information in the likelihood fit. In this section, we present a measurement of $\sin 2\phi_1$ using $B^0 \to J/\psi K^{*0}(K^{*0} \to K_S \pi^0)$ mode.

6.6.1 Signal PDF

Including the angular distribution of the decay products to separate different *CP* state, the theoretical distribution for $J/\psi K^{*0}(K^{*0} \to K_S \pi^0)$ mode is defined as [20]

$$\begin{aligned}
\mathcal{P}_{\rm sig}(\Delta t, \theta_{\rm tr}, q, w_l, \xi_f; \sin 2\phi_1) &= \\
(1 - f_{\rm odd})\frac{3}{8}(1 + \cos^2\theta_{\rm tr}) &\times \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}} A_{cp}(\Delta t, q, w_l, +1; \sin 2\phi_1) \\
+ \quad f_{\rm odd}\frac{3}{4}(1 - \cos^2\theta_{\rm tr}) &\times \frac{e^{-|\Delta t|/\tau_{B^0}}}{A}_{cp}(\Delta t, q, w_l, -1; \sin 2\phi_1), \qquad (6.29)
\end{aligned}$$

$$A_{cp}(\Delta t, q, w_l, \xi_f; \sin 2\phi_1) = 1 - \xi_f q(1 - 2w_l)\sin 2\phi_1 \sin(\Delta m_d \Delta t) \qquad (6.30)$$

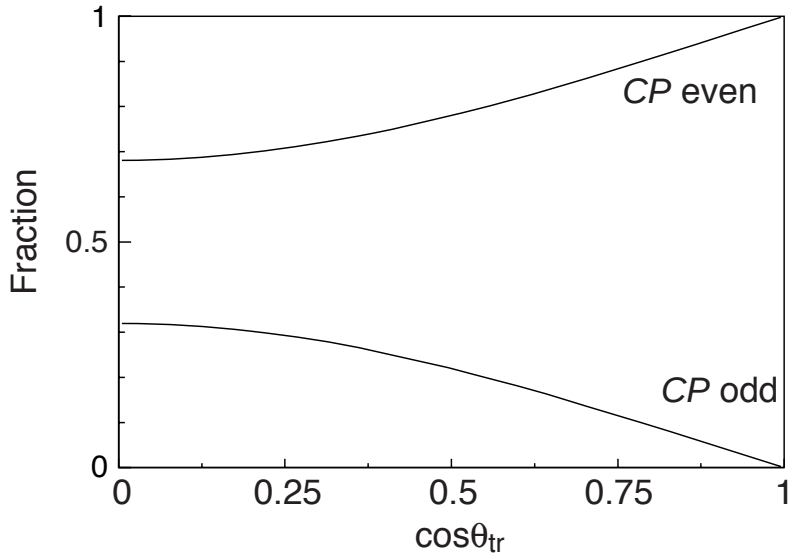

Figure 6.17: Fraction of the *CP* even and odd components for the case $f_{\rm odd}$=0.19, as a function of $\theta_{\rm tr}$.

where $f_{\rm odd}$ is the fraction of $\xi_f = -1$ component in $B^0 \to J/\psi K^{*0}(K^{*0} \to K_S\pi^0)$ mode and is measured to be 0.19 ± 0.06 from $B^0 \to J/\psi K^{*0}(K^{*0} \to K^+\pi^-)$ and $B^+ \to J/\psi K^{*+}(K^{*+} \to K_S\pi^+, K^+\pi^0)$ decays [21]. $\theta_{\rm tr}$ is the transversity angle, defined as the angle between the ℓ^+ direction in the J/ψ rest frame and the z axis, where the x-y plane is defined by the K^* decay products in the J/ψ rest frame and the x axis is defined as the direction of motion of the J/ψ in the $\Upsilon(4S)$ rest frame. Figure 6.16 shows the definition of the transversity angle. Figure 6.17 shows the fraction of the *CP* even and odd components for the case $f_{\rm odd}$=0.19, as a function of $\cos\theta_{\rm tr}$.

The signal PDF, $P_{\rm sig}$ is defined as

$$P_{\rm sig}(\Delta t_i; \sin 2\phi_1) = \int_{-\infty}^{+\infty} \mathcal{P}_{\rm sig}(\Delta t', \theta_{\rm tr}, q, w_l, \xi_f; \sin 2\phi_1) \cdot R_{\rm sig}(\Delta t_i - \Delta t') d(\Delta t') \quad (6.31)$$

using the above $\mathcal{P}_{\rm sig}$ and the same resolution function $R_{\rm sig}$ as $J/\psi K_L$ mode.

6.6.2 Background PDF

The background in $B \to J/\psi K^*$ mode is studied using MC simulation. Major sources of the background are found to be contamination from other $B \to J/\psi K^*$ modes (we call this type of background "feed-across" hereafter), non-resonant $B \to$

$J/\psi K\pi$ decay, and background from random combinations of particles (combinatorial background).

The background likelihood is defined as

$$P_{\rm bkg}(\Delta t_i) = \int_{-\infty}^{+\infty} \sum_j \left\{ \mathcal{P}^j_{\rm bkg}(\Delta t') \cdot R^j_{\rm bkg}(\Delta t_i - \Delta t') \right\} d(\Delta t'). \tag{6.32}$$

where j denotes each background component.

6.6.2.1 Construction of background PDF

For the feed-across and the non-resonant background, we use an exponential function as $\mathcal{P}_{\rm bkg}$. From an MC study, the effective lifetimes for these components are found to be consistent with the lifetime of B^0. The possible difference of the effective lifetime and B^0 lifetime is treated as the source of systematic error, as described later. For these backgrounds, we use the signal resolution function $R_{\rm sig}$ as $R_{\rm bkg}$.

The distribution of combinatorial background is parametrized with the same form as $B^0 \to J/\psi K_L$,

$$\begin{aligned} P_{\rm cmb} &= f^{\rm bkg}_{\rm Cmb} \Bigg[f_{\tau Cmb} \cdot \frac{e^{-|\Delta t'|/\tau^{\rm bkg}_{\rm Cmb}}}{2\tau^{\rm bkg}_{\rm Cmb}} \otimes R^{\tau}_{\rm bkg}(\Delta t - \Delta t') \\ &+ (1 - f_{\tau Cmb}) \cdot \delta(\Delta t') \otimes R^{\delta}_{\rm bkg}(\Delta t - \Delta t') \Bigg] \end{aligned} \tag{6.33}$$

where $R_{\rm bkg}$ is defined as Eq. (6.17). The background shape parameters for combinatorial background is determined from a fit to the $B \to J/\psi K$ ($K = K^{\pm}, K_S$) events in background-dominated regions of ΔE vs M_{bc} distribution, defined as:

$$\begin{aligned} &5.20 < M_{\rm bc} \leq 5.30 \text{ GeV}/c^2 \quad \text{if} \quad 0.12 < \Delta E \leq 0.20 \text{ GeV}, \\ &5.20 < M_{\rm bc} \leq 5.24 \text{ GeV}/c^2 \quad \text{if} \quad -0.12 < \Delta E \leq 0.12 \text{ GeV}, \\ &5.20 < M_{\rm bc} \leq 5.26 \text{ GeV}/c^2 \quad \text{if} \quad -0.20 \leq \Delta E \leq -0.12 \text{ GeV}. \end{aligned}$$

We use these modes because statistics of $B^0 \to J/\psi K^{*0}(K^{*0} \to K_S\pi^0)$ decays themselves is too poor to determine these parameters. Parameters obtained by $J/\psi K$ modes are valid for $B^0 \to J/\psi K^{*0}(K^{*0} \to K_S\pi^0)$ mode because of similar decay kinematics. From the fit we have found that f_τ is negligibly small. Thus we set $f_\tau = 0$ in the following analysis. The obtained parameters are summarized in Table 6.4.

Table 6.4: Background shape parameters for combinatorial background in $B^0 \to J/\psi K^{*0}$ mode.

Fit parameters		Fit values
$S_{\rm det}^{\rm bkg}$		1.08 ± 0.06
$S_{\rm tail}^{\rm bkg}$		3.31 ± 0.28
$f_{\rm tail}^{\rm bkg}$		0.14 ± 0.04
prompt component	$\mu_{\Delta t}^{\rm bkg}$(ps)	-0.05 ± 0.04
	$\mu_{\Delta t}^{\rm tail,bkg}$(ps)	-0.12 ± 0.26

6.6.2.2 determination of signal and background probabilities

The probability that an event is the combinatorial background, $f_{\rm cmb}$, is calculated based on ΔE and $M_{\rm bc}$ of the event, using functions determined from a two-dimensional fit to the ΔE-$M_{\rm bc}$ distribution of data. $f_{\rm cmb}$ is defined as

$$f_{\rm cmb} = \frac{F_{\rm CMB}(\Delta E, M_{\rm bc})}{F_{\rm CMB}(\Delta E, M_{\rm bc}) + F_{J/\psi K^*}(\Delta E, M_{\rm bc})}, \tag{6.34}$$

where $F_{\rm CMB}$ and $F_{J/\psi K^*}$ are functions which represent combinatorial background and $B \to J/\psi K^*$ component (which includes signal, feed-across and non-resonant background).

The distribution of combinatorial background is fit with a one-dimensional polynomial for ΔE and ARGUS function [58] for $M_{\rm bc}$ distribution, respectively.

$$F_{\rm CMB} = a \cdot (1 + c\Delta E) \cdot M_{\rm bc}\sqrt{1 - (M_{\rm bc}/E_{\rm beam})^2}\exp(N(1 - (M_{\rm bc}/E_{\rm beam})^2), \tag{6.35}$$

The $M_{\rm bc}$ distribution of $B \to J/\psi K^*$ component is fit with a Gaussian function with mean $\mu_{M_{\rm bc}}$ and sigma $\sigma_{M_{\rm bc}}$. The ΔE distribution is represented by the Crystal Ball function [59] $F_{\rm CB}$:

$$F_{CB}(\Delta E; \mu_{\Delta E}, \sigma_{\Delta E}, \alpha, n) = \begin{cases} A\exp(-\frac{(\Delta E - \mu_{\Delta E})^2}{2(\sigma_{\Delta E})^2}) & \text{for } \Delta E > \mu_{\Delta E} - \alpha\sigma_{\Delta E} \\ A\frac{\exp(-\alpha^2/2)}{(1+\frac{(\Delta E - \mu_{\Delta E})\alpha}{n\sigma_{\Delta E}} - \frac{\alpha^2}{n})^n} & \text{for } \Delta E < \mu_{\Delta E} - \alpha\sigma_{\Delta E} \end{cases} \tag{6.36}$$

which has long tail in the lower side of the peak. The values $\mu_{\Delta E}$, $\sigma_{\Delta E}$, α, n, $\mu_{M_{\rm bc}}$, and $\sigma_{M_{\rm bc}}$ are determined from MC simulation, because the statistics of real data is too small to determine these parameters. The values c and N are determined from two-dimensional maximum likelihood fit to real data. $F_{J/\psi K^*}$ is written as:

$$F_{J/\psi K^*}(\Delta E, M_{\rm bc}) = b \cdot F_{\rm CB}(\Delta E; \mu_{\Delta E}, \sigma_{\Delta E}, \alpha, n) \cdot G(M_{\rm bc}; \mu_{M_{\rm bc}}, \sigma_{M_{\rm bc}}). \tag{6.37}$$

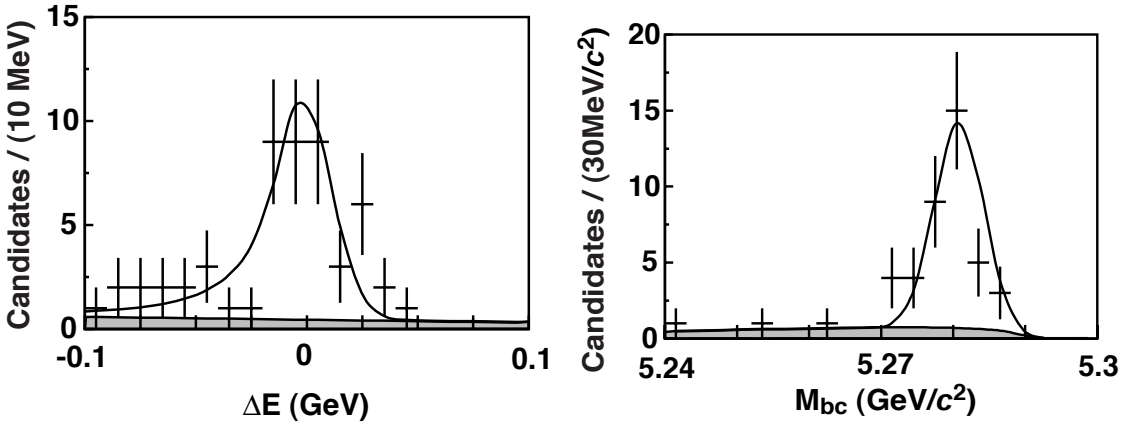

Figure 6.18: ΔE (left) and $M_{\rm bc}$ (right) distribution together with the fit result. In the ΔE ($M_{\rm bc}$) plot, $5.27 < M_{\rm bc} < 5.29$ GeV ($-0.05 < \Delta E < 0.03$ GeV/c^2) is required. The hatched areas indicate the expected distribution of combinatorial background.

The normalization factors a and b are determined such that the integration in the signal box is consistent to the overall signal-to-background ratio obtained from the fit to $M_{\rm bc}$ distribution for events with ΔE in the signal region. Figure 6.18 shows the result of the fit.

The contribution from the non-resonant $B \to J/\psi K\pi$ decay is determined from events in the K^* mass sideband in data. We use four $B \to J/\psi K^*$ modes ($B^0 \to J/\psi K^{*0}, K^{*0} \to K^+\pi^-/K_S\pi^0$ and $B^+ \to K^{*+}, K^{*+} \to K^+\pi^0/K_S\pi^+$) to estimate the non-resonant component, assuming the same fraction of non-resonant decay for these modes. Other three modes are reconstructed with similar procedure as $K^{*0} \to K_S\pi^0$ mode. Figure 6.19 shows the $K\pi$ invariant mass distribution for events selected with the criteria except for $K\pi$ invariant mass requirement. The $K\pi$ mass distribution is fit with two Breit-Wigner functions describing the $K^*(892)$ and $K_2^*(1430)$ mass peaks, and a second order polynomial. The peak positions and the widths of the Breit-Wigner functions are fixed at the world average values [18]. The probability that an event is non-resonant decay, $f_{\rm NR}$, is estimated from this distribution.

The $M_{\rm bc}$ distributions of the signal and the feed-across background are parametrized with Gaussian functions. The parameters of these function are determined with MC. Then, the probabilities of the signal and the feed-across background

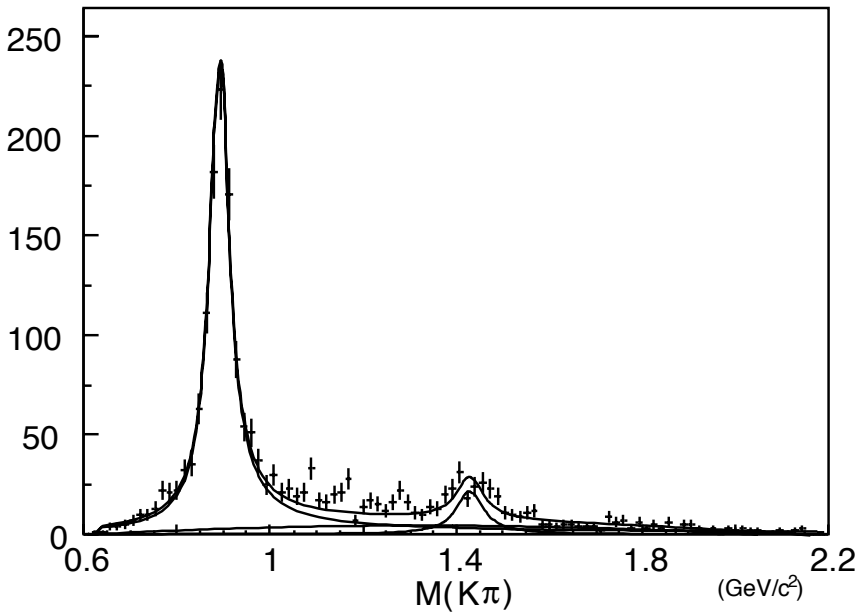

Figure 6.19: $K\pi$ mass distribution for $J/\psi K^*$ events without $K\pi$ invariant mass requirement. All modes are combined.

are calculated as a function of $M_{\rm bc}$:

$$f_{\rm sig} = (1 - f_{\rm cmb} - f_{\rm NR}) \cdot \frac{F_{\rm sig}(M_{\rm bc})}{F_{\rm sig}(M_{\rm bc}) + F_{\rm FA}(M_{\rm bc})}, \tag{6.38}$$

where $F_{\rm sig}$ and $F_{\rm FA}$ are functions which represent $M_{\rm bc}$ distributions of the signal and the feed-across background, respectively.

The signal reconstruction efficiency may depend on the transversity angle, $\theta_{\rm tr}$. This may result in the dependence of S/N ratio on $\theta_{\rm tr}$ and thus dilute the measured $\sin 2\phi_1$. We take into account this effect by introducing the efficiency function $\epsilon(\theta_{\rm tr})$, which is determined from the MC study and parametrized by a fourth order polynomial function. Figure 6.20 shows the signal efficiency as a function of $\cos\theta_{\rm tr}$ measured with MC. The efficiency function is found to be almost flat except for large $|\cos\theta_{\rm tr}|$ region. The efficiencies for the background modes are assumed to be flat based on the MC study.

We include this effect by scaling the signal probability $f'_{\rm sig}$ as

$$f'_{\rm sig} = \frac{\epsilon(\theta_{\rm tr}) f_{\rm sig}}{\epsilon(\theta_{\rm tr}) f_{\rm sig} + f_{\rm bkg}} \tag{6.39}$$

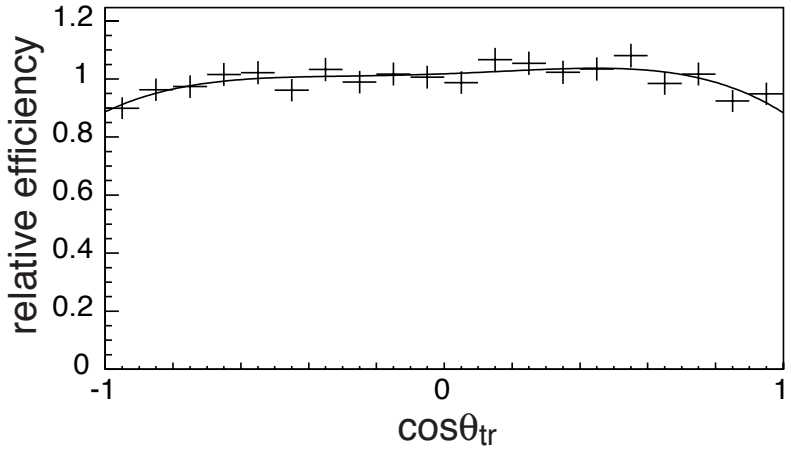

Figure 6.20: Relative signal efficiency as a function of $\cos\theta_{\rm tr}$. Fit curve is also shown.

and similarly for the background components.

