Observation of CP Violation with B^0 Meson Decaying to the $J/\psi K_S$ State

Takeo Higuchi

Department of Physics, Faculty of Science,
The University of Tokyo

December 2001

Abstract

We report observation of a time-dependent CP asymmetry in the neutral B meson system. A total of 31.3×10^6 B-$\overline{B}$ pairs produced by the KEKB accelerator are recorded by the Belle detector. We collect 387 neutral B mesons via $B^0 \to J/\psi K_S$ decay mode. The flavor of accompanying B meson is determined mainly from the leptons and kaons in the decay product of the associated side B meson to $J/\psi K_S$ decay. The proper-time difference of B meson decays is reconstructed from the distance of two decay vertices of both B mesons. We perform an unbinned-maximum-likelihood fit to determine $\sin 2\phi_1$ from the reconstructed proper-time difference distribution. We determine $\sin 2\phi_1 = 0.81 \pm 0.20(\mathrm{stat}) \pm 0.04(\mathrm{syst})$. Combining this mode with following neutral B decay modes: $\psi(2S)K_S$, $\chi_{c1}K_S$, $\eta_c K_S$, $J/\psi K_L$, and $J/\psi K^{*0}$ we also determine $\sin 2\phi_1 = 0.99 \pm 0.14(\mathrm{stat}) \pm 0.06(\mathrm{syst})$. We conclude that we have observed CP violation in the neutral B meson system.

Contents

Preface

It is one of the great puzzles in the cosmology that the Universe is composed entirely of matter, while equal amount of matter and antimatter are produced during the Big Bang in the present theory. The CP asymmetry, *i.e.* matter-antimatter asymmetry is a key to understand this puzzle.

The CP violation was first discovered in neutral K meson system in 1964 by Christenson *et al.* [1]. This discovery initiated many theoretical efforts to understand the CP violation phenomenon. In 1973, Kobayashi and Maskawa, two Japanese physicists, proposed a model (KM-model) to explain the CP violation [2]. They concluded that one or more irreducible complex phases remain in the quark-mixing matrix in the weak charged current when six or more quarks exist, even though only three types of quarks, up (u), down (d), and strange (s) are known at that period. These complex phases violate the CP symmetry in the framework of the Standard Model. Subsequently J/ψ was discovered, which proved the existence the charm quark, c, in 1974 [3], and the observations of the b- and t-quarks establish the KM-model as a essential part of the Standard Model. However, the magnitude and the phase of the quark-mixing matrix known as CKM-matrix (Cabbibo-Kobayashi-Maskawa matrix) are not experimentally well measured. It is of great importance for the particle physics to precisely measure the CKM matrix and test its consistency.

In 1980, Carter, Bigi and Sanda pointed out that the sizable CP violation can be observed in the B decays in the framework of the Standard Model. $B^0 \to J/\psi K_S$ is one of the most promising modes to observe the CP violation, because of its low experimental backgrounds and negligible theoretical uncertainties. Moreover, the branching fraction of this mode is relatively large, $(3.6 \pm 0.1) \times 10^{-5}$. In this decay mode, CP violation occurs through interference of two decay amplitudes, $B^0 \to J/\psi K_S$ and $B^0 \to \overline{B}^0 \to J/\psi K_S$. When B^0 and $\overline{B}^0$ are produced by $\Upsilon(4S)$ decay, the time dependence of the $B^0 \to J/\psi K_S$ decay rate can be expressed as

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}(\Delta t)} = \Gamma \mathrm{e}^{-\frac{|\Delta t|}{\tau_{B^0}}} \left[1 + q \cdot \sin 2\phi_1 \cdot \sin(\Delta M \Delta t)\right], \tag{1}$$

where Δt is proper-time difference of two B decays, q is $+1$ for $\overline{B}^0 \to J/\psi K_S$ and -1 for

$B^0 \to J/\psi K_S$, ϕ_1 is a phase difference of some CKM-matrix elements, and ΔM is the mass difference of two mass eigenstates of neutral B mesons.

To observe the CP violation in the B meson system and to test the KM-model, B-factory experiment [4] referred as Belle experiment has been started since 1994 at KEK, Japan. As the asymmetry equation states, the CP asymmetry would vanish with the time integration. Therefore, precise determination of the proper-time difference is required to observe the CP violation. To meet the requirement, the $\Upsilon(4S)$ is produced from e^+e^- collision by the KEKB accelerator [5] with asymmetric energy to give the $\Upsilon(4S)$ a motion. Consequently the $B^0\overline{B}^0$ pair possess the motion and thus Δt can be calculated from the distance of the decay vertices of the two B mesons along with the beam direction:

$$\Delta t \simeq \frac{\Delta z}{c\beta\gamma},$$

where Δz is the decay vertex distance, $\beta\gamma$ is the Lorentz boost factor of the $\Upsilon(4S)$ due to the asymmetric beam energy, and c is the light velocity. We configure the beam energy as 8.0 GeV for e^- and 3.5 GeV for e^+, which corresponds to $\beta\gamma = 0.425$. For B^0 meson lifetime of ~ 1.5 ps, typical Δz becomes 200 μm in distance, which is sizable length under the current experimental technique. In addition, the collider is expected to provide enormous number of $B^0\overline{B}^0$ pairs because of a small branching fraction of $\mathcal{O}(10^{-4})$. The Belle detector [6] is required to provide precise measurement of Δz, and good particle reconstruction and identification capability for full reconstruction of B^0 mesons and for B^0 or $\overline{B}^0$ flavor determination. The event triggering and data acquisition systems are required to withstand high event rate. We started taking data from June 1999. By August 2001, we recorded 29.1 fb^{-1} data with the Belle detector, which corresponds to 31.3×10^6 B meson pairs.