The average signal fraction in the signal region, $5.27 < M_{\rm bc} < 5.29$ GeV/c^2 and $-50 < \Delta E < 30$ MeV, is estimated to be 80.6%. The average fractions of feed-across, non-resonant and combinatorial background are estimated to be 9.8%, 4.0% and 2.7%, respectively.

6.6.3 Result of fit

After flavor tagging and vertex reconstruction, 36 events remain. The unbinned maximum likelihood fit yields

$$\sin 2\phi_1 = 0.97^{+1.38}_{-1.40}(stat.) \tag{6.40}$$

If we use the averaged CP eigenvalue $\langle\xi_f\rangle_{J/\psi K^*}$=+0.62 instead of using the event-by-event angular information of decay daughters, we obtain $\sin 2\phi_1 = 1.10^{+1.50}_{-1.59}$. Thus, we can reduce about 10% of statistical error by using the angular information.

The $-2\ln\mathcal{L}$ as a function of $\sin 2\phi_1$ is shown in Fig. 6.21. The Δt distribution together with the fit result is shown in Fig. 6.22.

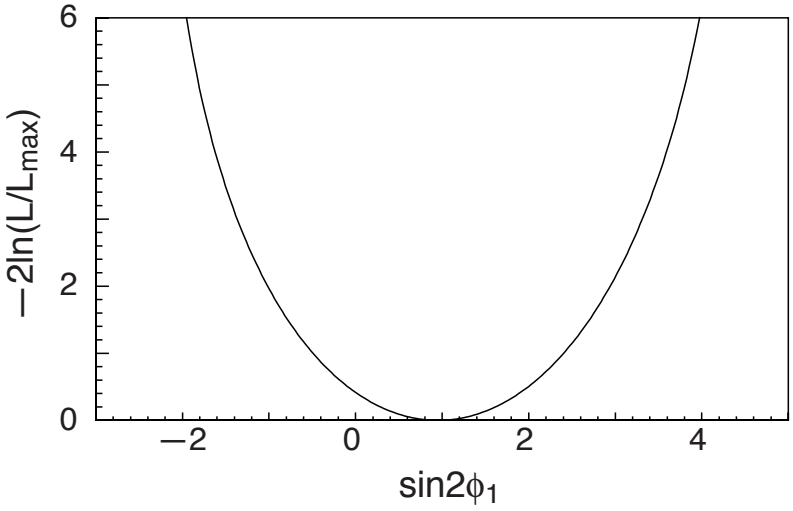

Figure 6.21: $-2\ln\mathcal{L}$ as a function of $\sin 2\phi_1$ for $B^0 \to J/\psi K^{*0}$ mode.

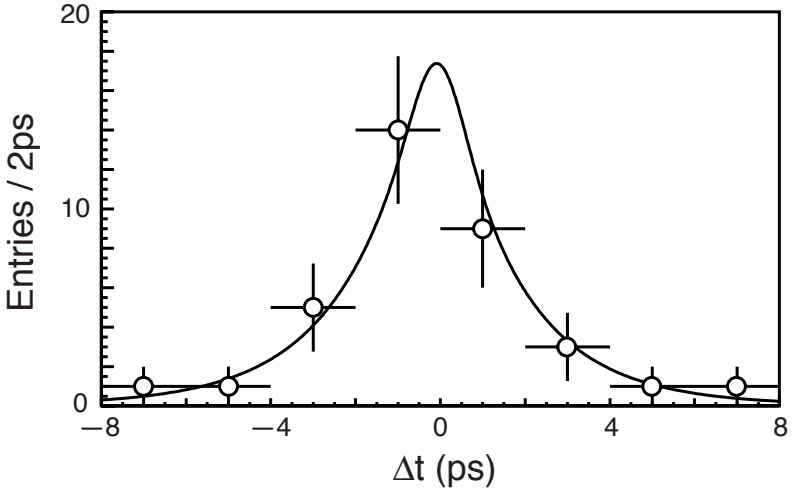

Figure 6.22: Δt distribution for $B^0 \to J/\psi K^{*0}$ mode. The result of the fit is superimposed.

Chapter 7

Check of Possible Systematics

In this chapter, we discuss the results of various checks we made to validate our results. The first two sections describe the evaluation of systematic uncertainties of $\sin 2\phi_1$ measurement with $B^0 \to J/\psi\, K_L$ and $B^0 \to J/\psi\, K^{*0}$ decays, respectively. The last section provides the detailed study of $J/\psi\, K_L$ mode to check the validation of the result.

7.1 Evaluation of Systematic Uncertainties in $B^0 \to J/\psi\, K_L$ Mode

We consider the following sources contributing to systematic errors in $B^0 \to J/\psi\, K_L$ mode. The resulting systematic errors are summarized in Table 7.2.

Background fraction

The largest contribution to the systematic error comes from the uncertainty of the background fraction. As described in the earlier chapter, the background fraction is estimated by fitting p_B^* distribution. In this procedure, because the sum of components are constrained to the total number of observed events, the derived signal and background yields are anti-correlated. To estimate the size of variation of background fraction, we have repeated the fit with the signal yield fixed to the value which is $+1\sigma$ or -1σ away from the central value. The same procedure is repeated by fixing the yield of background with K_L and without K_L. Using the obtained signal/background yield, an unbinned maximum likelihood fit is carried out to obtain $\sin 2\phi_1$. The result is summarized in Table 7.1. We take the maximum deviation

Table 7.1: Deviation of $\sin 2\phi_1$ value with signal/background yield variation.

variation	deviation (KLM)	deviation (ECL)
signal $+1\sigma$	-0.060	-0.015
signal -1σ	$+0.075$	$+0.013$
BG w/ K_L $+1\sigma$	$+0.062$	$+0.001$
BG w/ K_L -1σ	-0.056	-0.004
BG w/o K_L $+1\sigma$	-0.033	$+0.000$
BG w/o K_L -1σ	$+0.023$	-0.007

from the nominal value as the systematic error.

We also check the systematic error due to the uncertainty of the *CP* component in the background. We have repeated the fit with varying parameters of functions that determine the fractions of components inside background. Because these parameters are obtained from the MC simulation, we conservatively estimate the systematic errors by changing each parameter by $\pm 2\sigma$ taking into account the possible difference between the MC and real data. We take a quadratic sum of all contributions as the final systematic error.

Wrong tagging fraction

The contribution from the uncertainties of the wrong tagging fractions, which are summarized in Table 5.1, is studied by varying the wrong tag fractions by 1σ individually for each r bin.

Polarization of $J/\psi K^{*0}$

The fraction of *CP* even component in $B^0 \to J/\psi K^{*0}(K^{*0} \to K_L \pi^0)$ is measured from the analysis of the decay angles in $B \to J/\psi K^*$ decays [21]. We estimate the systematic error by varying this fraction by 1σ.

B lifetime and mixing parameter

The lifetime and mixing parameters of neutral B meson, τ_{B^0} and Δm_d, are fixed to the world average values, $\tau_{B^0} = 1.548 \pm 0.032$ ps and $\Delta m_d = (0.472 \pm 0.017) \times$

10^{12} $\hbar s^{-1}$. We estimate the systematic error by varying these parameters by the quoted errors.

Resolution function

We estimate the contribution from the uncertainty in the resolution function by varying parameters by 1σ and 2σ for parameters determined from data and MC, respectively, and repeating the fitting procedure.

Background shape

We have varied the parameters describing Δt distribution of combinatorial background components by their errors to estimate the systematic uncertainty.

Vertex reconstruction

Possible systematic effects due to the track quality requirements for the tagging-side B decay vertices have also been studied. The track selection criteria have been varied by 10% to estimate the systematic error associated with the estimation of the fraction of poorly reconstructed events. We further study the effect of vertex reconstruction using two different tagging-side vertex reconstruction algorithms.

We also check the effect of events which fall in the tail region of the resolution function. The selection criteria for the reduced χ^2 of the vertex fit is tightened to study the effect of events with poorly reconstructed vertices. $|\Delta t|$ region is also restricted to $|\Delta t| < 9$ ps and 6 ps in the fit to see the goodness of the description of the tail in our resolution function.

B decay point smearing

The IP constrained fit includes the uncertainty of the B decay point due to B flight length in r-ϕ plane. The uncertainty is estimated, using an MC simulation, to be $\sim$20 μm assuming a Gaussian distribution, although it is actually an exponential function. We have varied the uncertainty by $\pm$10 μm and repeated the analysis to estimate the error due to the approximation of B decay point smearing effect.

Fit bias

We check possible fit bias using signal MC. We generate MC events with $\sin 2\phi_1 = 0.8$. Using about 100,000 reconstructed MC events, we perform the fit in the same way as we do for real data. We obtain $\sin 2\phi_1 = 0.796 \pm 0.014$, which is consistent with the input value within 1σ. We then assign the statistical error as the uncertainty of the fit bias.

IP drift

As mentioned in Section 5.2.1, the IP position moves during the run. We use the event-dependent IP position to take into account this effect. We check possible remaining effect by varying the IP position by the nominal error (5 μm for x and y, 100 μm for z) and repeating the analysis.

Summary of systematic errors

The estimated systematic errors are summarized in Table 7.2.

Finally, we obtain

$$\sin 2\phi_1 = 1.31^{+0.19}_{-0.23}(stat.) \pm 0.12(syst.) \tag{7.1}$$

from $B^0 \to J/\psi\, K_L$ decays, where the first error is statistical and the second error is systematic.

7.2 Systematic Uncertainties in $B^0 \to J/\psi\, K^{*0}$ Mode

We consider the following sources contributing to systematic errors for $B^0 \to J/\psi\, K^{*0}$ mode. The results are summarized in Table 7.3.

Background fraction

For $B^0 \to J/\psi\, K^{*0}$ mode, we determine the background probability based on the fit to ΔE and $M_{\rm bc}$ distributions. We estimate the systematic error due to the uncertainty of background fraction by repeating fit with varying the parameters of these functions by $\pm 1\sigma$.

We also consider the systematic error due to the uncertainty of non-resonant and feed-across background fractions. We estimate the systematic error by changing

Table 7.2: List of systematic errors for $J/\psi K_L$ mode.

Source	+error	−error
Background fraction	+0.08	−0.08
CP contents in BG	+0.04	−0.03
Wrong tag fractions	+0.03	−0.03
Polarization of $J/\psi K^{*0}$	+0.01	−0.01
τ_{B^0} & Δm_d	+0.02	−0.02
Resolution function	+0.04	−0.04
Background shape	+0.01	−0.01
Vertex reconstruction	+0.05	−0.05
B decay point	+0.01	−0.01
MC statistics	+0.01	−0.01
IP position	< 0.01	< 0.01
Total	+0.12	−0.12

these fractions by ±50%. After calculating deviation for each parameter, we add all deviations in quadrature.

Wrong tag fraction

The contribution from the uncertainty of wrong tag fractions is estimated similarly to the $J/\psi K_L$ mode.

Polarization of $J/\psi K^{*0}$

We estimate the systematic error from the uncertainty of the fraction of CP odd component by varying this fraction by 1σ.

B lifetime and mass difference, resolution function, background shape

The systematic uncertainties originating from these sources are estimated in the same way as $J/\psi K_L$ mode.

Effective lifetime of feed-across background

The lifetime of B^0 is used as the effective lifetime of feed-across background. However, for the feed-across background, we may have contamination from charged particles from B_{CP} into tagging-side vertex reconstruction, as we experienced with $J/\psi K_L$ background. Therefore, we study the effect of possible difference of the effective lifetime by varying the parameter by $+0.1$ and -0.2 ps and repeating the fit.

Summary of systematic errors

The estimated systematic errors are summarized in Table 7.3.

Finally, we obtain

$$\sin 2\phi_1 = 0.97^{+1.38}_{-1.40}(stat.) \pm 0.19(syst.) \tag{7.2}$$

from $B^0 \to J/\psi K^{*0}$ decays, where the first error is statistical and the second error is systematic.

Table 7.3: List of systematic errors for $B^0 \to J/\psi K^{*0}$ mode.

Source	+error	−error
Background fraction	+0.13	−0.09
Wrong tagging fractions	+0.12	−0.14
Polarization of $J/\psi K^{*0}$	+0.05	−0.07
τ_{B^0} & Δm_d	+0.06	−0.06
Resolution function	+0.02	−0.02
Background shape	+0.01	−0.01
Feed-across lifetime	+0.02	−0.01
Total	+0.19	−0.19

7.3 Validation of the Result

In the previous sections, we present the measurement of $\sin 2\phi_1$ using $B^0 \to J/\psi K_L$ and $B^0 \to J/\psi K^{*0}$ decays. In this section, we present the result of various checks we made to confirm our measurement. We discuss only $B^0 \to J/\psi K_L$ mode here,

because it has much smaller measurement error and thus larger significance than $B^0 \to J/\psi K^{*0}$.

Control samples

Control samples with non-CP eigenstates are analyzed in order to verify the proper time reconstruction, tagging algorithm and likelihood fitting procedures. The samples for $B^0 \to J/\psi K^*(K^+\pi^-)$, $B^0 \to D^{(*)+}\pi^-$, $B^0 \to D^+\rho^-$, and $B^0 \to D^{*-}\ell^+\nu$ decays are selected in the same way as in flavor tagging study described in Appendix C. We perform the same fit as the CP-eigenstate modes to these control samples. The results are summarized in Table 7.4. The results show no systematic tendency and they are consistent with zero. When all the modes are combined, we obtain asymmetry $= 0.05 \pm 0.04$. Thus, we find no statistically significant bias from the control samples. Figure 7.1 shows the asymmetry plot for control samples made with the same procedure as Fig. 6.15.

Table 7.4: CP fit results for control samples.

	$J/\psi K^{*0}$	$D^{(*)^+}\pi^- + D^+\rho^-$	$D^{*-}\ell^+\nu$
No. of fitted events	816	5560	10232
Asymmetry	0.01 ± 0.14	0.12 ± 0.06	0.01 ± 0.05

Flavor blind fit

If we do not identify the flavor of B mesons, there should be no asymmetry. This provides a check of possible bias in the fitting procedure. By forcing q value of all events to be $+1$, we obtain

$$\text{Asymmetry (flavor blind)} = -0.26 \pm 0.27 \tag{7.3}$$

which is consistent with zero.

Ensemble test

We use a *toy Monte Carlo* to test the likelihood fit procedure. It is designed to create a sample of events whose distribution is exactly (within statistical fluctuations) described by the likelihood function. In this way, the expected values of the fit

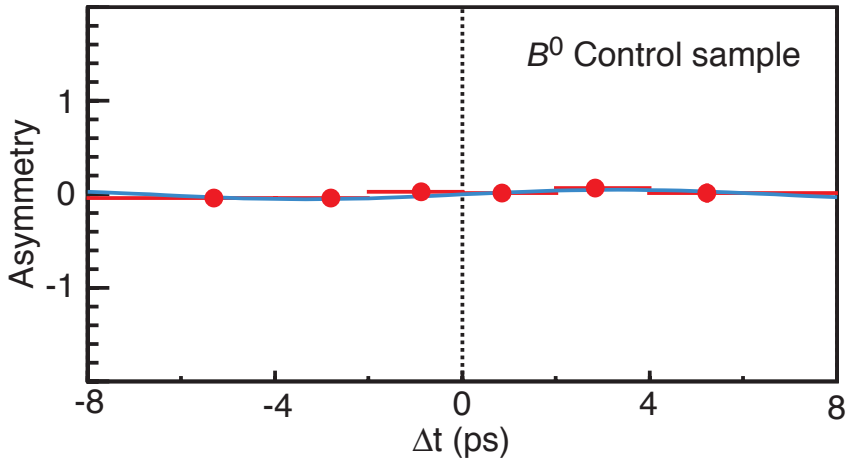

Figure 7.1: The asymmetry plot for control samples. Separate fits are made to each Δt bin.The curve is the result of the global fit, Asymmetry= 0.05.

parameters are known exactly, and one can thus determine directly whether the values returned by the likelihood fit are unbiased estimation of the fit parameters.

We generate 5000 sets of toy Monte Carlo data samples, each containing 523 events, based on the PDF we use. We cannot generate MC data with $\sin 2\phi_1$ larger than unity, because the PDF may become negative for such input values. We still can obtain $\sin 2\phi_1$ value larger than one as a result of real measurement, due to the finite measurement error. The check is done for the input value of $\sin 2\phi_1 = 0.99$, which is the value we obtain from combined fit with all CP eigenstate available [62].

Figure 7.2 shows the distributions of the (output value of the fit)−(input value), (output−input) / (error of fit), the positive and negative errors, and the likelihood value returned from the fits. The vertical lines indicate the value we obtain from the fit to the real data. The fit provides correct $\sin 2\phi_1$ value and the width of the distribution is consistent with the error of the fit, even if we observe $\sin 2\phi_1$ greater than one.

Fit with sideband region

We check the possible bias on CP asymmetry in background which we treat with no asymmetry. We fit $J/\psi K_L$ candidates in background region ($1.0 < p^*_B < 2.0$ GeV/c) treating all the events as $J/\psi K_L$ signal (the fraction of events with their own CP is expected to be negligible). The result is (Asymmetry) = 0.23 $\pm$ 0.26 and thus no bias is seen.

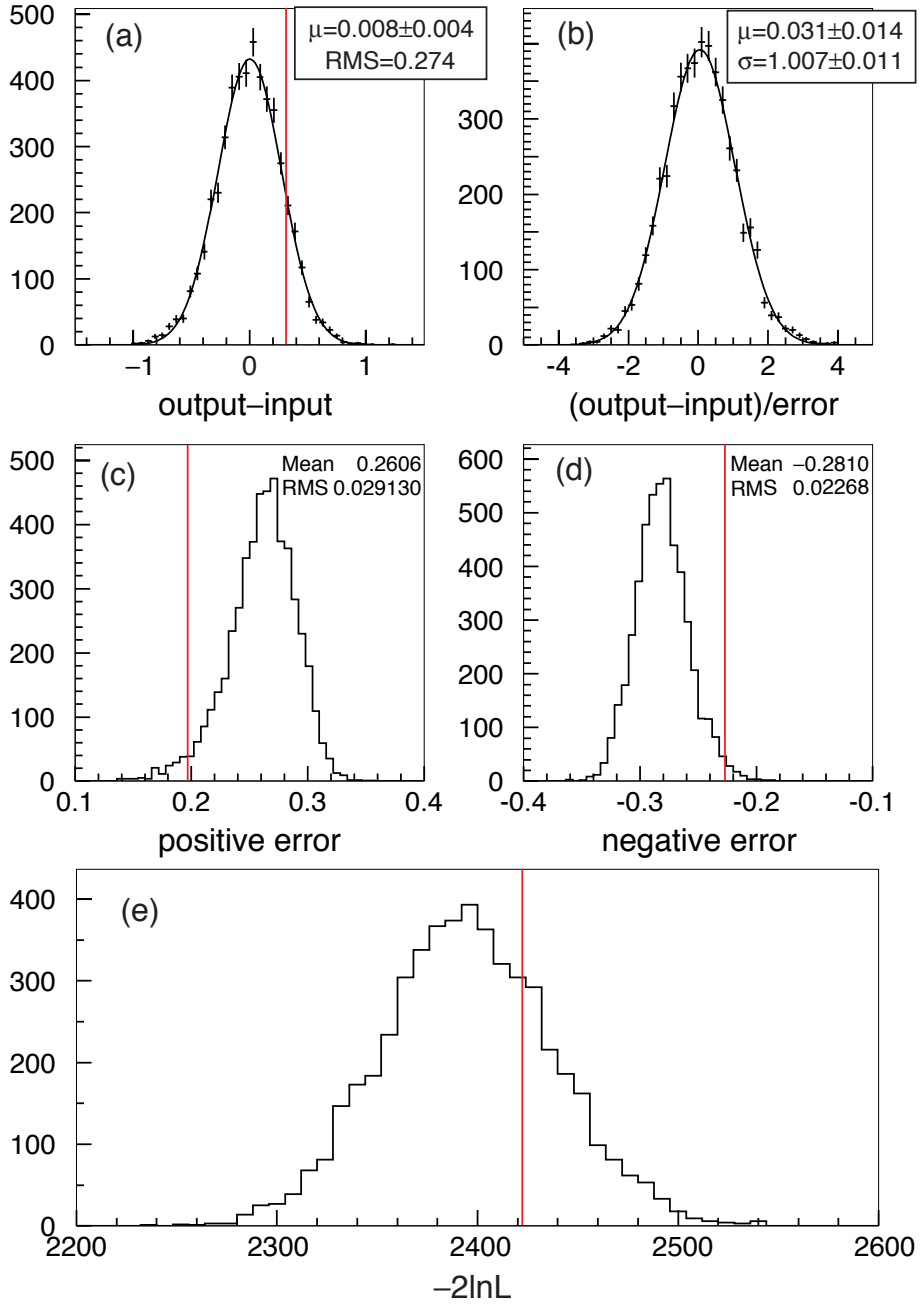

Figure 7.2: The result of ensemble test. Each plot shows the distribution of (a) result of fit, (b) (output−input) / (error of fit), (c) positive and (d) negative errors of $\sin 2\phi_1$, and likelihood value ($-2\ln\mathcal{L}$) returned from the fits.

Result with floating τ_B and Δm_d

We try to fit $\sin 2\phi_1$ and other B meson parameters, τ_{B^0} or Δm_d simultaneously to check possible bias.

If we fit $\sin 2\phi_1$ and τ_{B^0}, we obtain

$$\sin 2\phi_1 = 1.29 \pm 0.23, \qquad \tau_{B^0} = 1.60 \pm 0.10\,\text{ps}. \tag{7.4}$$

In case of $\sin 2\phi_1$ and Δm_d, we obtain

$$\sin 2\phi_1 = 1.46 \pm 0.25, \qquad \Delta m_d = (0.632 \pm 0.084) \times 10^{12}\,\hbar\text{s}^{-1}. \tag{7.5}$$

These values are consistent with the nominal values.

Fit without $|\lambda| = 1$ assumption

If we do not assume $|\lambda| = 1$ (see Section 2.3.2), the signal PDF (Eq. 6.6 and 6.7) is expressed as:

$$\mathcal{P}_{\text{sig}}(\Delta t, q, w_l, \xi_f; \sin 2\phi_1) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}(1+|\lambda|^2)} \times \left\{ \frac{1+|\lambda|^2}{2} + q\xi_f(1-2w_l) \left[\text{Im}\lambda \sin(\Delta m_d \Delta t) \; - \frac{1-|\lambda|^2}{2} \cos(\Delta m_d \Delta t) \right] \right\}. \tag{7.6}$$

We assume $|\lambda| = 1$ throughout this analysis based on the theoretical expectation of the SM. We also perform a fit using above general PDF with $\text{Im}(\lambda)/|\lambda| =$ "$\sin 2\phi_1$" and $|\lambda|$ as free parameters, otherwise everything is kept the same. We obtain $|\lambda| = 1.09 \pm 0.24$ and "$\sin 2\phi_1$"$= 1.31 \pm 0.23$, consistent with the SM prediction.

Flavor tag quality dependence

We perform the fit for $B_{tag} = B^0$ ($q = +1$) and $B_{tag} = \overline{B}^0$ ($q = -1$) samples separately. We also check the possible dependence of $\sin 2\phi_1$ on r intervals of the flavor tagging. The result is listed in Table 7.5. No systematic variation is seen.

We further check the possible bias due to the possible difference of wrong-tag fractions for $q = +1$ and $q = -1$. We have measured wrong tag fractions for $q = +1$ and $q = -1$ separately and repeated the fit. The resulting $\sin 2\phi_1$ is 1.30 ± 0.23 and no bias is seen.

Data set dependence

We check consistency between different data set, dividing data set into three subsamples listed in Fig. 7.6. No statistically significant variation is seen.

Table 7.5: q and r dependence of $\sin 2\phi_1$ fit result. Only statistical errors are shown.

q	$+1$	-1
No. of events	246	277
$\sin 2\phi_1$	1.20 ± 0.36	1.40 ± 0.30

r interval	0.0-0.5	0.5-0.75	0.75-0.875	0.875-1.0
No. of events	266	129	59	69
$\sin 2\phi_1$	2.52 ± 1.36	0.93 ± 0.57	1.62 ± 0.43	1.23 ± 0.33

Table 7.6: Three sub data set to check data set dependence.

	Set 1	Set 2	Set 3
Period	01/2000–12/2000	01/2001–04/2001	04/2001–07/2001
Integrated luminosity (fb^{-1})	10.2	8.2	10.7
No. of events	168	156	199
$\sin 2\phi_1$	0.43 ± 0.49	2.03 ± 0.59	1.65 ± 0.28

S/N ratio

If we underestimate S/N ratio, we may over-subtract the background and find larger asymmetry (and *vice versa*). We perform several checks to ensure our evaluation of S/N ratio is correct.

dependence on flavor tag quality There could exist some correlation between the quality of flavor tagging and S/N. For example, in $B^{\pm} \to J/\psi\, K^{*\pm}$ ($K^{*\pm} \to K_L\, \pi^{\pm}$) background event, $\pi^{\pm}$ could be used for the flavor-tagging and bias the r value. We check the possible r dependence of S/N. Figure 7.3 shows the p_B^* distributions for different r intervals. We further check the r dependence using the MC. The result is summarized in Table 7.7. No clear dependence is seen for both cases.

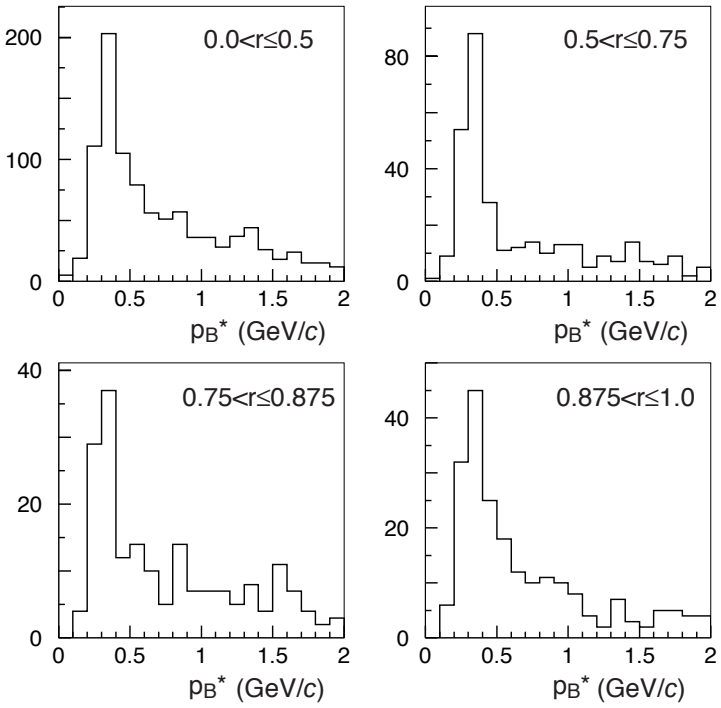

Figure 7.3: p_B^* distributions for different r interval.

Table 7.7: S/N for different r intervals in the MC.

r	0.0-0.5	0.5-0.75	0.75-0.875	0.875-1.0	all
S/N	1.51 ± 0.06	1.47 ± 0.10	1.50 ± 0.12	1.38 ± 0.10	1.48 ± 0.04

Table 7.8: The purity for each data set.

data set	Set 1	Set 2	Set 3	all
purity(%)	63.0	61.1	68.2	60.9

data set dependence Variation of K_L detection efficiency would result in unstable S/N ratio and bias the result. We estimate the purity for three sub-data sets as defined in Table 7.6. The result is shown in Table 7.8. Considering the fact that we have about 20% error on the number of signal for each sub-data set, the fluctuation of the signal fraction is well within the statistical error.

Chapter 8

Discussions and Conclusions

8.1 Comments on Measured $\sin 2\phi_1$ Greater than Unity

As presented in the previous chapters, we have measured the *CP* asymmetry parameter $\sin 2\phi_1$ to be $1.31^{+0.19}_{-0.23} \pm 0.12$ using the $B^0 \to J/\psi K_L$ decays. It might seem strange that we have observed the $\sin 2\phi_1$ value beyond the physics boundary, $|\sin 2\phi_1| \leq 1$. However, the observable asymmetry is smaller than the true value due to the dilution such as imperfect flavor tagging, existence of background, and finite vertex resolution. Even though the asymmetry in the number of B^0 and $\overline{B}^0$ mesons are always within ± 1, the statistical fluctuation of the dilution can result in the measured $\sin 2\phi_1$ value greater than one, especially when the true value is close to the physical boundary.

As described in Section 7.3, our fitting procedure is confirmed to yield correct $\sin 2\phi_1$ value and errors even if the true value is very close to the physical boundary. Thus, we conclude the observed $\sin 2\phi_1$ greater than one is due to the statistical fluctuation, not due to the incorrectness of our analysis.

8.2 Significance of the Measurements

We have calculated the statistical significance of the *CP* asymmetry observed using $B^0 \to J/\psi K_L$ decays. We calculate 95% confidence interval using the method proposed by Feldman and Cousins [60]. The detail of Feldman-Cousins method is described in Appendix D. We assume that the error is a symmetric Gaussian and take the σ to be the larger one of the asymmetric errors. We find the lower bound

of $\sin 2\phi_1$ to be 0.76 at 95% confidence level. A zero value for $\sin 2\phi_1$ is ruled out at a level greater than 5σ. Thus, we have observed *CP* violation in the neutral B meson system with sufficient statistical significance.

For $B^0 \to J/\psi K^{*0}$ mode, we cannot make any statistically significant statement from our measurement for present data set. However, this is the first attempt to measure $\sin 2\phi_1$ using the decay angle information of the decay daughters. Although the current result is statistically limited, this analysis demonstrates the capability of *CP* asymmetry measurement in the case the *CP* eigenvalue depends on the decay angles.

8.3 Comparison with Other Measurements

Other measurements at Belle

Measurement of $\sin 2\phi_1$ using charmonium+K_S modes, namely $J/\psi K_S$, $\psi(2S)K_S$, $\chi_{c1}K_S$, and $\eta_c K_S$ is also reported by Belle. The detail of this measurement is described in [61, 62]. The results are summarized in Fig. 8.1, together with the results presented in this thesis. The χ^2 of the three measurement, except $B^0 \to J/\psi K^{*0}$ which error is large and has no statistical significance, is 2.3 for the number of degrees of freedom 2. Thus, our measurements are consistent with others. Combining all modes, we obtain

$$\sin 2\phi_1 = 0.99 \pm 0.14(stat.) \pm 0.06(syst.). \tag{8.1}$$

The lower bound is found to be 0.70 at the 95% confidence level with the combined result. Figure 8.2 shows the 68% and 95% confidence level regions of $\sin 2\phi_1$, together with constraints of $(\overline{\rho}, \overline{\eta})$ from other measurements of K and B decays [22]. Our result is consistent with the values allowed by the constraints of the KM model. Note that there are two allowed region of ϕ in the shown area due to two-fold ambiguity from $\sin 2\phi_1$ [1]. Because our measurement has resulted in the $\sin 2\phi_1$ value close to the unity, these two regions are overlapped.

Measurements at other experiments

Figure 8.3 shows the comparison of measured $\sin 2\phi_1$ value with other experiments. Our result is consistent with other measurements [23, 24, 25, 26, 63].

[1] Exactly speaking, we have another two allowed region in $\overline{\eta} < 0$ region which is not displayed.

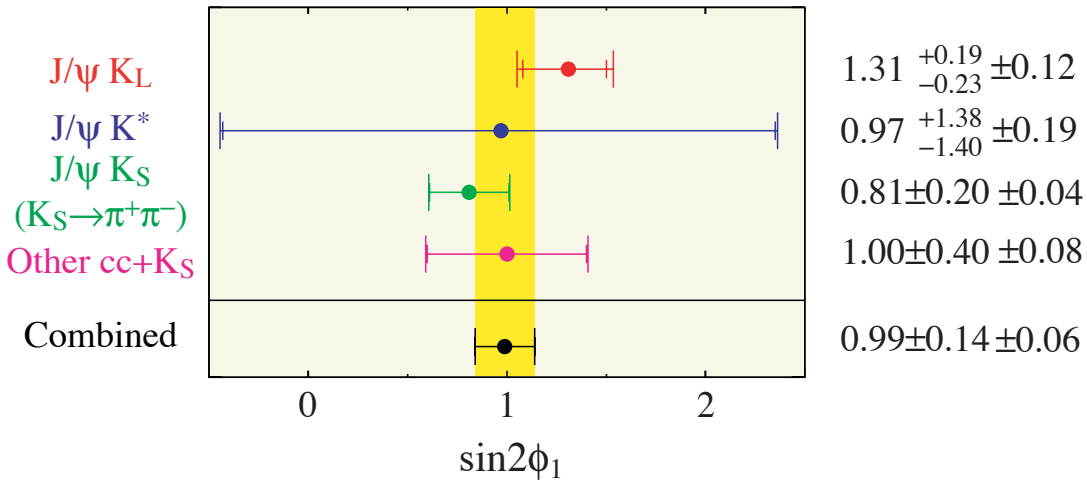

Figure 8.1: Comparison of $\sin 2\phi_1$ measurements in Belle. The first errors are statistical and the second are systematic.

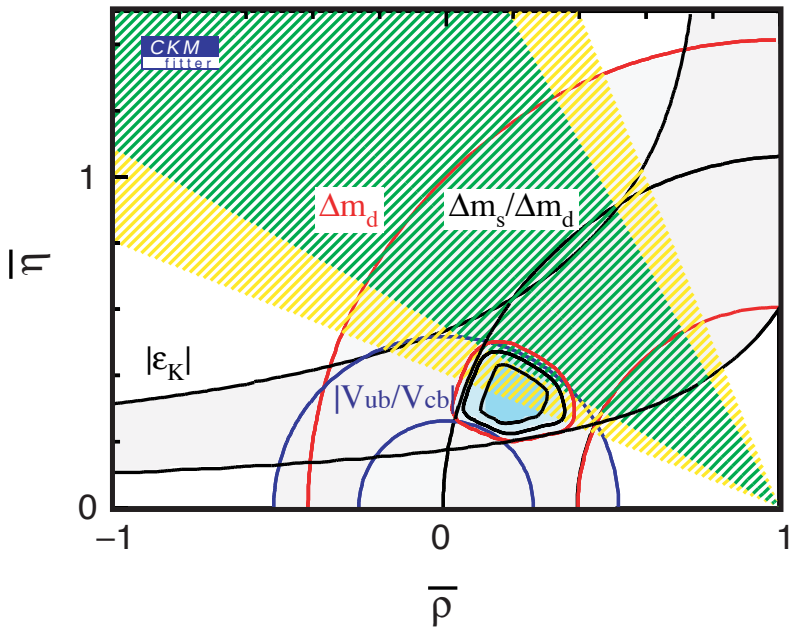

Figure 8.2: 68% and 95% confidence level regions of $\sin 2\phi_1$ in $(\overline{\rho}, \overline{\eta})$ plane for Belle combined result. Also shown are constraints from other measurements of K and B decays [22].

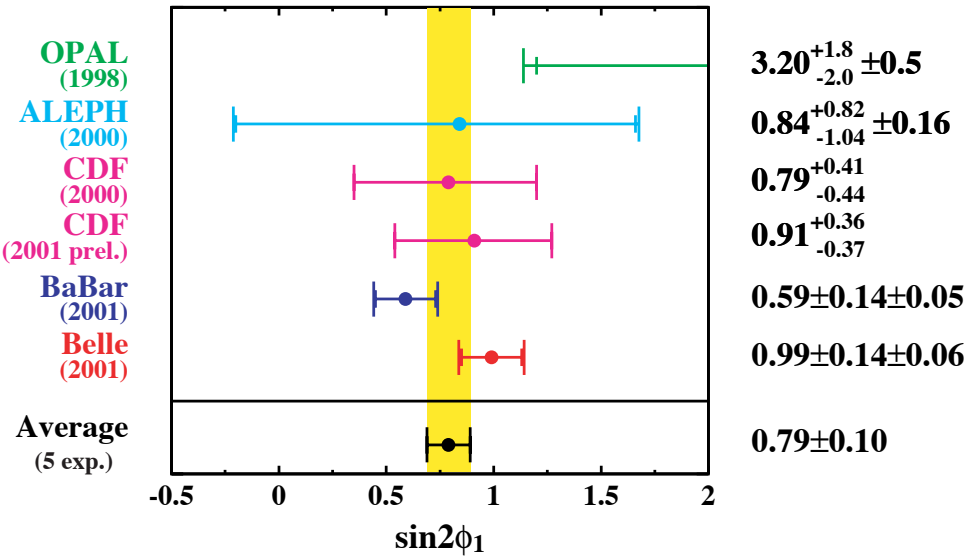

Figure 8.3: Comparison of measured $\sin 2\phi_1$ values at each experiment. The preliminary result of CDF group is not included in the average.