The importance of the research of the CP violation in B decays is reflected in the number of laboratories engaged in the measurement. Several measurements of the CP violation in the B meson system were presented by the OPAL, ALEPH, and the CDF collaboration, however a conclusive evidence of the CP violation in the B meson system has not been observed by this study.

In this thesis, we present the first significant observation of the CP violation in the B meson system through the $B^0 \to J/\psi K_S$ decay with the Belle detector. This thesis is organized as follows: the CP violation in the B decay is discussed in Chapter 1. An experimental apparatus, consisting of the KEKB accelerator, the Belle detector including the trigger and the data acquisition system is described in Chapter 2. In Chapter 3, we describe the event reconstruction of $B^0 \to J/\psi K_S$ decay. The procedure to determine the CP violation parameter is described in Chapter 4. We discuss the significance of observed asymmetry in Chapter 5. Chapter 6 concludes this thesis.

Chapter 1

CP Violation in the B Meson System

The violation of CP symmetry is one of the most interesting topics of high-energy physics today. Experimentally, it is one of the least tested properties of the Standard Model. This chapter explains a basic theory of the CP violation in B decays and measurement technique of the unitarity triangle. In Section 1.1, we give an introduction of the CP violation reviewing the definition of C, P, and T, and the discovery of the CP violation in K decays. Next we review the weak interaction in the Standard Model and describe the KM-mechanism in terms of the Cabbibo-Kobayashi-Maskawa matrix, which is regarded as the origin of the CP violation in Section 1.2. We describe the CP violation through the B^0-$\overline{B}^0$ mixing in Section 1.3. In the end the of this chapter, experimental measurement of the CP violation at an asymmetric B-factory is described.

1.1 Introduction to CP Violation

In the quantum theory, there are conservation laws corresponding to discrete transformations. In this section, we give the definitions of P, C, and T operation, followed by the discovery of the CP violation, which was made in neutral K meson system.

1.1.1 Definitions of P, C, and T Operators

A P transformation is a parity transformation. The operator P changes the sign of a spatial vector $\vec{r}$. The sign of a momentum vector $\vec{p}$ is also changed, while the sign of an angular momentum vector, $\vec{L} = \vec{r} \times \vec{p}$ is invariant by this operation. Let us define a spin and momentum of particle f as $\vec{s}$ and $\vec{p}$, respectively. The particle f at the state $|f; \vec{s}, \vec{p}\rangle$ is transformed by the operator P as

$$P|f; \vec{s}, \vec{p}\rangle = \eta_P |f; \vec{s}, -\vec{p}\rangle. \tag{1.1}$$

Since P^2 is an identical operator, η_P^2 is 1, which is followed by $\eta_P = \pm 1$. The sign of η_P is chosen arbitrary as far as the definition is consistent through the discussion.
C operator changes the sign of internal charges (like electric charge, baryon number, etc). It is called charge conjugation operator. The spin and the momentum of f are unchanged by the C operator. C operates on the state f as

$$C|f;\vec{s},\vec{p}\rangle = \eta_C|\overline{f};\vec{s},\vec{p}\rangle. \tag{1.2}$$

The third fundamental operator is T transformation. The operator is defined as time reversal. T operator flips the direction of time evolution. The time reversal changes momentum vector $\vec{p}$ to $-\vec{p}$ and spin vector $\vec{s}$ to $-\vec{s}$.

$$T|f;\vec{s},\vec{p}\rangle = \eta_T|f;-\vec{s},-\vec{p}\rangle, \tag{1.3}$$

where η_T is a phase.

According to the CPT theorem of Lüders and Pauli [7], the combined operations of C, P and, T is invariant under rather weak assumption in a local field theory.