8.4 Comments for Future Measurement

The uncertainty of signal/background fractions in $J/\psi K_L$ mode is the dominant source of systematic error for the measurement of $\sin 2\phi_1$ using the $J/\psi K_L$ mode. It is also one of the major sources of the systematic error in the measurement using all *CP* eigenstates. Table 8.1 shows the major sources of the systematic errors for the combined result.

Belle plans to accumulate about 80 fb^{-1}of data by the summer 2002. The statistical error will be less than 0.1 by combining all modes and $\simeq$ 0.14 by $J/\psi K_L$ only. Therefore the significance of the systematic error will considerably increase at that time. Among the sources of the systematic error listed in Table 8.1, the vertex reconstruction and the resolution function are now being studied as a consequence of the B meson lifetime measurement. It requires more precise understanding of the proper time resolution than $\sin 2\phi_1$ measurement. The uncertainty of wrong tagging fraction is expected to be limited by the statistics of samples used for the measurement, which scales with the amount of total data.

The uncertainty of signal/background fractions in the $J/\psi K_L$ mode could be the dominant source of the systematic error for the combined $\sin 2\phi_1$ measurement in near future. Furthermore, it would become a dominant source of the total (*stat.* + *syst.*) error in the measurement of $\sin 2\phi_1$ in the $J/\psi K_L$ mode. The experience

Table 8.1: Major sources of systematic errors for combined result.

Source	systematic error of $\sin 2\phi_1$
Vertex reconstruction	0.04
Wrong tag fraction	0.03
Resolution function	0.02
Background fraction in $J/\psi K_L$	0.02

tells us that the signal fraction can be determined in a better precision when we have more data, hence the systematic uncertainty will be reduced to some extent. However, the contents of background are determined by the MC simulation and limited by the correctness of the simulation. Most of the background comes from $B \to J/\psi X$ processes with known branching fractions and can be studied with the MC simulation. We need to understand the performance of the detector, especially the KLM, to improve our understanding of the backgrounds.

The systematic uncertainty is also improved by reducing the fraction of background itself, *i.e.* by increasing the S/N ratio. When the systematic uncertainty become dominant in the total error, we have to optimize our analysis to maximize the total significance of the measurement. We can improve S/N by tightening the requirement on the likelihood ratio and/or by adding more discriminating variables to the likelihood ratio. Either case will require a careful study using an MC simulation and the real data to maximize the significance of the measurement. For example, we may be able to utilize the correlation between the missing momentum and the K_L cluster direction. Because the energy of the K_L is not measured by the KLM, these two values should be correlated in the signal. However, the missing momentum is also correlated with semi-leptonic decays of tagging-side B meson in which neutrinos are produced but undetected. Hence careful study will be necessary to identify the effect on flavor tagging when we utilize this variable.

As for $B^0 \to J/\psi K^{*0}$ mode, the measurement precision continues to be dominated by the statistical error. We need $\mathcal{O}(300)$ fb^{-1} to make statistically meaningful measurement if further significant improvement in the reconstruction and the fitting procedure are not achieved. We will need careful optimization of selection criteria and inclusion of full angular information of decay products to improve the precision in this mode.

8.5 Conclusions

We have measured the *CP* asymmetry parameter $\sin 2\phi_1$ at the KEKB asymmetric e^+e^- collider using a data sample of 29.1 fb^{-1} recorded on the $\Upsilon(4S)$ resonance with the Belle detector.

Using 523 candidate events in decays of neutral B mesons to the *CP* eigenstate $J/\psi K_L$, where flavor tagging and the vertex reconstruction have been successful, an unbinned maximum likelihood fit is carried out to extract $\sin 2\phi_1$. We obtain

$$\sin 2\phi_1 = 1.31^{+0.19}_{-0.23}(stat.) \pm 0.12(syst.). \quad [J/\psi K_L]$$

We have also measured $\sin 2\phi_1$ using 36 candidate events in decays of neutral B mesons to $J/\psi K^{*0}$, followed by K^{*0} decays into $K_S\pi^0$. We use the angular information of the decay daughters of the B mesons to separate *CP* odd and even components in the final state. Using the fraction of *CP* odd component $f_{odd} = 0.19 \pm 0.06$ which is determined from a full-angular analysis of B decays into $J/\psi K^*$, we obtain

$$\sin 2\phi_1 = 0.97^{+1.38}_{-1.40}(stat.) \pm 0.19(syst.). \quad [J/\psi K^{*0}]$$

These results are consistent with the results obtained using other modes at Belle [61, 62].

Combining the decay modes presented in this thesis with decay modes involving charmonium and K_S, we obtain $\sin 2\phi_1 = 0.99 \pm 0.14(stat.) \pm 0.06(syst.)$. These results are consistent with the values allowed by the constraints of the KM model as well as with other measurements [23, 24, 25, 26, 63].

We conclude that we have observed a large *CP* violation in the neutral B meson system. A zero value for $\sin 2\phi_1$ is ruled out at a level greater than 5σ and the lower bound of $\sin 2\phi_1$ is calculated to be 0.76 at 95% confidence level by the $B^0 \to J/\psi K_L$ mode alone. This is the first observation of *CP* violation outside the kaon system.

The precision of these measurements will be still dominated by statistical uncertainty for a while. However, the systematic uncertainties in the $J/\psi K_L$ mode has to be carefully studied for the future precision measurement. In particular, we have to reduce the uncertainties of the signal fraction and background components determination. They are the dominant source of the systematic uncertainty of $\sin 2\phi_1$ measurement in the $J/\psi K_L$ mode at present. They may become major source of the systematic error in the combined measurement in a few years.

Given the systematic error is controlled within reasonable level, this measurement will provide a severe constraint to the CKM triangle in near future.

Appendix A

Alignment of the SVD

The Belle SVD is designed to provide precise vertex information which is necessary for the measurement of time dependent CP asymmetry. However, we must align the position of DSSD's during operation with sufficient precision to achieve the designed vertex resolution. In this chapter, we describe the alignment procedure of the SVD.

A.1 Definitions of Parameters

We present in this section definitions of parameters that are used in the alignment study of the SVD.

A.1.1 Definition of DSSD local coordinate

The Belle SVD consists of 102 double-sided silicon strip detectors (DSSDs). The deviation of position of one DSSD is described by shift and rotation from its nominal position. We define the DSSD local coordinate system as shown in Fig. A.1. The DSSD plane is defined as the x-z plane, where x is ϕ-direction (i.e. shorter side of DSSD). Local z coordinate is along longer side of the DSSD, which is the same direction as the Belle global z direction when the DSSD is in its nominal position. y axis of the DSSD local frame is perpendicular to the DSSD plane. Since SVD consists of 102 DSSD's, we have 612 parameters in total.

A.1.2 Definition of helix parameters

The definition of helix parameters used to represent tracks in Belle experiment is described in [65]. Here, we extract and summarize only necessary information to

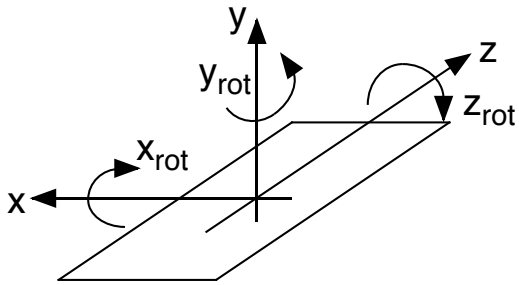

Figure A.1: DSSD local coordinate system

explain our alignment study.

The helix parameters are defined with respect to a pivotal reference point (pivot), $\mathbf{x_0} = (x_0, y_0, z_0)$. Usually in the SVD region, the pivots for tracks are taken to be the origin of the Belle coordinate system, $\mathbf{x_0} = (0, 0, 0)$. There are five helix parameters:

d_ρ: the signed distance of the helix from the pivot in x-y plane;

ϕ_0: the azimuthal angle to specify the pivot with respect to the helix center;

κ: $1/p_t$ (reciprocal of the transverse momentum). the sign of κ represents the charge of the track assigned by the track fitting;

d_z: the signed distance of the helix from the pivot in the z direction;

$\tan\lambda$: the slope of the track, tangent of the dip angle.

Charged particle in a uniform magnetic field is represented by a helical trajectory. The position along the helix is given by

$$\begin{cases} x = & x_0 + d_\rho \cos\phi_0 & +\frac{\alpha}{\kappa}\{\cos\phi_0 - \cos(\phi_0 + \phi)\} \\ y = & y_0 + d_\rho \sin\phi_0 & +\frac{\alpha}{\kappa}\{\sin\phi_0 - \sin(\phi_0 + \phi)\} \\ z = & z_0 + d_z & -\frac{\alpha}{\kappa}\tan\lambda \cdot \phi, \end{cases} \tag{A.1}$$

where α is the magnetic-field-constant, $\alpha = 1/c\mathrm{B}$, ϕ is the turning angle, that is an internal parameter with a sign, (for instance, ϕ has a negative sign for out-going positive tracks from the pivot.) and determines the location. Figure A.2 shows schematic representations of the helix parametrization for positive and negative charged tracks.

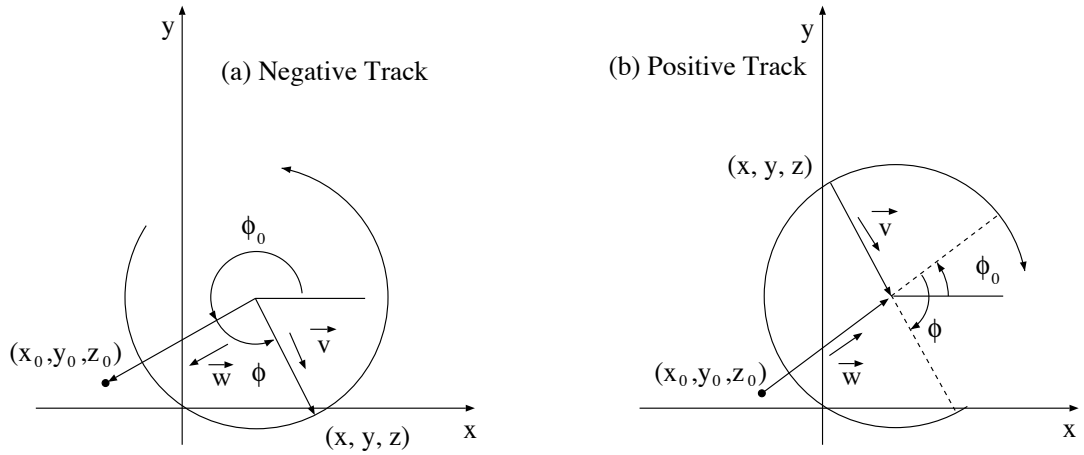

Figure A.2: Schematic representations of the helix parametrization for (a) negative and (b) positive charged tracks.

A.2 Overview of SVD Alignment

The alignment of Belle SVD consists of two steps. One is *internal* alignment of SVD, in which each DSSD is aligned with each other. As noted in the previous section, we have 612 parameters in total. The other is *CDC-SVD* alignment (also referred as *global* alignment). In the SVD-CDC alignment, the SVD is treated as one rigid body and aligned with respect to the CDC. Number of parameter is six (three shifts+three rotation angles) in this case.

We store alignment parameters in a database which can manage run dependence of alignment constants. We define a table based on PANTHER data management system used in Belle to store the alignment parameters. It contains six parameters, (xshift, yshift, zshift, xrot, yrot, zrot) for each DSSD and this information is used for simulation and reconstruction. Global alignment parameters are merged with internal alignment parameter by global transformation of SVD and only merged parameters are stored in the database.

The effect of misalignment of the SVD on $\sin 2\phi_1$ measurement is studied [64]. If there is imperfection of alignment, it may become a source of systematic error. From a Monte Carlo simulation study, it is found that the effect on $\sin 2\phi_1$ measurement is negligible if we could align DSSDs in the accuracy of < 25 μm for shifts and < 1 mrad for rotations (in σ). We set the initial goal of our alignment precision based on this study to be 25 μm for shifts and < 1 mrad for rotations.

A.3 Internal Alignment

Internal alignment is based on a minimization of χ^2 of residuals between tracks constrained by the SVD and hits recorded on the SVD. Charged particle tracks are found by the CDC and track helix parameters and their error are calculated by Kalman filtering method [39] using only CDC information. The internal alignment of the SVD is strongly sensitive to systematics of input data, *e.g.* track parameters defined by CDC. In order to be as much independent of such systematics as possible, the SVD internal alignment is based on the concept of a constrained residual. This is a residual where only a part of the track information is calculated by the outer tracking device, while the rest is determined with SVD hits. Three types of constrained residuals are used in the alignment procedure:

single constrained residual Taking the track direction and curvature from the CDC, the track is constrained to pass through one of the SVD hits and the residual with the other hit is measured.

double constrained residual Only the curvature of track is taken from the CDC and the direction and position is defined by SVD hits by constraining track to pass through two SVD hits. We use this constrained track and the other hit which is not used in the constraint to calculate the double constrained residual. The double constrained residual is more robust against the systematics of CDC parameters than the single constraint residual, because it only take the curvature from the CDC.

multi-constrained residual Taking the curvature from the CDC, a track is constrained to more than two SVD hits to minimize the residual between the track and the hits. The multi-constrained residual is the residual between this constrained track and the other hit which is not used in the constraint. Because we can use as many hits as possible in this type of constraint, multi-constrained residual could have the largest power in the alignment.

Figure A.3 shows a schematic view of constrained tracks and residuals.

We have to calculate the intersection of the constrained-track with a DSSD plane. Since a track is not a straight line but a helix in the magnetic field, we cannot calculate the intersection analytically when the plane is not parallel to the magnetic field. We numerically calculate intersection by approximating the track with a straight line near the DSSD plane and iterating calculation.

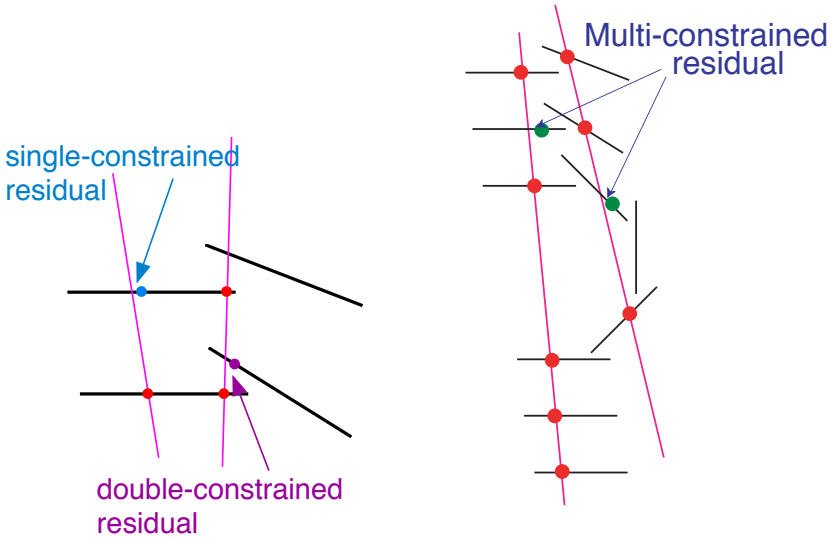

Figure A.3: Schematic view of constrained tracks and residuals.

From the calculated intersection point of the track with the DSSD, $\mathbf{x_{trk}} = (x_{trk}, 0, z_{trk})$ and the reconstructed hit position on the DSSD, $\mathbf{x_{hit}} = (x_{hit}, 0, z_{hit})$, we define the χ^2 as

$$\chi^2 = \sum_{hit} \left(\frac{(res_x)^2}{{\sigma_x}^2} + \frac{(res_z)^2}{{\sigma_z}^2} \right) = \sum_{hit} \left(\frac{(x_{trk} - x_{hit})^2}{{\sigma_x}^2} + \frac{(z_{trk} - z_{hit})^2}{{\sigma_z}^2} \right) \tag{A.2}$$

where res_x and res_z are residual along x and z direction in the DSSD local frame and σ_x and σ_z are errors of residual. We can find the best alignment parameters by minimizing this χ^2.

We can directly minimize χ^2 using minimization packages like MINUIT, but it generally consumes a lot of time. To reduce the computation time, we use a linear approximation to find the answer by an analytical calculation.

The intersection point of a track with the DSSD depends on the alignment parameters of the DSSD. When all the parameter are small, we can express trk'_x and trk'_z, the track intersection in the DSSD local frame after changing alignment parameters, neglecting the higher order terms as:

$$trk'_x \simeq trk_x - x_{shift} + \frac{p_x}{p_y} y_{shift} - \frac{p_x}{p_y} z x_{rot} - z y_{rot} + \frac{p_x}{p_y} x z_{rot}, \tag{A.3}$$

$$trk'_z \simeq trk_z - z_{shift} + \frac{p_z}{p_y} y_{shift} - \frac{p_z}{p_y} z x_{rot} + x y_{rot} + \frac{p_z}{p_y} x z_{rot}. \tag{A.4}$$

Using these approximation, we can analytically solve the equations and calculate the alignment parameters which minimize the χ^2 defined as Eq. (A.2).