1.1.2 Discovery of CP Violation

For a long time it was assumed that all elementary processes are invariant under the application of each of the three operations, C, P, and T separately. However, the work of Lee and Yang questioned the assumption in 1956 [8], and the subsequent experiments by Wu *et al.* and Garwin *et al.* in 1957 independently demonstrated the violation of P and C invariance in weak decays of nuclei and of pions and muons [9] [10]. This violation can be visualized by the longitudinal polarization of neutrinos emerging from a weak vertex: they are left-handed when they are particles and right-handed when antiparticles. Application of P or C to a neutrino leads to an unphysical state (Figure 1.1). The combined operation CP, however, transforms a left-handed neutrino into right-handed antineutrino, thus connecting two physics states. CP invariance was therefore considered to be replacing the separate P and C invariance of weak interactions.
In 1964, Christenson, Cronin, Fitch and Turlay discovered the violation of the CP symmetry in K^0 meson decays. There are two neutral strange mesons, K^0 and $\overline{K}^0$. The quark content of the K^0 is $\overline{s}d$, and $\overline{K}^0$ is $s\overline{d}$. With definition of $CP|K^0\rangle = -|\overline{K}^0\rangle$, we can create CP eigenstates of neutral K mesons, labeled with 1 and 2 as an linear combinations of $|K^0\rangle$ and $|\overline{K}^0\rangle$ as

$$|K_1\rangle = \frac{1}{\sqrt{2}}\left(|K^0\rangle - |\overline{K}^0\rangle\right), \qquad |K_2\rangle = \frac{1}{\sqrt{2}}\left(|K^0\rangle + |\overline{K}^0\rangle\right), \tag{1.4}$$

where CP eigenvalues of $|K_1\rangle$ and $|K_2\rangle$ are $+1$ and -1, respectively. Experimentally, it was observed that there are two lifetimes of neutral K mesons, where the shorter lived

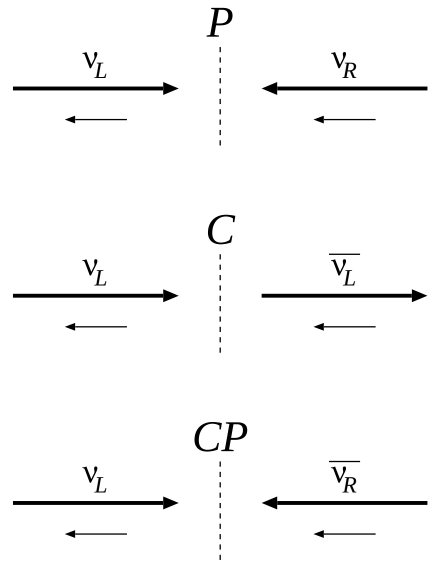


Figure 1.1: Mirror images of a left-handed neutrino under P, C and CP symmetry. Long and short arrows represent momenta and spins of the neutrinos.

was named K_S, and the longer lived was named K_L. The lifetime of K_S is ~ 90 ps and that of K_L is ~ 50 ns. It was also measured that K_S decays into two-pions and K_L decays into three pions. Since CP eigenvalue of two pion state is $+1$ and that of three pions was -1, the common consideration at that period was that the K_S corresponds to the K_1, and the K_L corresponds to the K_2, respectively. The difference of lifetimes was considered to be due to the phase space difference. The discovery made by Christenson *et al.* group was the small fractional decay of $K_L \to \pi^+\pi^-$. The branching fraction was order of $\sim 10^{-3}$. Even though it was small, it was an evidence for that K_S and K_L are not "real" CP eigenstates and should be rewritten as

$$|K_S\rangle = \frac{1}{\sqrt{1+|\epsilon_m|^2}}\left(|K_1\rangle + \epsilon_m |K_2\rangle\right), \tag{1.5}$$

$$|K_L\rangle = \frac{1}{\sqrt{1+|\epsilon_m|^2}}\left(|K_2\rangle + \epsilon_m |K_1\rangle\right). \tag{1.6}$$

Subsequent observations of $K_L \to \pi^0\pi^0$ decay [11], and charge asymmetries in $K_L \to \pi^\pm e^\mp \nu$[12] and $K_L \to \pi^\pm \mu^\mp \nu$[13] confirmed the CP violation in neutral K meson system.

The theoretical explanation within the Standard Model for the origin of the CP violation was proposed by Kobayashi and Maskawa in 1973 [2], which is described in Section 1.2.

1.2 CP Violation in Standard Model

In terms of mass eigenstates, a Lagrangian of charged-current weak interaction forms

$$-\mathcal{L}_{\mathrm{cc}} \;=\; -\frac{g_2}{\sqrt{2}}\,\overline{\boldsymbol{u}}_L\gamma^\mu V\,\boldsymbol{d}_L\,W_\mu^+ + [h.c.]. \tag{1.7}$$

In general, $\overline{\boldsymbol{u}}_L$ and $\boldsymbol{d}_L$ are an "array" of mass eigenstates of quark-fermions and V is a mixing matrix, which is unitary: $V^\dagger V = I$. When irreducible phases exist in the mixing matrix, they remain in the amplitude with different signs between $K^0 \to \overline{K}^0$ and $\overline{K}^0 \to K^0$, and then they enable the CP violation to occur. This is a key idea of the KM-mechanism to describe the CP violation within the Standard Model. In this section, first we trace the historical development of weak interaction within the Standard Model. In the history, we see the KM-mechanism is a natural expansion of the theory at that period. In the successive paragraphs, we describe the "unitarity triangles", which are derived from the unitarity condition of V. They are worth to be described because the unitarity triangles are related to the CP violation. In the description, we review the approximated parameterization of the matrix elements of V to discuss the shapes of the unitarity triangles. After the discussion, we find more feasible CP violation in the B meson system rather than the K meson system.

The first proposal of "quark" model was proposed by Gell-Mann and Zweig in 1964 including three flavors of quarks, up (u), down (d), and strange (s) [14]. Citing a paper [15] written by Cabbibo, Gell-Mann also proposed that weak interactions would couple the u-quark to a combined state of d and s as $|d'\rangle = \cos\theta|d\rangle + \sin\theta|s\rangle$. With this definition, all (known) weak interactions could be described by a single coupling constant. In 1970, S. L. Glashow, J. Illiopoulos, and L. Maiani proposed a theory that a fourth flavor of quark, which they labeled "charm" (c), can explained why K^0 meson was not observed to decay to $\mu^+\mu^-$ [16]. Figure 1.2 (a) shows a diagram of K^0 decay to $\mu^+\mu^-$ whose amplitude is proportional to $\sin\theta_C\cos\theta_C$. To explain the small decay rate of $K^0 \to \mu^+\mu^-$ process, they introduced a coupling of c quark and $|s'\rangle$ state defined as $|s'\rangle = -\sin\theta_C|d\rangle + \cos\theta_C|s\rangle$ to "cancel" the amplitude corresponding to the diagram of Figure 1.2 (a). The canceling diagram is shown in Figure 1.2 (b), which has proportional amplitude to $-\sin\theta_C\cos\theta_C$. Now, the combined states of d- and s-quarks are described by "mixing matrix" as

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}, \tag{1.8}$$

where θ_C is known as Cabbibo-angle. The angle is measured to be $\sin\theta_C \sim 0.22$.

In 1973, M. Kobayashi and T. Maskawa proposed that when at least three quark "generations" exist, the CP violation could be explained within the Standard Model. It was

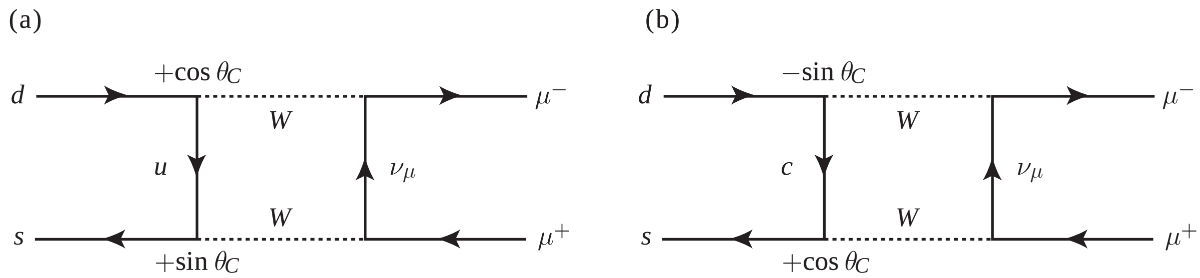


Figure 1.2: Paths for the decay of $K^0 \to \mu^+\mu^-$. The diagram (b) exchanging c quark for d quark is introduced by the GIM-mechanism.

the extension of a dimension of the quark-mixing matrix from two to three or more. A general $n \times n$ complex matrix, $V = \{V_{ij}\}$, has $2n^2$ free parameters. The $SU(n)$ symmetry of the mixing matrix requires $\sum_j V_{ij}V^*_{jk} = \delta_{ik}$, where δ_{ik} is Kronecker's δ, and it yields n constraints for $i = k$ and $n^2 - n$ constraints for $i \neq k$. Thus, $n \times n$ unitary matrix has n^2 free parameters. The phases of the quark fields can be rotated freely. Since the overall phase is irrelevant, $2n - 1$ relative phase can be removed from V. Accordingly, V has $(n-1)^2$ free observables. On the other hand, a general $n \times n$ orthogonal matrix has $n(n-1)/2$ rotational Euler angle. Thus, $(n-1)(n-2)/2$ parameters corresponding to irreducible "phase" remain in V. To summarize, general $n \times n$ matrix representing mixing matrix consists of $n(n-1)/2$ rotational Euler angles and $(n-1)(n-2)/2$ irreducible phases independently. The least number for n to generate irreducible phase in V is three. The expanded mixing matrix from equation (1.8) was named Cabbibo-Kobayashi-Maskawa matrix (CKM matrix). Subsequent discoveries of c, bottom (b), and top (t) quarks make the KM-mechanism a essential part of the Standard Model. We define the matrix elements of V as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V \begin{pmatrix} d \\ s \\ b \end{pmatrix} \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.9}$$

To clarify a hierarchy in the size of the matrix elements, let us review an approximated representation of V, proposed by Wolfenstein [17]. The diagonal elements of the matrix are nearly one, and among the off-diagonal elements, dominating terms are $V_{us} \simeq V_{cd} \simeq \sin\theta_C \equiv \lambda$. Thus, to order of λ^2, the upper left 2×2 sub-matrix in the CKM matrix can be constructed as the GIM four-quark model and we have

$$V \simeq \begin{pmatrix} 1 - (\lambda^2/2) & \lambda & \star \\ -\lambda & 1 - (\lambda^2/2) & \star \\ \star & \star & 1 \end{pmatrix}. \tag{1.10}$$