At the minimum point of χ^2,

$$\frac{\partial \chi^2}{\partial p^a} = 0 \tag{A.5}$$

should hold, where p^a is xshift, yshift, zshift, xrot, yrot, and zrot. This results in six linear equations, which can be easily solved. With defining necessary vectors and matrix as

$$\mathbf{a_x} = \left(-1, \frac{p_x}{p_y}, 0, -\frac{p_x}{p_y}z, -z, \frac{p_x}{p_y}x\right)^T \tag{A.6}$$

$$\mathbf{a_z} = \left(0, \frac{p_z}{p_y}, -1, -\frac{p_z}{p_y}z, x, \frac{p_z}{p_y}x\right)^T \tag{A.7}$$

$$\mathbf{v} = \sum_{hits}\left(\frac{res_x}{{\sigma_x}^2}\mathbf{a_x} + \frac{res_z}{{\sigma_z}^2}\mathbf{a_z}\right) \tag{A.8}$$

$$\mathbf{M} = \sum_{hits}\left(\frac{\mathbf{a_x}\cdot\mathbf{a_x}^T}{{\sigma_x}^2} + \frac{\mathbf{a_z}\cdot\mathbf{a_z}^T}{{\sigma_z}^2}\right) \tag{A.9}$$

where p_x, p_y and p_z are the momentum of the track at the intersection, x and z are x and z position of the intersection point, res_x and res_z are the residuals along x and z axis, σ_x and σ_z are the errors of each residual (Note that all variables are defined in DSSD local frame). The alignment parameters $\mathbf{p}$ can be obtained with just one matrix inversion and production,

$$\mathbf{p} = \mathbf{M}^{-1}\cdot\mathbf{v} \tag{A.10}$$

where $\mathbf{p} = (xshift, yshift, zshift, xrot, yrot, zrot)^T$.

We have to estimate errors of residuals to obtain good alignment parameters. Errors of residual has two components; error of the track and that of the hit. The error of a single-constrained track is considered to be dominated by the error of the direction of the track, which is determined by the CDC. We can estimate the error of single-constrained tracks from the error matrix of track parameters determined by the CDC. For the double-constrained track, error is derived from the error of the SVD hit used for the constraint and the effect of the multiple scattering. We can estimate the error from the SVD intrinsic resolution and the momentum of track. In the case of the multi-constrained track, we can estimate the error from the result of the fit.

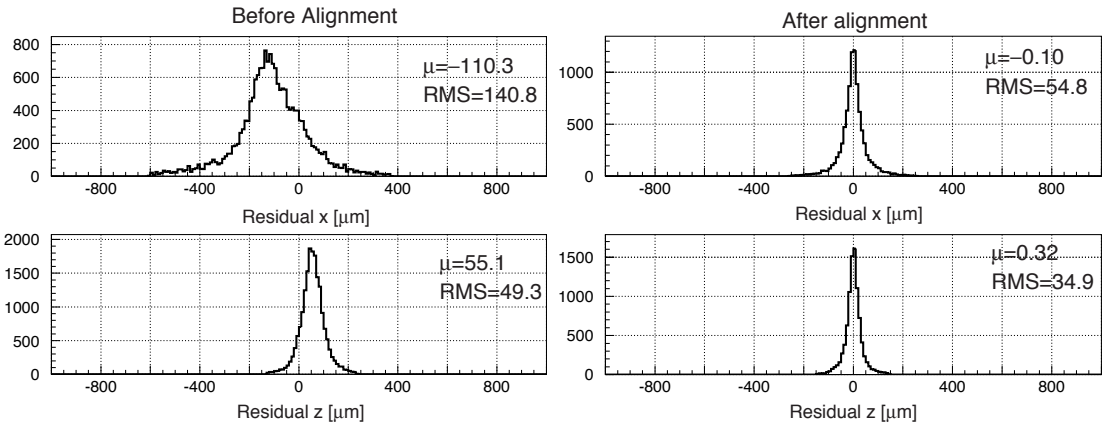

Figure A.4: Residual distributions of an inner layer DSSD before (left) and after (right) alignment with real data.

The error of SVD hit depends on the track incident angle. The nominal error evaluated from Monte Carlo simulation is stored in a file and used for reconstruction. We can use this data for hit error with known track incident angle with DSSD.

Figure A.4 shows the residual distributions of an inner layer DSSD before and after the alignment with real data. We can see the clear improvement in residual distributions.

For the internal alignment, we have tried several method (or strategy) to find the way to get the best alignment precision.[1] We describe here how we come up with some alignment methods and how they work or do not work. Since we have learned many things from experience with earlier experiments, we believe it is useful for the future similar activities to describe our experience in detail.

A.3.1 Original method

Our original alignment method is based on what was used in DELPHI experiment at LEP [68]. Basic idea of the *DELPHI method* is to divide the SVD into two half-shells. We first align each half-shell separately and align half-shells with each other in the end. The major profit of this method is less accumulation of error. If we align all DSSD's one by one, starting from one reference DSSD to neighboring,

[1]There are also other SVD alignment activities that are similar to, but independent of the procedure described here [66, 67]. All of them result in similar alignment precision. The alignment constants obtained by the procedure described here are used to analyze most of data used in the measurement described in this thesis.

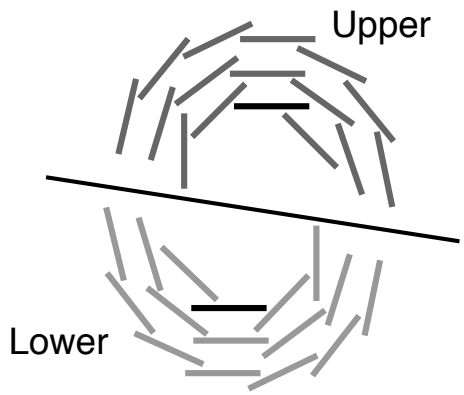

Figure A.5: The SVD divided into two half-shells

error should eventually accumulate and become large at the end. By aligning only half of SVD at the same time, accumulated error should be at most half of the case all DSSD's are aligned at once. Then, we can align two half-shells with each other, regarding each half-shells as well-aligned rigid body. In this half-shell alignment, we can use larger statistics because of large number of DSSD included.

A.3.1.1 Half-shell internal alignment

Among many way to divide SVD into two half-shells, we choose to divide it by approximately horizontal plane. (See Figure A.5) This is because we need tracks passing through both of half-shells for mutual alignment. We can use tracks in the overlap region, but we also need tracks perpendicular to the dividing plane to decide all the parameters precisely. In usual experiment with symmetric collider, we can use *e.g.* $e^+e^- \to \mu^+\mu^-$ events for such purpose, but we have only cosmic-ray tracks in Belle because of the asymmetric and finite-angle collision of KEKB.

Our basic strategy for half-shell internal alignment is as follows:

1. Select a reference DSSD for each half-shell. All other DSSD's in a half-shell are aligned with respect to the reference.

2. Align DSSD adjacent to the reference. The first DSSD is aligned using single constrained residual.

3. Then, align DSSD's adjacent to DSSD's already aligned. To make our procedure independent from the CDC systematics, we use double-constrained residual as much as possible.

4. Repeat this procedure until all DSSD's become aligned.

We select first layer DSSDs at the top and the bottom part as reference DSSD. We could select the other DSSD, like second or third layer DSSD, as the reference. However, we have found the first layer DSSD is the best candidate since it has more hits than others.

The alignment order of DSSD might affect the accuracy of alignment in this scheme. We have made our alignment program such that we can control the order of DSSD by an external control file. We have tried several ordering strategy. The general "guidelines" we find from our experience are:

- Constraint by the overlap region of DSSD is not so efficient in Belle SVD geometry. Due to small distance between layers, constraints from different layers have more statistical power.
- We have to use double-constraint method rather than single-constraint method as much as possible in order to reduce the effect from CDC.
- ~ 1000 hits are sufficient to align a DSSD.

According to these guidelines, we have developed half-shell internal alignment program. We have tested our program with the MC simulation and confirmed the capability of alignment within required precision.

A.3.1.2 Half-shell mutual alignment

By constraining tracks by one of half-shells and minimizing residual between those tracks and hits in the other half-shell, we can align half-shells with each other. For this purpose, we should use tracks penetrating both of half-shells. We have two types of such tracks, one is tracks going through overlap region of two half-shells and the other is cosmic-ray.

We have developed mutual alignment program using MINUIT [56] as a minimization package. Alignment procedure is as following:

1. Select tracks passing through both of half-shells.
2. Constrain tracks to SVD hits in the upper half-shell.
3. Minimize residuals in the lower half-shell by changing the relative alignment of the upper and the lower half shells.

4. Repeat the entire procedure several times (typically three times).

We have verified this program using MC data. Using 10,000 cosmic and 50,000 hadronic events, we can align half-shells with better than 10 μm and 1 mrad precision, if half-shell internal alignment was precise enough (σ of shift and rotation are within $\simeq$10 μm and 1 mrad level, respectively).

Then, we have tried to align half-shells using real data. However, the minimization have failed for many cases and result is not good even we can obtain it. Especially, resolution for tracks passing through the horizontal part is twice as bad as that for the vertical part. We investigated the reason by a detail study of the residual distribution for each DSSD. Finally, we find there is a large discrepancy of residual in the edge region of each half-shell. This indicates the accumulated error within (even) a half-shell are so large that half-shell mutual alignment does not work correctly. Half-shell mutual alignment works reasonably with MC data because half-shell internal alignment works better in that case due to the better tracking resolution in MC.

We may use this method when we get as good input as MC simulation, but we have to look for another way to improve alignment precision.

Table A.1 shows the cosmic ray helix mis-distance resolution obtained by the original method. The measurement procedure of the helix parameter resolution with cosmic data is the same as the measurement of impact parameter resolution, which is described in Appendix B.

Table A.1: Cosmic-ray helix parameter resolution with original method.

$d\rho(\mu m)$	ϕ_0 (mrad)	$\kappa(10^{-3})$	$dz(\mu m)$	$\tan\lambda(10^{-3})$
28.6	0.884	2.28	47.6	1.48

A.3.2 Improvement of the alignment method

Since we have found the half-shell method does not work well for the real data, we have finally abandoned the original *DELPHI* method and decided to search for a better alignment method.

A.3.2.1 Improvement of the alignment procedure

From a study of the residual distributions, we have concluded that our capability of aligning DSSD's in the close region is sufficient. Our weak point is the alignment between distant parts, like the upper and the lower parts. When we do not divide SVD into half-shells, we must consider how to align whole SVD without accumulating error. We use cosmic ray tracks with multi-constraint to constrain distant parts.

We also change the timing of updating alignment parameters after calculating constant. In the previous alignment procedure, we updated constants of every DSSD just after calculating them. In this scheme, error accumulated after aligning many DSSD's, and finally alignment parameter became inconsistent at the last DSSD to be aligned. We try to avoid such situation by updating parameters once for all DSSD's at the end of one procedure.

This new scheme has several advantages:

- The error will be distributed uniformly in all DSSD's, not accumulated in particular part.
- We no longer need to choose reference DSSD, nor worry about the order of alignment,
- We can use as many as possible constraint for all alignment steps.

We obtain helix parameter resolution shown in table A.2 with this method using cosmic ray events taken during collision. We used 100,000 cosmic ray events and 80,000 hadronic events. Comparing table A.2 with old result shown in table A.1,

Table A.2: Helix parameter resolution obtained with improved method

$d\rho(\mu m)$	ϕ_0 (mrad)	$\kappa(10^{-3})$	$dz(\mu m)$	$\tan\lambda(10^{-3})$
28.0	0.747	2.63	42.1	0.973

we can see an improvement of resolution with the new method, in spite of worse κ resolution, which is mostly determined by the CDC.

We also study the alignment precision using MC data. Using 50,000 cosmic ray and 50,000 continuum events, we have performed the alignment with the same procedure as with the real data. The helix parameter resolutions obtained by cosmic

ray MC events are shown in table A.3 for the cases where perfect alignment constants and constants obtained with alignment procedure are used. Distributions of difference of true and obtained alignment parameters for 102 DSSD's are shown in Fig. A.6.

Table A.3: Helix parameter resolution obtained with MC data.

	$d\rho(\mu m)$	ϕ_0 (mrad)	$\kappa(10^{-3})$	$dz(\mu m)$	$\tan\lambda(10^{-3})$
perfect	19.4	0.503	1.73	34.4	0.891
aligned	19.8	0.482	1.75	34.6	0.867

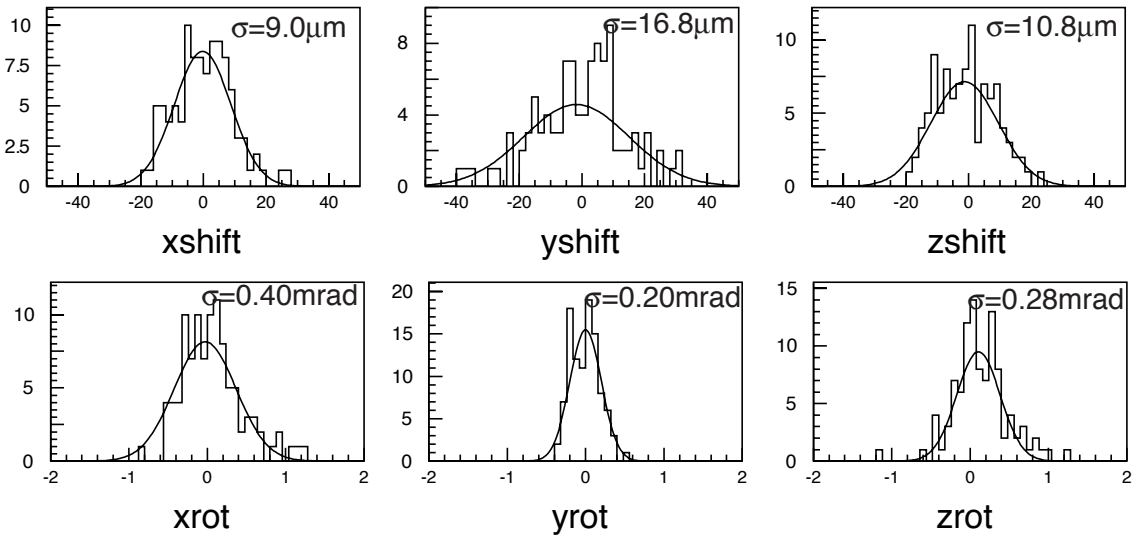

Figure A.6: Alignment parameter precision with MC

As seen from Table A.3 and Fig. A.6, we can align our SVD almost perfectly with the MC. This means our alignment program itself works very well, **if correct input is provided**. For the real data, our tracking quality has been continuously improving, but is not still as good as in ideal MC. Although our alignment program is designed as much independent from the CDC systematics as possible, it cannot be fully independent from the CDC as long as using tracks reconstructed by the CDC.

A.3.2.2 Alignment with no magnetic field data

We then try to be independent from the CDC by using cosmic ray tracks taken *without* a magnetic field. Since tracks are just straight lines without magnetic field, we can reconstruct tracks only with SVD. We developed alignment program which uses straight tracks reconstructed in the following way.

1. Extract information of tracks reconstructed by only the CDC and hits recorded by the SVD. Since our normal reconstruction program cannot reconstruct straight tracks correctly, we use special track fitter [69] for non-magnetic field data.

2. Select SVD hits which are close to the CDC track. Selection window is determined to be loose enough in order to avoid bias in the SVD hit selection.

3. Fit straight line with the selected SVD hits.

4. If the χ^2 of fit is too large, reject SVD hit which gives the worst χ^2. The threshold for the χ^2 is determined from the χ^2 distribution.

5. Re-fit track with remaining SVD hits. Repeat this procedure until the χ^2 becomes reasonable or number of SVD hits becomes less than three.

We took cosmic data without magnetic field several times. The first no magnetic field data for alignment purpose were taken in early January 2000, during the new year shutdown. We could reconstruct about 78,000 cosmic events passing through SVD with about one week of data taking.

When we use non-magnetic field data, we have to take into account the Lorentz effect. The Lorentz effect is caused by drift of carriers inside silicon by the magnetic field, which results in a shift of measured position orthogonal to the magnetic field. Usually, this shift is absorbed by the alignment parameters, since we are calculating alignment parameters using Lorentz-shifted data. When we calculate alignment parameters using data without magnetic field, we must add the effect of Lorentz shift to them. The Lorentz shift is parametrized by one global parameter added to $xshift$ of each DSSD. We have developed program to calculate the Lorentz shift by a fit to the data taken with a magnetic field using MINUIT. We obtain the Lorentz shift value as 6.7 μm.

Table A.4 shows the helix parameter resolution with the obtained constants. We used different data from what we used in the previous section to obtain this

resolution. MC results (perfect alignment case) with two different software library version are shown together. The main difference between Nov. 1999 and Feb. 2000 versions is the quality of tracking. Tracking group has introduced more "realistic" CDC resolution in newer versions, while SVD part has not been changed at all. From this result, we conclude our alignment precision is very close to the perfect case and sufficient for physics analysis use.

Table A.4: Helix parameter resolution for the final alignment parameter together with MC results. The difference between Nov. 1999 and Feb. 2000 versions is the tracking quality.

	$d\rho(\mu m)$	ϕ_0 (mrad)	$\kappa(10^{-3})$	$dz(\mu m)$	$\tan\lambda(10^{-3})$
data	18.3	0.471	1.81	40.7	0.909
MC(Nov. 1999)	16.4	0.431	1.28	30.7	0.781
MC(Feb. 2000)	17.8	0.507	1.66	36.1	0.932

We have been taking non-magnetic field data periodically and re-calculating the alignment constants. Especially, we replaced the SVD in summer 2000. After the replacement, we took cosmic data without magnetic field to align the new SVD. We could release the first alignment constants within one month after data taking was started. Our alignment precision with the new detector became the same level as the previous one after several iterations of alignment procedure and vertex resolution study.

A.4 CDC-SVD Alignment

In the CDC-SVD alignment, the SVD is considered as a rigid body and aligned with respect to the CDC. We have developed two methods for the CDC-SVD alignment. Considering the fact that the resolution of the CDC is a few hundred μm for r-ϕ direction and a few mm for z-direction, required precision of CDC-SVD alignment is $\simeq 10$ μm and $\simeq 100$ μm for r-ϕ and z-direction, respectively.

A.4.1 SVD-track method

The CDC-SVD alignment can be done by looking at the difference of tracks reconstructed by each detector. For the SVD, we can reconstruct track using position

of more than three SVD hits and curvature determined by the CDC. We call such track *SVD-track*. Similarly, tracks reconstructed using only CDC information are called *CDC-track*.