The observation that $V_{cb} \simeq 0.04$ allows us to express it as $A\lambda^2$, where A is an order of unity. Unitarity requires $V_{ts} \simeq -A\lambda^2$ as long as V_{td} and V_{ub} are small enough. Finally, V_{ub}

appears to be of order $A\lambda^3 \times \mathcal{O}(1)$. Here, we have to allow for one phase. We write V_{ub} as $V_{ub} = A\lambda^3(\rho - \mathrm{i}\eta)$, where ρ and η are also orders of unity. Finally, unitarity specifies uniquely the form $V_{td} = A\lambda^3(1-\rho-\mathrm{i}\eta)$. To summarize, the CKM matrix is approximated as

$$V \simeq \begin{pmatrix} 1-(\lambda^2/2) & \lambda & A\lambda^3(\rho-\mathrm{i}\eta) \\ -\lambda & 1-(\lambda^2/2) & A\lambda^2 \\ A\lambda^3(1-\rho-\mathrm{i}\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{1.11}$$

Three Euler angles and one phase are replaced by real numbers of λ, A, ρ, and η. This representation is known as Wolfenstein parameterization.

The geometrical interpretation of the unitarity of the CKM matrix is a good stage to comprehend the feasibility of the CP violation in the B meson system. The requirements that the CKM matrix must be unitary leads relationships among its elements as

$$\begin{aligned}
V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* &= 0, & (1.12)\\
V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* &= 0, & (1.13)\\
V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* &= 0, & (1.14)\\
V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* &= 0, & (1.15)\\
V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* &= 0, & (1.16)\\
V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* &= 0. & (1.17)
\end{aligned}$$

The relationships form six triangles in the complex plane, which are called "unitarity triangles". Figures 1.3 (a) through (f) show the unitarity triangles constructed by unitarity conditions from (1.12) through (1.17), in the same order. Lengths of the triangle sides are the magnitudes of products of the CKM-matrix elements that can be measured as decay rates. Angles correspond to the complex phase that can cause asymmetry of the amplitudes between quark and antiquark transition. Therefore, an observation of the CP violation can lead to a measurement of the angles. According to the matrix elements of each side, one can find the triangle in Figure 1.3 (a) corresponds to the decays in the neutral K meson system. Because two sides are order of $\mathcal{O}(\lambda)$ and one side is order of $\mathcal{O}(\lambda^5)$, this triangle is extremely squashed, and thus, it is hard to measure the side and the angles, except for two long sides. This causes relative smaller CP violation in K meson system. Because of the same reason, triangle represented by Figure 1.3 (b) also has an extremely squashed shape. The angles of the triangles of Figure 1.3 (e) and (f) are equally large, since the lengths of all of three sides are the same order of $\mathcal{O}(\lambda^3)$. The triangle of Figure 1.3 (f) relates to the B meson decay. Thus, we can expect relative larger CP violation in the B meson system than the K meson system. The triangle of Figure 1.3 (e) relates to the top quark decays. We can also image a large CP violation