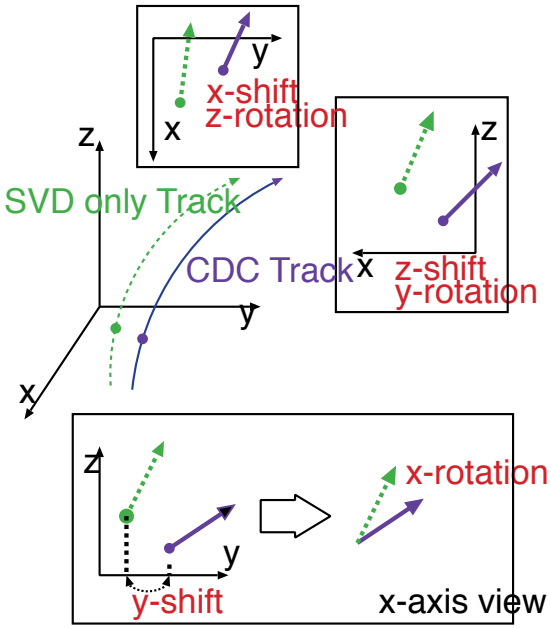

Figure A.7: CDC-SVD alignment with SVD-track method

By comparing the difference of position and direction between SVD-tracks and CDC-tracks near the origin, we can obtain the CDC-SVD alignment parameters. Figure A.7 shows a schematic view of this approach. When we project SVD- and CDC-tracks on some plane, we can obtain shift and rotation from the difference of two tracks. By averaging a few thousand events, we can get sufficient precision. We usually iterate this procedure for three times or more to ensure we obtain proper result.

This method was used in the early period and confirmed to provide sufficient alignment precision.

A.4.2 Track-hit residual method

Although the method described in the previous section provides us reasonable CDC-SVD alignment precision, it has a difficulty in the estimation of error. In order to solve this, we have developed another method for the CDC-SVD alignment based

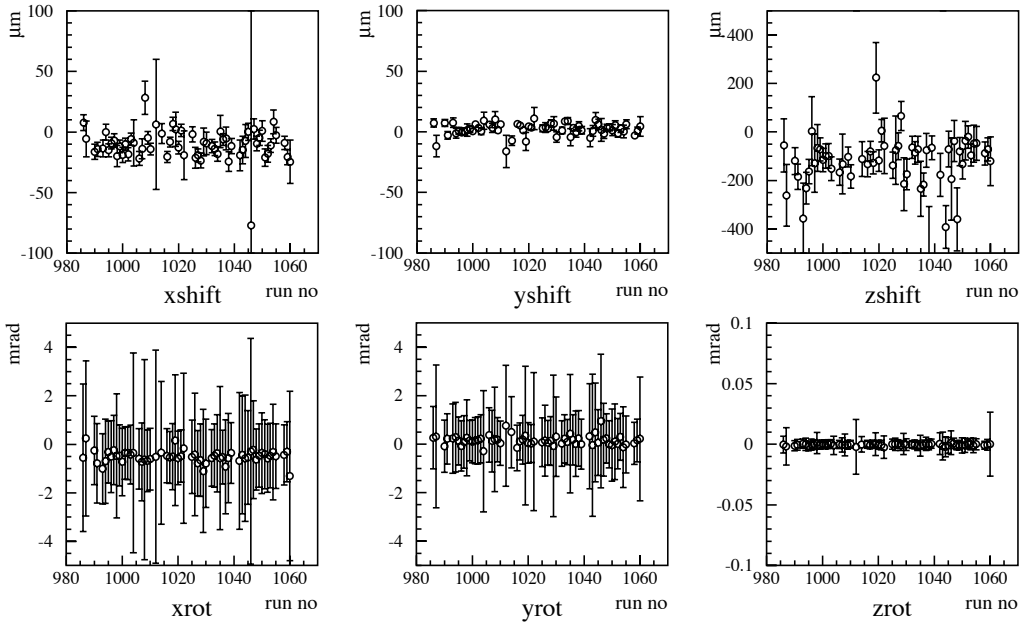

Figure A.8: Time variation of global alignment parameters.

on the concept of residual minimization.

Basic concept is the same as internal alignment. By minimizing

$$\chi^2 = \sum \frac{(\mathbf{x_{trk}} - \mathbf{x_{hit}})^2}{\sigma^2} \tag{A.11}$$

we can determine the CDC-SVD alignment constant.

We use the MINUIT minimizing package for minimization of Eq. A.11. This program works well and gives consistent result with the first method described in the previous section.

We have been checking the time dependence of CDC-SVD alignment constants. The vertex resolution and matching efficiency are continuously monitored along with data processing. In addition, we have checked possible time variation of relative position between SVD and CDC. Figure A.8 shows the time dependence of the CDC-SVD alignment parameters corresponding about hundred accelerator fills. We find no time variation from these plots.

Appendix B

Measurement of Impact Parameter Resolution

The impact parameter resolutions in both $r\phi$ and z planes are the most important quantities of SVD performance, because they are directly related to the vertex resolution. Thus its evaluation using real data is the essential part of the SVD performance study. Since the impact parameter resolution depends on the alignment precision, its evaluation is also important for the alignment study.

The impact parameter resolution can be estimated if the origins of the tracks in question are known. We use cosmic ray events taken during beam collisions to evaluate the impact parameter resolutions. One cosmic ray is reconstructed as an incoming and an outgoing track separated near the interaction point. Since the two reconstructed tracks belong actually to the same cosmic ray, the difference of the track positions near the interaction point represents the resolution. The resolution of helix parameter $d\rho$ and dz can be measured precisely from the $d\rho$ and dz residual distribution. The impact parameter resolution is calculated as the standard deviation of the difference of the two track positions divided by $\sqrt{2}$. A wide momentum range of cosmic rays can be used for study.

Once this is done, we can test the momentum dependence of the resolution which is approximately given by

$$\sigma = a \oplus b/\tilde{p}, \tag{B.1}$$

where a is determined by the SVD intrinsic resolution and the alignment accuracy, and b by the multiple Coulomb scattering. The quantity denoted by $\tilde{p}$ is called

"pseudo-momentum" and defined as follows:

$$\tilde{p} = \begin{cases} p\beta \sin^{\frac{3}{2}}\theta & (r\phi) \\ p\beta \sin^{\frac{5}{2}}\theta & (z) \end{cases} \tag{B.2}$$

Figure B.1 shows the residual distributions obtained from the distance between two reconstructed tracks in the $r\phi$ plane defined as $(d\rho_+ - d\rho_-)/\sqrt{2}$ (top-left) and in the rz plane defined as $(dz_+ - dz_-)/\sqrt{2}$ (bottom-left), where $d\rho_{+(-)}$ and $dz_{+(-)}$ are $d\rho$ and dz for the track with positive (negative) charge. Tracks with large momentum ($\tilde{p} > 5$ GeV/c) are selected to avoid the effect of multiple Coulomb scattering. The sigma of the fit to the distribution with the Gaussian function gives the impact parameter resolution. As the SVD alignment relies on cosmic rays, one needs to check if the difference in resolution between "vertical" tracks ($|\sin\phi| \sim 0$) and "horizontal" tracks ($|\sin\phi| \sim 1$) is seen. As shown in the figure (middle and right), we do not see sizable differences between two cases, indicating the goodness of the alignment.

Table B.1: $\sigma_{d\rho}$ and σ_{dz} in various pseudo momentum regions.

$\tilde{p}$(GeV/c)	$\sigma_{d\rho}(\mu$m)	$\sigma_{dz}(\mu$m)
0.5-0.75	68.3 ± 2.2	76.8 ± 1.3
0.75-1.0	55.3 ± 2.0	58.6 ± 1.3
1.0-1.5	43.4 ± 1.0	49.3 ± 1.3
1.5-2.0	34.7 ± 0.9	45.3 ± 1.0
2.0-2.5	30.6 ± 0.9	39.7 ± 1.2
2.5-3.0	25.1 ± 0.3	40.2 ± 1.0
3.0-4.0	24.5 ± 0.6	36.9 ± 1.5
4.0-5.0	20.9 ± 0.5	36.2 ± 1.1
5.0-6.0	19.9 ± 0.1	35.3 ± 1.3
6.0-7.0	17.0 ± 1.0	37.0 ± 1.4
7.0-8.0	18.7 ± 0.5	33.1 ± 1.4
8.0-10.0	14.9 ± 0.1	34.2 ± 1.3

Table B.1 summarizes the impact parameter resolutions $\sigma_{d\rho}$ and σ_{dz} in various pseudo momentum regions down to 0.5 GeV/c. These results are combined with dimuon and two-photon events ($\gamma\gamma \to \pi^+\pi^-\pi^+\pi^-$) [37]. The impact parameter

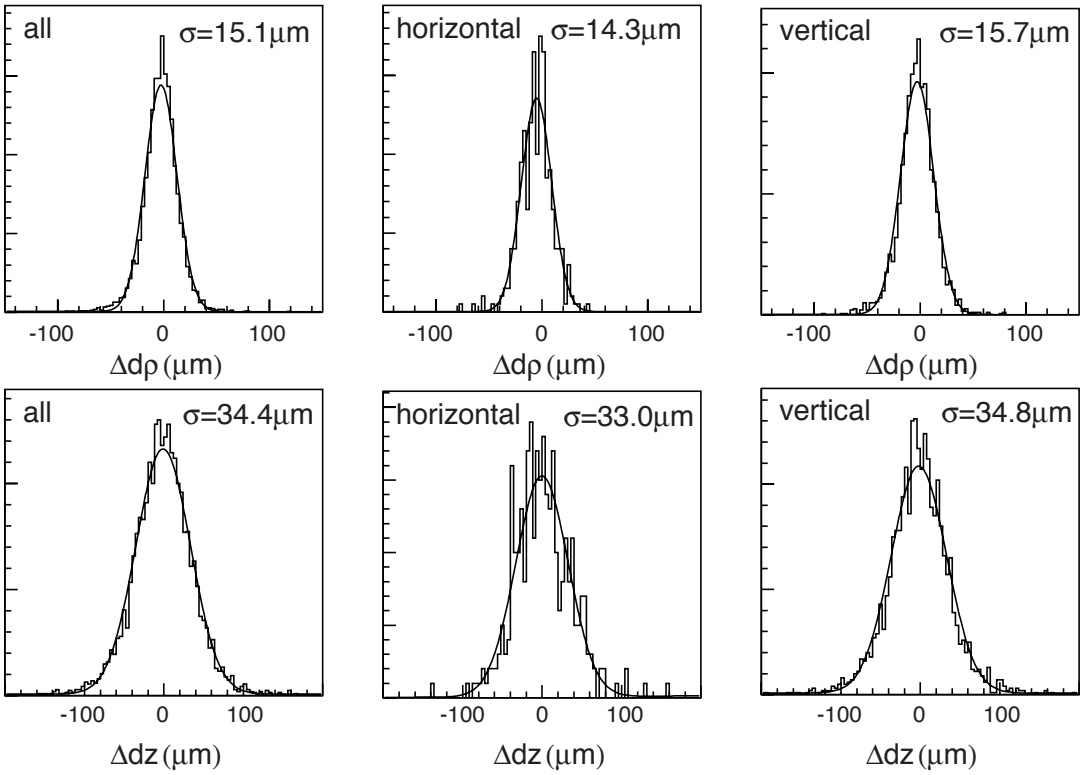

Figure B.1: Residual distributions in the $r\phi$ (top) and rz (bottom) plane. Tracks with large momentum ($\tilde{p} > 5$ GeV/c) are selected.

resolutions, $\sigma_{r\phi}$ and σ_z, for the $r\phi$ and the z directions, respectively, are shown in Fig. (3.10). The momentum and angular dependence as well are well represented by the following formulae:

$$\sigma_{r\phi} = 19 \oplus 50/(p\beta \sin^{3/2}\theta)\mu\text{m}, \tag{B.3}$$

$$\sigma_z = 36 \oplus 42/(p\beta \sin^{5/2}\theta)\mu\text{m}. \tag{B.4}$$

Appendix C

Measurement of Wrong Tagging Fractions

The wrong tagging fractions for each $r \equiv |Q|$ region, w_l, are determined directly from the data for the six r intervals using exclusively reconstructed, self-tagged $B^0 \to D^{*-}\ell^+\nu$, $D^-\pi^+$, $D^{*-}\pi^+$, and $D^{*-}\rho^+$ decays. These samples are also used as control samples in the validation of result.

C.1 Reconstruction of B Decays with Specific Flavor

For the semi-leptonic $B^0 \to D^{*-}\ell^+\nu$ decay mode, we use the decay chains $D^{*-} \to \overline{D}^0\pi^-$, where $\overline{D}^0 \to K^+\pi^-$, $K^+\pi^-\pi^0$, and $K^+\pi^+\pi^-\pi^-$. Candidate D^0 decays are selected requiring the invariant mass within a certain range from the nominal D^0 mass, where the actual range depends on the D^0 decay modes. The requirement on the positive side varies from +9 to +23 MeV/c^2 and the negative side varies from −9 to −37 MeV/c^2. For the reconstruction of $D^{*\pm}$, we impose a requirement on the mass difference, $\Delta M = M_{K\pi\pi_s} - M_{K\pi}$, where π_s denotes the low-momentum $\pi^\pm$ from $D^{*\pm}$ (slow-pion). By taking the difference of invariant mass, we can utilize the very small phase space of $D^{*\pm}$ decay to reduce background. The mass difference make very narrow peak for the signal, canceling the measurement error of the invariant mass of $K\pi$. The actual range of requirement varies from 0.8 to 1.75 MeV/c^2 depending on the D^0 decay modes. $D^{*\pm}$ candidates are combined with e or μ candidates which have opposite charge to form $D^{*-}\ell^+\nu$ candidates, with requirement on missing mass to select events consistent with existence of ν.

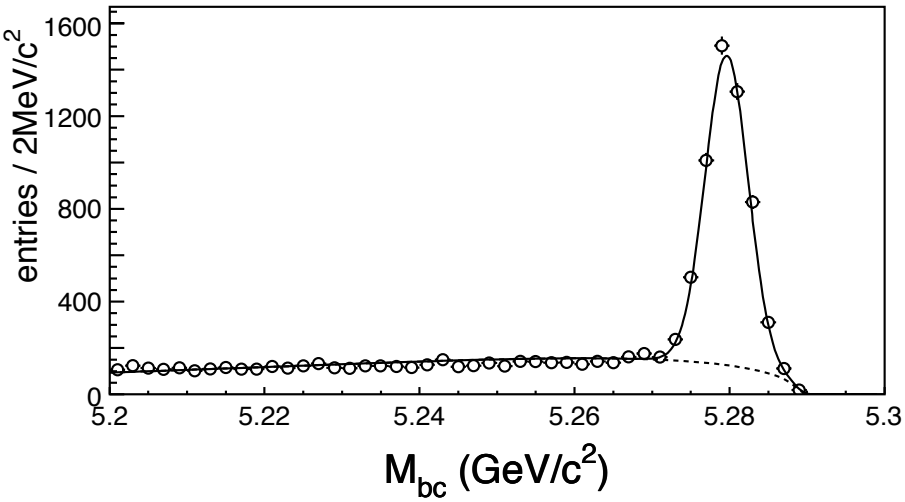

Figure C.1: $M_{\rm bc}$ distribution for all hadronic control samples ($B^0 \to D^-\pi^+$, $D^{*-}\pi^+$, $D^{*-}\rho^+$) with $|\Delta E| < 50$ MeV. Fit curves for background (dashed) and signal puls background (solid) are also shown.

For hadronic decays, we use decay modes $B^0 \to D^-\pi^+$, $D^{*-}\pi^+$, and $D^{*-}\rho^+$. D^{*-} is reconstructed in the same decay chain as $D^{*-}\ell^+\nu$ mode. D^- is reconstructed in $D^- \to K^+\pi^-\pi^-$. For D and D^{*-}, we impose decay-mode dependent requirement on the reconstructed D mass (from ± 30 to ± 60 MeV/c^2) and D^*-D mass difference (from ± 3 to ± 12 MeV/c^2), in similar way as semi-leptonic mode. We select ρ^+ by requiring an invariant mass of π^+ and π^0 to be within 150 MeV/c^2 from nominal ρ^+ mass. In order to suppress continuum background, we impose mode dependent requirements in R_2 (upper limit values 0.5–1.0) and the thrust angle $\cos\theta_{thr}$ (upper limit values 0.92–1.0). [1] We select B^0 by requiring $5.27 < M_{\rm bc} < 5.29$ GeV/c^2 and $|\Delta E| < 50$ MeV. Figure C.1 shows the $M_{\rm bc}$ distribution for three modes combined.

The number of reconstructed events and the signal fraction for each mode are summarized in Table C.1.

[1] The thrust axis is defied as the axis which maximizes the sum of momentum of each particle projected onto it. The thrust angle used here is defined as the angle between two axes defined by particles used and not-used for the B meson reconstruction.

Table C.1: Number of events and signal fraction for modes used in the wrong-tag fraction measurement.

	mode	yield	purity(%)
semi-leptonic	$D^{*-}\ell^+\nu$	16101	79.4
hadronic	$D^-\pi^+$	2241	85.5
	$D^{*-}\pi^+$	2126	87.1
	$D^{*-}\rho^+$	1620	72.2

C.2 Measurement of Wrong-Tagging Fractions

We evaluate the wrong tagging fractions by reconstructing flavor-specific decays on one side, and tagging the b-flavor for the other side with the flavor-tagging algorithm described in Section 5.1.

Since we know the flavors of both of B mesons in this case, we can observe the time evolution of the neutral B meson pair with the opposite flavor (OF) and the same flavor (SF), which originates from B^0-$\overline{B}^0$ mixing. The observed OF-SF asymmetry is expressed as:

$$A_{mix} = \frac{N_{OF} - N_{SF}}{N_{OF} + N_{SF}} = (1 - 2w_l)\cos(\Delta m_d \Delta t), \tag{C.1}$$

where Δm_d is the mass-difference of the two B meson mass eigenstates and w_l is the wrong-tagging fraction. Thus we can obtain the w_l by measuring the amplitude of the OF-SF asymmetry.

The vertex position of the reconstructed B meson is obtained as follows. First, we obtain D vertex by vertex fit using charged K and π tracks and calculate D momentum. Then, we obtain B vertex using the pseudo-D track and remaining charged tracks (ℓ or π). The vertex position of tagging-side B meson is obtained using the same algorithm as CP eigenstate candidates described in Section 5.2.

We fit the Δz distribution of the SF and OF events to obtain the wrong-tag fraction, fixing Δm_d to the world average value of 0.472 ps^{-1} [18]. The PDF used for this measurement is similar to that used for the measurement of $\sin 2\phi_1$. The parameters of resolution function and background shape are determined separately for modes used in this measurement.