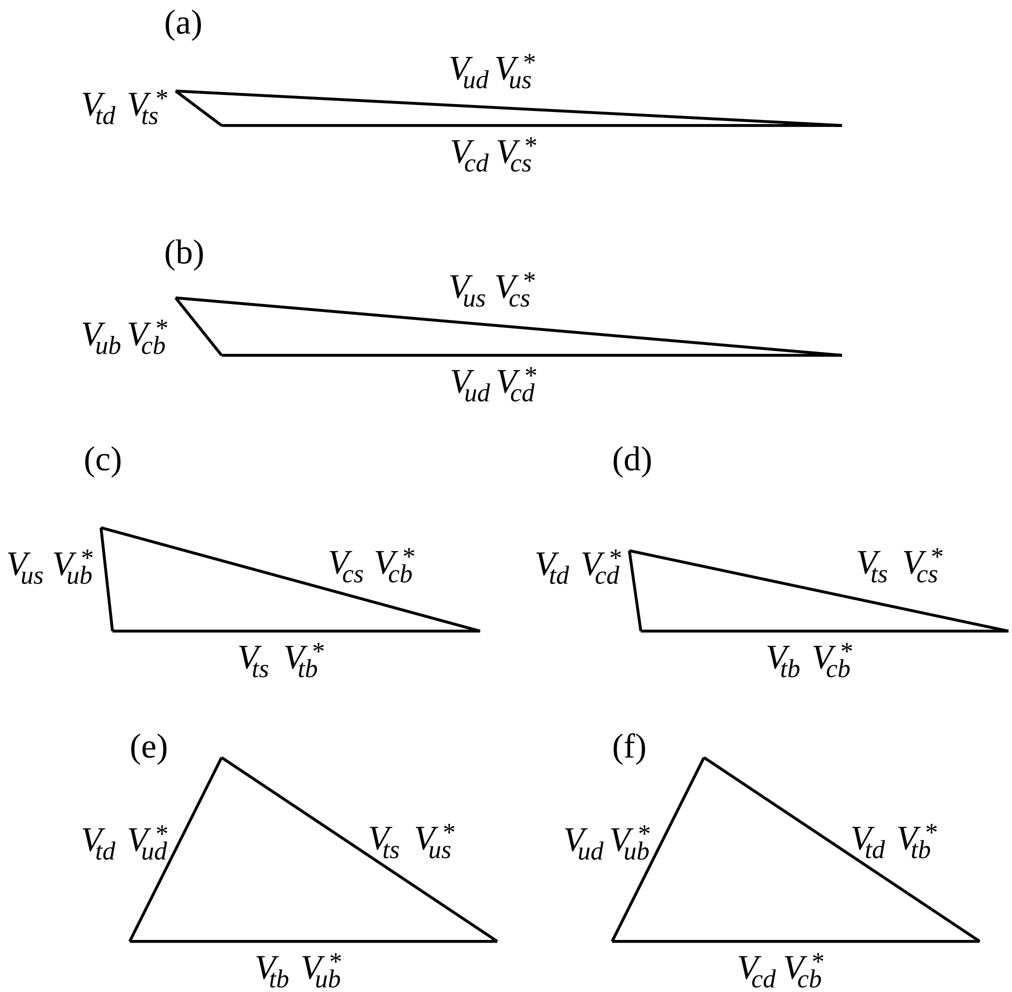


Figure 1.3: Six unitarity triangles constructed by the unitarity conditions from (1.12) through (1.17).

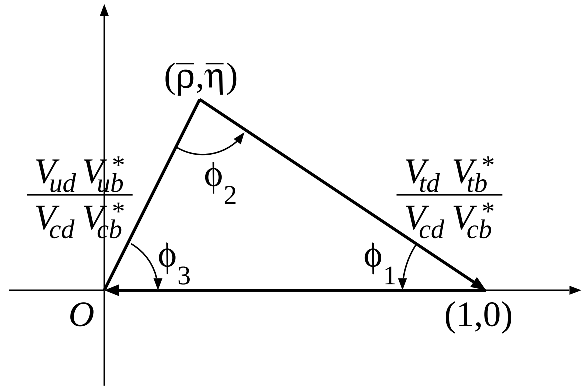


Figure 1.4: Scaled triangle of Figure 1.3 (f) by $1/|V_{cd}V_{cb}^*|$. The definitions of three angles, ϕ_1, ϕ_2, and ϕ_3 are also shown.

in the t-quark system, however, the examination of the t-quark decays is far beyond our current experimental technique.

It is convenient to divide all sides of the triangle of Figure 1.3 (f) by $|V_{cd}V_{cb}^*|$. Coordinates of the triangle are $(0,0)$, $(1,0)$ and $(\tilde{\rho},\tilde{\eta})$, where $\tilde{\rho}$ and $\tilde{\eta}$ are defined with Wolfenstein's parameterization as

$$\tilde{\rho} \equiv \left(1 - \frac{\lambda^2}{2}\right)\rho, \qquad \tilde{\eta} \equiv \left(1 - \frac{\lambda^2}{2}\right)\eta. \tag{1.18}$$

The three angles of the triangle shown in Figure 1.4 are defined as

$$\phi_1 = \pi - \arg\left(\frac{-V_{td}V_{tb}^*}{-V_{cb}^*V_{cd}}\right), \quad \phi_2 = \arg\left(\frac{V_{tb}V_{td}^*}{-V_{ub}^*V_{ud}}\right), \quad \phi_3 = \arg\left(\frac{V_{ub}V_{ud}^*}{-V_{cb}^*V_{cd}}\right). \tag{1.19}$$

A crucial test of the Standard Model is to evaluate the consistency of all sides and angles of this unitarity triangle. This test may provide information on the dynamical origin of the quark-mixing matrix.

1.3 *CP* Violation in *B* Decays

In 1980, A. Carter, I. I. Bigi, and A. I. Sanda pointed out that the KM-mechanism indicates the possibility of sizable CP violation in the B meson system [18]. As we see in this section, neutral mesons, namely B meson, mix with their antiparticles. In the B meson system, sizable CP asymmetries are expected in the interference between decays into a same final state with and without the B^0-$\overline{B}^0$ mixing. The CP asymmetry is observed in the difference between the time-dependent decay rates of B^0 and $\overline{B}^0$ mesons into a common CP eigenstate. In this section, we first explain the phenomenology of time evolution of neutral B mesons. Then we consider the case that both B^0 and $\overline{B}^0$

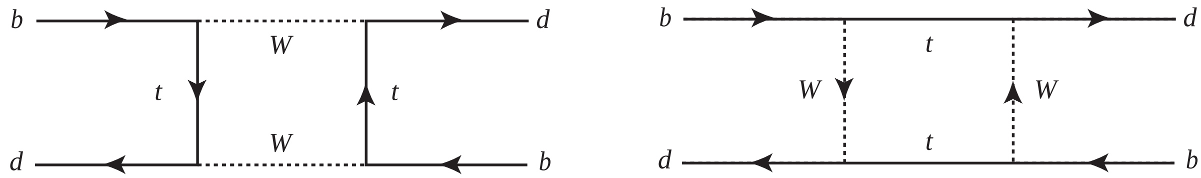


Figure 1.5: Two diagrams contributing for B^0-$\overline{B}^0$ mixing.

decay into the same CP eigenstate and the CP violation in these decays. Finally, we relate the CP violation in $B^0 \to J/\psi K_S$, of which final state is the CP eigenstate.

1.3.1 B^0-$\overline{B}^0$ Mixing

B^0 and $\overline{B}^0$ can mix through second order weak interactions via diagrams shown in Figure 1.5 known as "box diagrams". In this subsection, we see the time-evolution of B^0 and $\overline{B}^0$. Due to the mixing, an arbitrary neutral B meson state is written as the admixture of B^0 and $\overline{B}^0$,

$$|\Phi(t)\rangle = |B^0(t)\rangle + |\overline{B}^0(t)\rangle, \tag{1.20}$$

where $CP|B^0\rangle = -|\overline{B}^0\rangle$. The time-dependent Schrödinger equation of $|\Phi(t)\rangle$ is

$$\mathrm{i}\frac{\partial}{\partial t}|\Phi(t)\rangle = \mathcal{H}|\Phi(t)\rangle. \tag{1.21}$$

$\mathcal{H}$ is an Hamiltonian defined as

$$\mathcal{H} \equiv \mathcal{M} - \mathrm{i}\Gamma \equiv \begin{pmatrix} M_{11} - \frac{\mathrm{i}}{2}\Gamma_{11} & M_{12} - \frac{\mathrm{i}}{2}\Gamma_{12} \\ M_{21} - \frac{\mathrm{i}}{2}\Gamma_{21} & M_{22} - \frac{\mathrm{i}}{2}\Gamma_{22} \end{pmatrix}. \tag{1.22}$$

Using $\mathcal{H}^\dagger = \mathcal{H}$, eigenvalues $\mu_\pm$, and eigenvectors $|B_\pm\rangle$ of above equation can be given as

$$\mu_\pm = M_0 - \frac{\mathrm{i}}{2}\Gamma_0 \pm \sqrt{\left(M_{12} - \frac{\mathrm{i}}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{\mathrm{i}}{2}\Gamma_{12}^*\right)}, \tag{1.23}$$

and

$$|B_\pm\rangle = \frac{1}{\sqrt{|p|^2 + |q|^2}}\left(p|B^0\rangle \pm q|\overline{B}^0\rangle\right), \tag{1.24}$$

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{\mathrm{i}}{2}\Gamma_{12}^*}{M_{12} - \frac{\mathrm{i}}{2}\Gamma_{12}}}. \tag{1.25}$$

In the following discussion, we consider p and q are normalized as $|p|^2 + |q|^2 = 1$. Here the lifetimes and masses of B^0 and $\overline{B}^0$ are assumed to be equal because of the CPT theorem,

i.e. $M_0 \equiv M_{11} = M_{22}$, and $\Gamma_0 \equiv \Gamma_{11} = \Gamma_{22}$. Equation (1.24) indicates that the general mass eigenstates are deviated from the eigenstates of the weak interaction.
We redefine mass eigenvectors to fit to usual convention,

$$|B_H\rangle = p|B^0\rangle - q|\overline{B}^0\rangle, \tag{1.26}$$

$$|B_L\rangle = p|B^0\rangle + q|\overline{B}^0\rangle, \tag{1.27}$$

where corresponding eigenvalues are $\mu_H \equiv \mu_-$ and $\mu_L \equiv \mu_+$, respectively. Defining mass and width of $B_{H,L}$ as $M_{H,L}$ and $\Gamma_{H,L}$, time evolutions of $|B_H\rangle$ and $|B_L\rangle$ are given by

$$|B_H(t)\rangle = \mathrm{e}^{-\mathrm{i}\mu_H t}|B_H\rangle \equiv \mathrm{e}^{-\mathrm{i}M_H t}\mathrm{e}^{-\frac{\Gamma_H t}{2}}|B_H\rangle, \tag{1.28}$$

$$|B_L(t)\rangle = \mathrm{e}^{-\mathrm{i}\mu_L t}|B_L\rangle \equiv \mathrm{e}^{-\mathrm{i}M_L t}\mathrm{e}^{-\frac{\Gamma_L t}{2}}|B_L\rangle. \tag{1.29}$$