We first obtain results with semi-leptonic and hadronic modes separately, since the background component of these two categories are totally different and therefore

Table C.2: Summary of measured wrong-tag fractions (w_l) and event fractions (f_l) for each r interval. The errors include both statistical and systematic uncertainties.

l	r	f_l	w_l(semi-leptonic)	w_l(hadronic)	w_l(combined)
1	$0.000 - 0.250$	0.405	$0.463^{+0.011}_{-0.011}$	$0.469^{+0.015}_{-0.016}$	$0.465^{+0.010}_{-0.009}$
2	$0.250 - 0.500$	0.149	$0.351^{+0.019}_{-0.017}$	$0.352^{+0.026}_{-0.026}$	$0.352^{+0.015}_{-0.014}$
3	$0.500 - 0.625$	0.081	$0.254^{+0.021}_{-0.020}$	$0.219^{+0.031}_{-0.030}$	$0.243^{+0.021}_{-0.030}$
4	$0.625 - 0.750$	0.099	$0.169^{+0.019}_{-0.018}$	$0.192^{+0.028}_{-0.027}$	$0.176^{+0.022}_{-0.017}$
5	$0.750 - 0.875$	0.123	$0.107^{+0.015}_{-0.015}$	$0.127^{+0.032}_{-0.031}$	$0.110^{+0.022}_{-0.014}$
6	$0.875 - 1.000$	0.140	$0.041^{+0.012}_{-0.011}$	$0.041^{+0.024}_{-0.023}$	$0.041^{+0.011}_{-0.010}$

treatment of background is different as well. Then, we combine results from these two categories by taking weighted average, taking into account the correlated systematic errors correctly. We conservatively treat the difference between the weighted average and individual measurement as a systematic error as well, and add them in quadrature.

Figure C.2 shows the measured asymmetries as a function of Δt for hadronic modes together with fit curves for the six r intervals. The measured wrong-tag fractions (w_l) and event fractions (f_l) for each r interval ($l = 1, 6$) are summarized in Table C.2. The overall efficiency and effective efficiency are estimated to be 99.7% and $(27.0 \pm 0.8(\text{stat})^{+0.6}_{-0.9}(\text{syst}))\%$, respectively.

We check for a possible bias in flavor tagging by measuring the effective tagging efficiency for B^0 and $\overline{B}^0$ self-tagged samples separately, and $q = -1$ and $q = +1$ samples separately. We find no statistically significant difference.

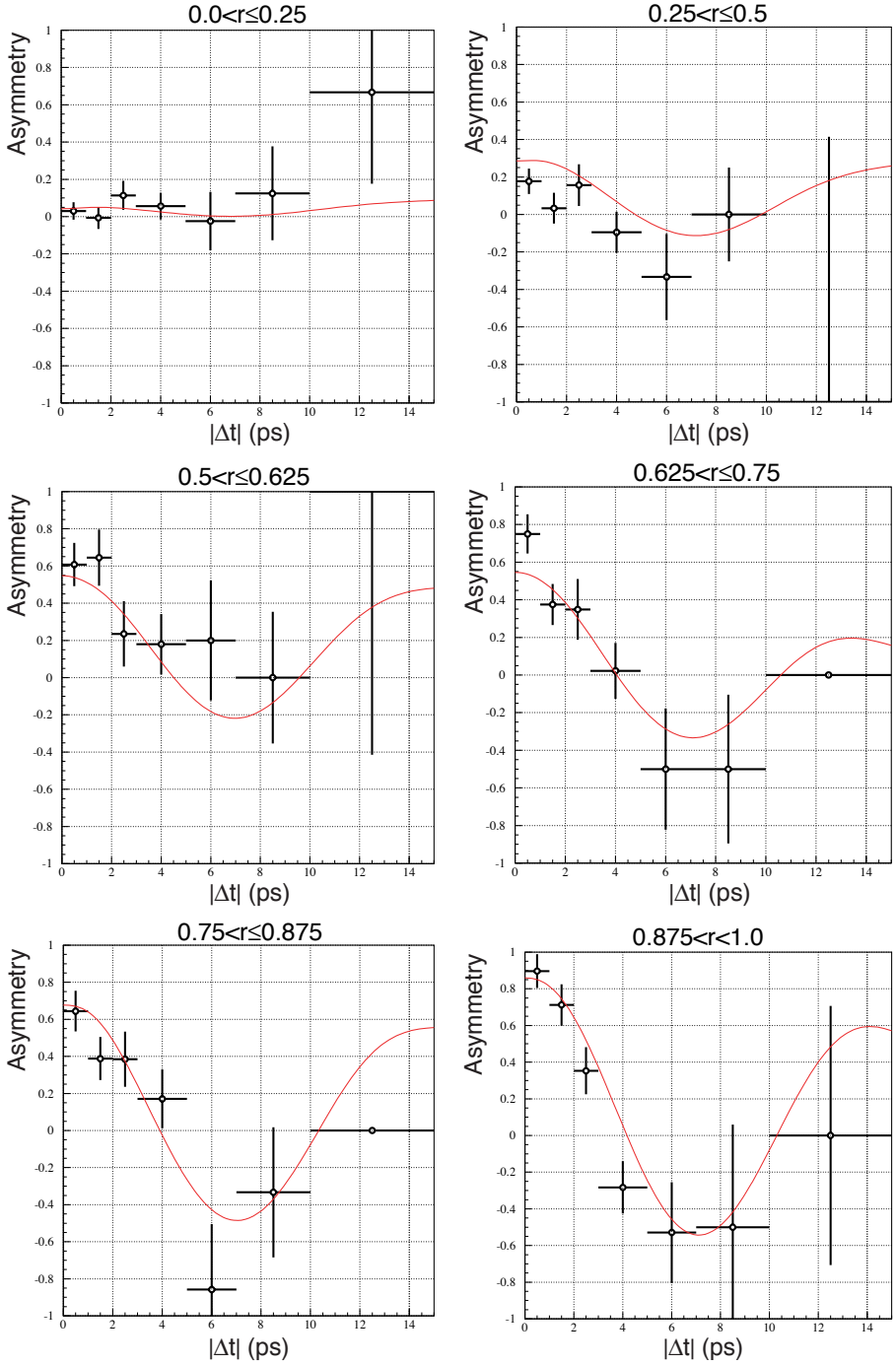

Figure C.2: Asymmetries as a function of Δt for hadronic modes together with fit curves for the six r intervals.

Appendix D

Statistical Significance of the Result

In this appendix, we discuss the significance of $\sin 2\phi_1$ measurement based on the method proposed by Feldman and Cousins [60]. We first explain the method of Feldman-Cousins in general, and then apply it to our result.

The *confidence intervals* are intervals constructed so as to include the true value of the parameter with a probability greater than or equal to a specified level.

Consider the case that the true value of the parameter α is fixed but unknown. The properties of our experimental apparatus are expressed in a function $f(x;\alpha)$ which gives the probability of observing data x if the true value of the parameter is α. $f(x;\alpha)$ has to be normalized such that $\int_{-\infty}^{\infty} f(x;\alpha)dx = 1$. Using $f(x;\alpha)$, we can find for every value of α, a set of values $x_1(\alpha,\varepsilon)$ and $x_2(\alpha,\varepsilon)$ such that

$$P(x_1 < x < x_2;\alpha) = 1-\varepsilon = \int_{x_1}^{x_2} f(x;\alpha)dx, \tag{D.1}$$

This is illustrated in Fig. D.1: a horizontal line segment $[x_1(\alpha,\varepsilon), x_2(\alpha,\varepsilon)]$ is drawn for representative values of α. The union of such intervals for all values of α, designated as $D(\varepsilon)$, is known as the *confidence belt*.

Upon performing an experiment to measure x and obtaining a value x_0, one draws a vertical line through x_0. The confidence interval for α is the set of all values of α for which the corresponding line segment $[x_1(\alpha,\varepsilon), x_2(\alpha,\varepsilon)]$ is intercepted by this vertical line. Such confidence intervals are said to have a *confidence level* (CL) equal to $(1-\varepsilon)$.

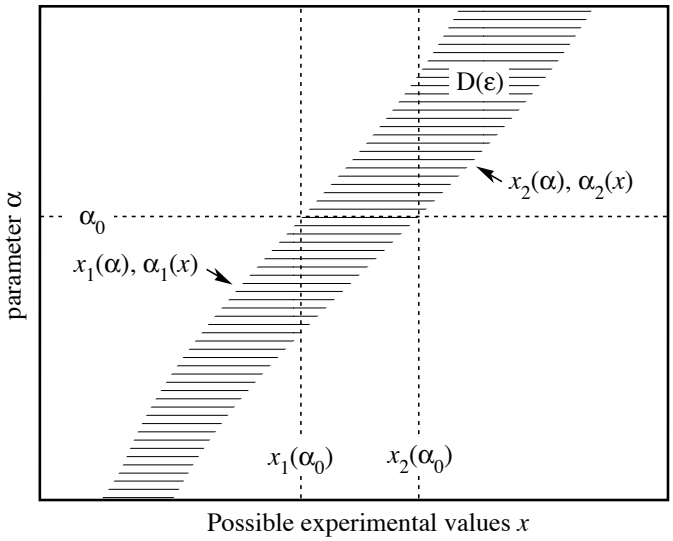

Figure D.1: Construction of the confidence belt.

Now suppose that the true value of α is α_0, indicated in the figure. We see from the figure that α_0 lies between $\alpha_1(x)$ and $\alpha_2(x)$ if and only if x lies between $x_1(\alpha_0)$ and $x_2(\alpha_0)$. The two events thus have the same probability, and since this is true for any value α_0, we can drop the subscript 0 and obtain

$$1 - \varepsilon = P(x_1(\alpha) < x < x_2(\alpha)) = P(\alpha_2(x) < \alpha < \alpha_1(x)). \tag{D.2}$$

The condition of coverage Eq. (D.1) does not determine x_1 and x_2 uniquely and additional criteria are needed. The most common criterion is to choose *central intervals* such that the probabilities excluded below x_1 and above x_2 are each $\varepsilon/2$. In other cases one may want to report only an upper or lower limit, in which case the probability excluded below x_1 or above x_2 can be set to zero. These criteria are sufficient in most cases, but there is a more general *ordering principle* which produces confidence intervals with better properties when in the neighborhood of a physical boundaries.

One of such ordering principles, which was suggested by Feldman and Cousins [60], is that based on *likelihood ratio ordering*. We define the likelihood ratio

$$\lambda(x;\alpha) = \frac{f(x;\alpha)}{f(x;\hat{\alpha})}, \tag{D.3}$$

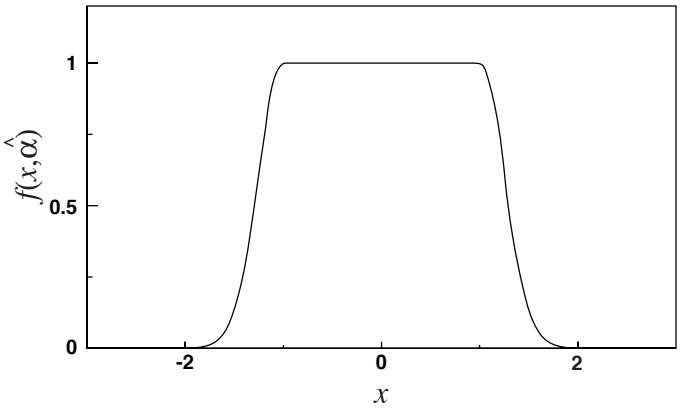

Figure D.2: Distribution of the function $f(x;\hat{\alpha})$.

where $\hat{\alpha}$ is the value of the parameter which, out of allowed values, maximizes $f(x;\alpha)$. Then we determine x_1 and x_2 such that

$$\lambda(x_1;\alpha) = \lambda(x_2;\alpha). \tag{D.4}$$

The Feldman-Cousins construction of confidence intervals ensures that a vertical line drawn in the x-α plane for any measurement x_0 will intercept the confidence belt to determine a well behaved confidence interval with correct coverage. This means that intervals $[x_1, x_2]$ at the confidence level $(1-\varepsilon)$ really would contain the unknown true value of α in a fraction of $(1-\varepsilon)$ if we repeat a set of the same experiments. The Feldman-Cousins method has another nice feature. If we make decision to report an upper (or lower) limit or two-sided interval by looking at the data, then it introduces bias and the resulting intervals will not in general cover the parameter with the probability $(1-\varepsilon)$. With the Feldman-Cousins method, we can avoid such biasing effect because it *unifies* the one- and two-sided intervals in a way which preserves the coverage probability.

We calculate 95% confidence interval using the Feldman-Cousins method for our measurement of $\sin 2\phi_1$ in $B^0 \to J/\psi K_L$ mode. We assume a symmetric Gaussian uncertainty with σ conservatively taken to be the larger one of the asymmetric errors. In this case, $f(x;\alpha)$ is a Gaussian function with the mean being α and the sigma being the error of the measurement. $f(x;\hat{\alpha})$ is a constant, corresponding to $\hat{\alpha} = x$, for x within the physical region. For $|x| > 1$, $f(x;(\hat{\alpha}))$ falls according to the Gaussian because $\hat{\alpha}$ is constrained by the physical boundary ± 1. Figure D.2 shows the function $f(x;\hat{\alpha})$. Figure D.3 shows the function $\lambda(x;\alpha) = f(x;\alpha)/f(x;\hat{\alpha})$ for

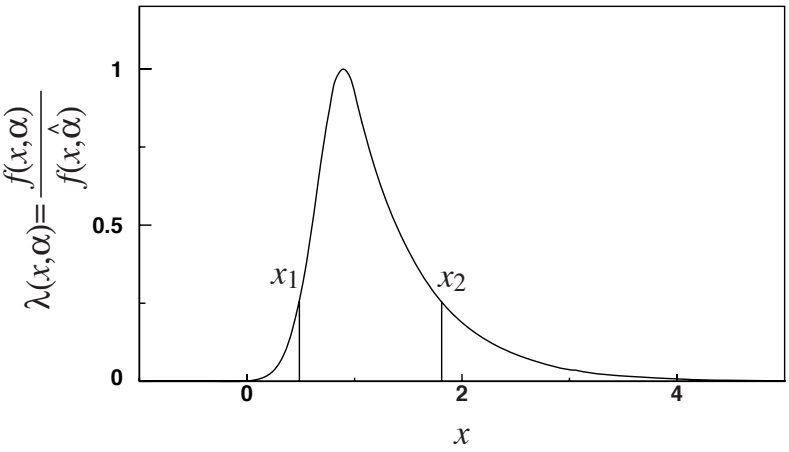

Figure D.3: Likelihood ratio $\lambda(x;\alpha) = f(x;\alpha)/f(x;\hat{\alpha})$ for the case $\alpha = 0.9$. x_1 and x_2, which are determined such that $\lambda(x_1;\alpha) = \lambda(x_2;\alpha)$ and $\int_{x_1}^{x_2} f(x;\alpha)dx = 0.95$, are also shown.

the case $\alpha = 0.9$. The x_1 and x_2, which are determined such that $\lambda(x_1;\alpha) = \lambda(x_2;\alpha)$ and $\int_{x_1}^{x_2} f(x;\alpha)dx = 0.95$, are also shown in the figure.

Figure D.4 shows the 95% Feldman-Cousins confidence belt for our measurement. The vertical line indicates the value we obtain from fit, $\sin 2\phi_1 = 1.31$. The intersections of this line with the confidence belts determine the confidence interval, from which we find the lower bound of $\sin 2\phi_1$ to be 0.76 at 95% confidence level.

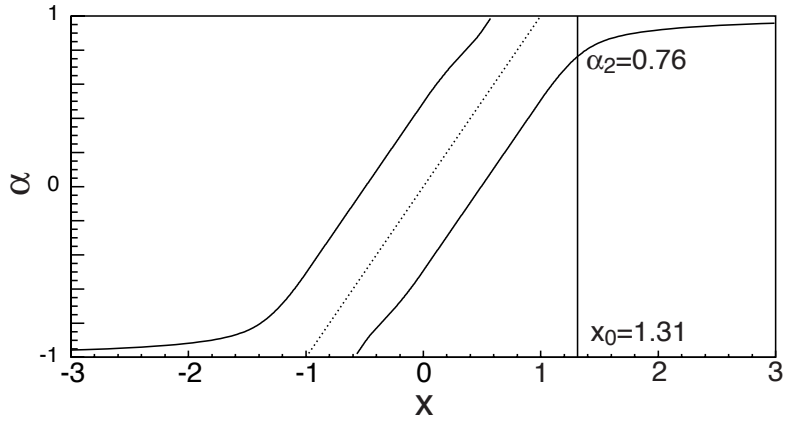

Figure D.4: 95% confidence interval constructed by Feldman-Cousins method [60].

Appendix E

Development of Radiation Hard Preamplifier Chip for Belle SVD

The radiation hardness of readout electronics is the most critical limitation for a long-term operation of the silicon vertex detector (SVD). The VA1 CMOS VLSI chip, a low noise amplifier with a shaper designed for readout of silicon strip detectors, is used to readout the Belle SVD. We have measured the radiation hardness of VA1 chip using γ ray from ^{60}Co and found that the VA1 chip is radiation tolerant up to about 200 krad. Operation experience has shown that the radiation dose in the SVD is about 10 krad per month. We need to improve the radiation hardness of the readout chip to 2 Mrad or more for 5 years of stable operation. We have been working on the design of a new detector which addresses the weak points of the current SVD and serves as the vertex detector for several years. Since the performance of the VA1 chips has been satisfactory in the current SVD, we plan to continue with the same basic VA1 architecture in the upgraded SVD [70] (so-called SVD 2).

In this appendix, we describe the result of radiation hardness measurements we have performed in our R&D activity. We start from a brief note of radiation damage to a MOS FET, and then present results of measurements.

E.1 Radiation Effects on MOS Devices

Radiation damage to a MOS FET is mainly caused by ionization charges trapped near or in the interface between the SiO_2 and the Si bulk [71]. Increased oxide trapping (fixed charge in the oxide layer) and interface trapping (localized electronic state very near the Si-SiO_2 interface) cause increased $1/f$ noise, mobility degradation

Table E.1: Thickness of gate oxide for each process used for the samples.

feature size (μm)	1.2	0.8	0.35
oxide thickness (nm)	25	16	7.5

(which then degrades transconductance and increases white noise) and threshold voltage shifts.

It is well known that the intrinsic radiation hardness of CMOS ICs depends on the gate-oxide thickness, since a thinner oxide provides fewer ionization charges. In addition, trapped charges can escape via a tunneling effect. The threshold voltage shift in a MOS transistor will scale like the square of the gate oxide thickness (t_{ox}), which in turn is (approximately) proportional to the feature size of the process being used. Furthermore, it is known that the radiation hardness scales much better than t_{ox}^{-2} in the deep sub-micron region [72].

Our basic strategy is to improve the intrinsic radiation hardness of the VA1 by taking advantage of the thinner gate-oxide layers used in sub-micron CMOS processes. The standard VA1 chip is fabricated in the AMS 1.2 μm process. We implemented the VA1 in 0.8 μm and 0.35 μm processes. The design work for these tests was done by the IDE AS company in Oslo Norway, who sell the standard VA1 [73]. Table E.1 shows the oxide thickness of a MOS FET in each process technology.

E.2 Measurement of Radiation Effects on VA1 Chip

VA1 samples are irradiated by a ^{60}Co gamma-ray source in the Research Center for Nuclear Science and Technology at the University of Tokyo. The average dose rate is approximately 1 krad/min. In order to simulate a running environment, proper DC bias currents and voltages are supplied to the VA1 chips during the irradiation. The samples are annealed for a few days after each irradiation. Bias voltages which controls feed back resisters are adjusted to minimize equivalent noise charge (ENC) and to keep constant peaking time before the measurement. The bias currents for the input transistors of the preamplifier and shaper are kept constant.

Figure E.1 (a) and (b) show measured noise and gain of the VA1 fabricated in

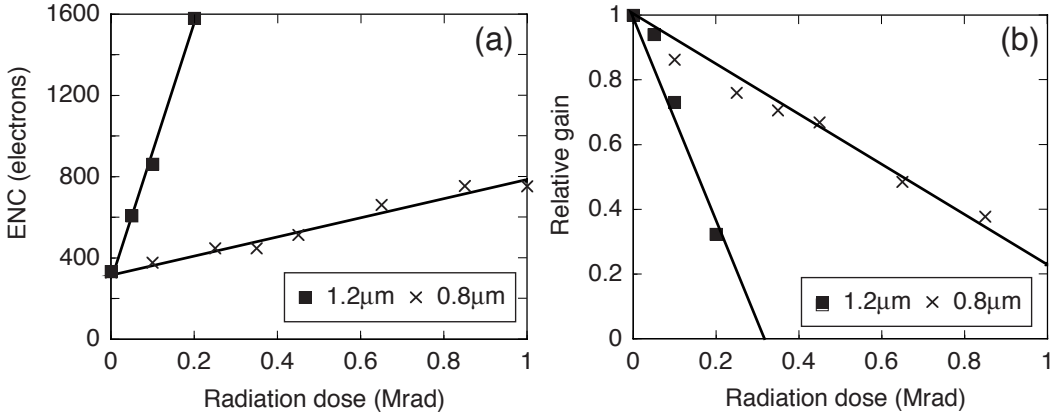

Figure E.1: Measured noise (a) and gain (b) of the 1.2 μm and 0.8 μm VA1 at a peaking time of 1.7 μs with 20 pF of load capacitance as a function of the total radiation dose.

the 1.2 μm and 0.8 μm process at a peaking time of 1.7 μs with 20 pF of load capacitance as a function of the total radiation dose. The fractional noise increase of the standard VA1 is measured to be about 0.3%/krad. After irradiation of over 200 krad, standard VA1 shows no gain although the digital part is still functioning. The improvement in radiation hardness for the 0.8 μm technology is quite significant. The fractional noise increase of the 0.8 μm VA1 is measured to be 0.08%/krad. The 0.8 μm VA1 shows no gain at 1.2 Mrad.

Figure E.2 (a) and (b) show the measured noise and gain of the VA1 fabricated in the 0.35 μm process at a peaking time of 0.5 μs with 20 pF of load capacitance as a function of the total radiation dose. Note that the initial noise is larger for the 0.35 μm VA1 than the other VA1s due to the different shaping times. Taking into account the $1/\sqrt{\tau}$ scaling, the noise performance of the 0.35 μm VA1 is equivalent to the original VA1. After a small degradation of the noise and gain, no further significant degradation is observed. The sample is still functional after 20 Mrad of total radiation dose.

Figure E.3 shows the fractional noise increase and the fractional gain degradation as a function of the oxide thickness (t_{ox}). In these plots, the fractional noise increase and the fractional gain degradation for the 0.35 μm VA1 are obtained assuming the linear dependence of the noise and the gain on the radiation dose. Lines with $y = at_{ox}^6$ and $y = at_{ox}^3$ are also shown as reference for the fractional noise increase and the fractional gain degradation, respectively. We find that the radiation hardness

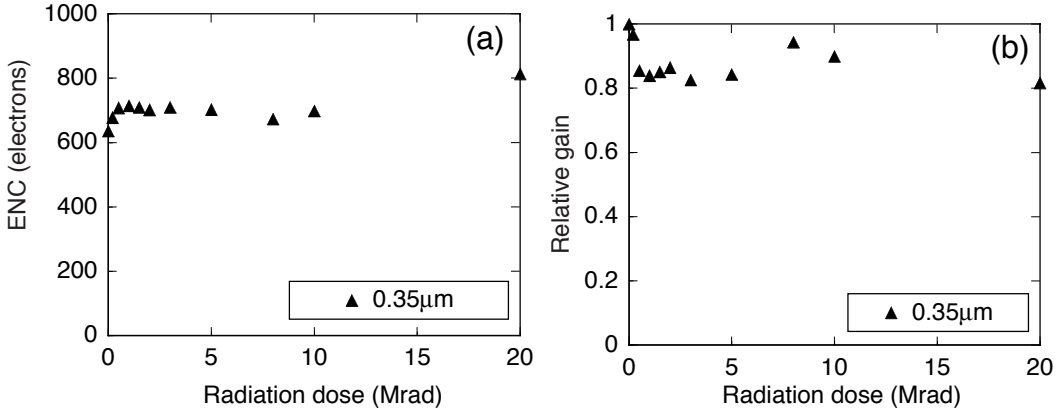

Figure E.2: Measured noise (a) and gain (b) of the VA1 chip fabricated in the 0.35 μm process at a peaking time of 0.5 μs with 20 pF of load capacitance as a function of the total radiation dose.

scales as t_{ox}^{-6} in the case of the noise and t_{ox}^{-3} in the case of the gain, for oxide thicknesses from 1.2 μm to 0.8 μm. The difference may indicate the difference in the damage mechanism. The change in noise and gain from 0.8 μm to 0.35 μm appears to be larger than a sixth power dependence on the thickness of the oxide. Escape of hole traps by the tunneling effect must be dominant for sub-micron technologies. Furthermore, almost all traps escape by tunneling for the device fabricated in the 0.35 μm process and the total dose effect is negligible.

E.3 Expected Performance of SVD After Irradiation

Figure E.4 shows the expected S/N ratio for VA1 fabricated with 1.2 μm, 0.8 μm, and 0.35 μm connected with two DSSDs, as a function of radiation dose. For the case of VA1 with 0.8 μm and 0.35 μm, the improvement of DSSD [74] is also taken into account. Thanks to the improvement of the radiation hardness of VA1, we can expect S/N ratio greater than 10 even after irradiation of 1 Mrad for VA1 with 0.8 μm. In the case of VA1 with 0.35 μm process, the noise increase of VA1 is negligible and the dominant source of noise increase after irradiation is the shot noise from DSSD leakage current.

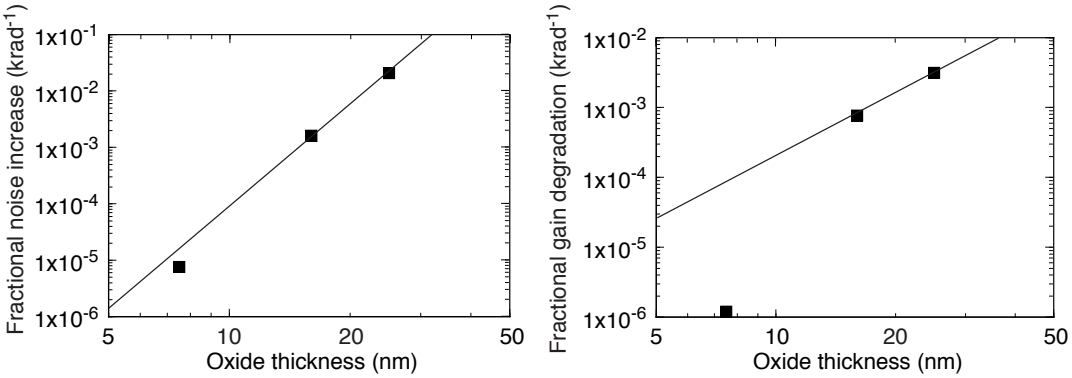

Figure E.3: Fractional noise increase (left) and fractional gain degradation (right) as a function of the oxide thickness (t_{ox}). Lines with $y = at_{ox}^6$ $y = at_{ox}^3$ are also shown as a reference.

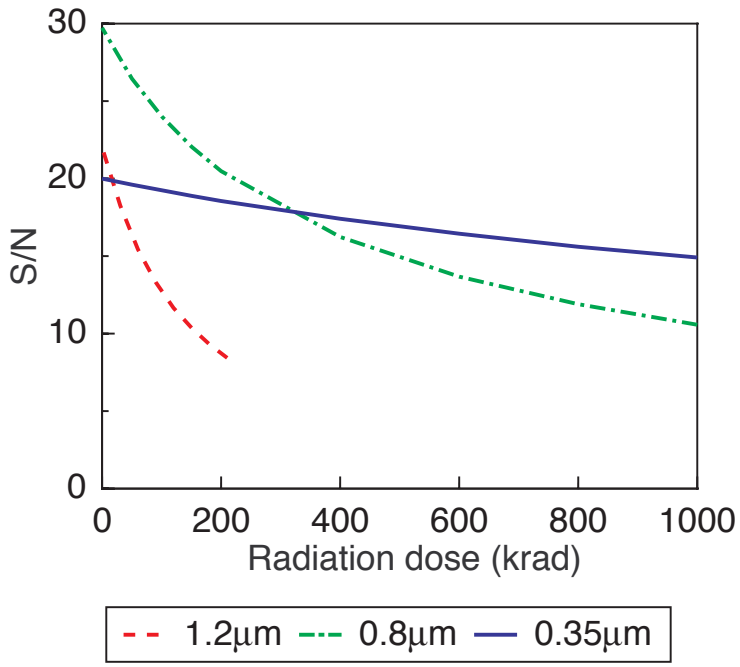

Figure E.4: Expected S/N for VA1 fabricated with 1.2 μm (dashed), 0.8 μm (dot-dashed), and 0.35 μm (solid) connected with two DSSDs as a function of radiation dose. For the case of VA1 with 0.8 μm and 0.35 μm, the improvement of DSSD [74] is also taken into account.

Bibliography

[1] A. D. Sakharov, Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967).

[2] J. H. Christenson, Cronin, Fitch and Turlay, Phys. Rev. **13** 138 (1964).

[3] M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **49** 652 (1973).

[4] A. B. Carter and A. I. Sanda, Phys. Rev. D **23** 1567 (1981); I. I. Bigi and A. I. Sanda, Nucl. Phys. B **193** 85 (1981).

[5] E. Fernandez *et al.* (MAC Collaboration), Phys. Rev. Lett. **51** 1022 (1983); N. Lockyer *et al.* (Mark II Collaboration), Phys. Rev. Lett. **51** 1316 (1983).

[6] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **192** 245 (1987).

[7] I. I. Bigi and A. I. Sanda, *CP violation*, Cambridge university press, Cambridge (2000).

[8] E. Noether, *Nach. v. d. Ges. d. Wiss. zu Göttingen*, 253-257 (1918).

[9] T. D. Lee and C. N. Yang, Phys. Rev. Lett. **98** 1501 (1955); Phys. Rev. Lett. **104** 254 (1956).

[10] C. S. Wu *et al.*, Phys. Rev. Lett. **105** 1413 (1957).

[11] M. Gell-Mann and A. Pais, Phys. Rev. D **97** 1387 (1955).

[12] K. Lande *et al.*, Phys. Rev. D **103** 1901 (1956).

[13] N. Cabibbo, Phys. Rev. Lett. **10** 531 (1963).

[14] M. Gell-Mann, Phys. Lett. **8** 214 (1964).

[15] S. Glashow, J. Illiopolous and L. Maiani, Phys. Rev. D **2** 1285 (1970).

[16] M. K. Gaillard and B. W. Lee, Phys. Rev. D **10** 897 (1974).

[17] L. Wolfenstein, Phys. Rev. Lett. **51** 1945 (1983).

[18] D.E. Groom *et al.* (Particle Data Group), Eur. Phys. J. **C15** 1 (2000). Eur. Phys. J. **C15**, 1 (2000).

[19] T. Kamae, *Proceedings of the Workshop on Physics and Detectors for the KEK Asymmetric B-Factory*, KEK proceedings 91-7, p.506.

[20] I. Dunietz, H. Quinn, A. Snyder, W. Toki, and H. Lipkin, Phys. Rev. D **43** 2193 (1991).

[21] Belle Collaboration, BELLE-CONF-0105 (2001).

[22] H. Höcker, H. Lacker, S. Laplace and F. Le Diberder, Eur. Phys. J. **C21** 225 (2001), hep-ph/00104062. CKM fitter home page: `http://www.slac.stanford.edu/~laplace/ckmfitter.html`

[23] K. Ackerstaff *et al.* (OPAL collaboration), Eur. Phys. J. **C 5** 379 (1998).

[24] T. Affolder *et al.* (CDF collaboration), Phys. Rev. D **61** 0702005 (2001).

[25] V. Papadimitriou, presentation at 5th KEK topical conference, Nov. 2001, KEK, Tsukuba.

[26] R. Barate *et al.* (ALEPH collaboration), Phys. Lett. B **492** 259 (2000).

[27] Belle Collaboration, Technical Design Report, KEK Report 95-7 (1995); Belle Collaboration, BELLE Progress Report, KEK Report 96-1; Belle Collaboration, BELLE Progress Report, KEK Report 97-1.

[28] A. Abashian *et al.*, KEK Progress Report 2000-4, submitted to Nucl. Instr. Meth.

[29] KEKB accelerator group, KEK B-Factory Design Report, KEK Report 95-7 (1995).

[30] Belle SVD group, Belle SVD Technical Design Report; G. Alimonti *et al.*, Nucl. Instr. Meth. **A453** 71 (2000); R. Abe *et al.*, IEEE Trans. Nucl. Sci. **48** 997 (2001).

[31] M. Yokoyama, "Design, Construction and Performance of the Belle Silicon Vertex Detector", Master thesis.

[32] V. Chabaud *et al.* (DELPHI Collaboration), Nucl. Instr. Meth. **A368** 314 (1996).

[33] M. Tanaka *et al.*, Nucl. Instr. Meth. **A432** 422 (1999).

[34] E. Nygård *et al.*, Nucl. Instr. Meth. **A301** 506 (1991); O. Toker *et al.*, Nucl. Instr. Meth. **A340** 572 (1994).

[35] M. Yokoyama and H. Tajima, Belle note 196, unpublished.

[36] M. Yokoyama *et al.*, IEEE Trans. Nucl. Sci. **48** 440 (2001).

[37] M. Hazumi, H. Tagomori, A. Bożek, and M. Yokoyama, Belle note 325, unpublished.

[38] H. Hirano *et al.*, Nucl. Instr. Meth. **A455** 294 (2000); M.Akatsu *et al.*, Nucl. Instr. Meth. **A454** 322 (2000).

[39] R. Frühwirth, Nucl. Instr. Meth. **A262** 444 (1987).

[40] T. Iijima *et al.*, Nucl. Instr. Meth. **A453** 321 (2000).

[41] H. Kichimi *et al.*, Nucl. Instr. Meth. **A453** 315 (2000).

[42] H. Hamasaki *et al.* (Belle KID group), Belle note 321, unpublished.

[43] H. Ikeda *et al.*, Nucl. Instr. Meth. **A441** 401 (2000).

[44] K. Hanagaki *et al.*, submitted to Nucl. Instr. Meth.

[45] A. Abashian *et al.*, Nucl. Instr. Meth. **A449** 112 (2000).

[46] S. Kumer, Belle note 351, unpublished.

[47] L. Piilonen *et al.*, Belle note 338, unpublished.

[48] Y. Ushiroda *et al.*, Nucl. Instr. Meth. **A438** 460 (1999).

[49] QQ event generator, `http://www.lns.cornell.edu/public/CLEO/soft/QQ` .

[50] T. Sjöstrand, Comput. Phys. Commun. **82** (1994) 74; T. Sjöstrand and M. Bengston *ibid* **43** (1987) 367; T. Sjöstrand, *ibid* **39** (1986) 347.

[51] GEANT Detector Description and Simulation Tool, CERN Program Library Long Writeup W5013.

[52] B. Casey, Belle note 296 and 390, unpublished.

[53] G. Fox and S. Wolfram, Phys. Rev. Lett. **41** 1581 (1978).

[54] J. Tanaka, Belle note 193 and 194 , unpublished.

[55] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, New Jersey (1958).

[56] MINUIT, Function Minimization and Error Analysis, CERN Program Library Long Writeup D506.

[57] H.Tajima (Belle Collaboration), Measurements of Heavy Meson Lifetimes with Belle, in the proceedings of 30th International Conference on High-Energy Physics.

[58] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **241** 278 (1990); Phys. Lett. B **254** 288 (1991).

[59] T. Skwarnicki, Ph.D. thesis, Institute for Nuclear Physicsm Krakow, 1986; DESY Internal Report, DESY-F31-86-02 (1986).

[60] G. J. Feldman and R. D. Cousins, Phys. Rev. D **57** 3873 (1998).

[61] T. Higuchi, Ph.D thesis.

[62] K. Abe *et al.* (Belle collaboration), Phys. Rev. Lett. **87** 091802 (2001).

[63] B. Aubert *et al.* (BaBar collaboration), Phys. Rev. Lett. **87** 091801 (2001).

[64] T. Nakadaira (U. Tokyo), private communication.

[65] Y. Ohnishi, Belle Note 148, unpublished.

[66] M. Różańska, Belle note 150, unpublished; Belle note 256, unpublished.

[67] H. W. Zhao, Belle note 289, unpublished.

[68] V. Chabaud *et al.* Alignment of DELPHI vertex detector, DELPHI 96-177

[69] K. Inami (Nagoya U.), private communication.

[70] Belle SVD group, Technical Design Report of Belle SVD2, unpublished.

[71] T. P. Ma and P. V. Dressendorfer, *Ionizing Radiation effects in MOS Devices and Circuits.* Wiley, New York (1989).

[72] W. Snoeys *et al.*, Nucl. Instr. Meth. **A439** 349 (2000).

[73] `http://www.ideas.no`

[74] J. Kaneko, Improvement of radiation-hardness of Double-sided Silicon Strip Detector, presentation at 2000 IEEE Nuclear Science Symposium and Medical Imaging Conference, Oct. 2000, Lyon, France.