Next, we transform the base from the mass eigenstates in equations (1.28) and (1.29) to the weak interaction eigenstates. First, suppose a neutral B meson being $|B^0\rangle$ at time $t = 0$. The wave function of the particle can be expressed with equations (1.24), (1.26), and (1.27) as $|B^0\rangle = (|B_H\rangle + |B_L\rangle)/2p$. and therefore, the time-evolution of $|B^0(t)\rangle$ is written using equations from (1.26) through (1.29) as

$$|B^0(t)\rangle = \frac{1}{2}\Big[\left(\mathrm{e}^{-\mathrm{i}\mu_H t} + \mathrm{e}^{-\mathrm{i}\mu_L t}\right)|B^0\rangle - \frac{q}{p}\left(\mathrm{e}^{-\mathrm{i}\mu_H t} - \mathrm{e}^{-\mathrm{i}\mu_L t}\right)|\overline{B}^0\rangle\Big]. \tag{1.30}$$

In the same manner, the time-evolution of $|\overline{B}^0(t)\rangle$ is written as

$$|\overline{B}^0(t)\rangle = -\frac{1}{2}\Big[\frac{p}{q}\left(\mathrm{e}^{-\mathrm{i}\mu_H t} - \mathrm{e}^{-\mathrm{i}\mu_L t}\right)|B^0\rangle - \left(\mathrm{e}^{-\mathrm{i}\mu_H t} + \mathrm{e}^{-\mathrm{i}\mu_L t}\right)|\overline{B}^0\rangle\Big]. \tag{1.31}$$

1.3.2 B Meson Decay into CP Eigenstate

In this subsection, we see the time evolution of B meson decay into CP eigenstate. We start from the calculation of four amplitudes of generic decay final state for $B^0 \to f$ and its CP conjugation modes: $\overline{B}^0 \to f$, $B^0 \to \overline{f}$, and $\overline{B}^0 \to \overline{f}$. Then we set $f = \overline{f} = f_{CP}$. After calculating the decay rates, we discuss the difference of them, and finally we see that the difference is induced by the B^0-$\overline{B}^0$ mixing.

We define instantaneous decay amplitudes of B^0 and $\overline{B}^0$ to f and $\overline{f}$ as

$$\begin{array}{lllll} a \equiv A(B^0 \to f) = \left\langle f|\mathcal{H}_\mathrm{w}|B^0\right\rangle, & & a' \equiv A(\overline{B}^0 \to f) = \left\langle f|\mathcal{H}_\mathrm{w}|\overline{B}^0\right\rangle, \\ b \equiv A(\overline{B}^0 \to \overline{f}) = \left\langle \overline{f}|\mathcal{H}_\mathrm{w}|\overline{B}^0\right\rangle, & & b' \equiv A(B^0 \to \overline{f}) = \left\langle \overline{f}|\mathcal{H}_\mathrm{w}|B^0\right\rangle, \end{array} \tag{1.32}$$

where $\mathcal{H}_\mathrm{w}$ is a weak-decay Hamiltonian.
Let $A_{B\to f}(t)$ be the time-dependent decay amplitude for a pure B^0 state at $t = 0$ to decay into a final state f at time t. It can be obtained by the replacement of $|B^0\rangle$ and $|\overline{B}^0\rangle$ in equation (1.30) by a and a', respectively. Similarly, $A_{\overline{B}\to\overline{f}}(t)$ can be obtained by

replacing $|B^0\rangle$ and $|\overline{B}^0\rangle$ in equation (1.31) by b' and b, respectively. The amplitudes can be written as

$$A_{B\to f}(t) = \frac{a}{2}\Big[\left(\mathrm{e}^{-\mathrm{i}\mu_H t}+\mathrm{e}^{-\mathrm{i}\mu_L t}\right)-\alpha\left(\mathrm{e}^{-\mathrm{i}\mu_H t}-\mathrm{e}^{-\mathrm{i}\mu_L t}\right)\Big], \tag{1.33}$$

$$A_{\overline{B}\to\overline{f}}(t) = \frac{b}{2}\Big[\left(\mathrm{e}^{-\mathrm{i}\mu_H t}+\mathrm{e}^{-\mathrm{i}\mu_L t}\right)-\beta\left(\mathrm{e}^{-\mathrm{i}\mu_H t}-\mathrm{e}^{-\mathrm{i}\mu_L t}\right)\Big], \tag{1.34}$$

where

$$\alpha \equiv \frac{qa'}{pa}, \qquad \beta \equiv \frac{pb'}{qb}. \tag{1.35}$$

The amplitudes for the CP conjugation modes, $A_{\overline{B}\to f}$ and $A_{B\to\overline{f}}$, are similarly expressed as

$$A_{\overline{B}\to f}(t) = -\frac{a}{2}\frac{p}{q}\left[\left(\mathrm{e}^{-\mathrm{i}\mu_H t}-\mathrm{e}^{-\mathrm{i}\mu_L t}\right)-\alpha\left(\mathrm{e}^{-\mathrm{i}\mu_H t}+\mathrm{e}^{-\mathrm{i}\mu_L t}\right)\right], \tag{1.36}$$

$$A_{B\to\overline{f}}(t) = -\frac{b}{2}\frac{q}{p}\left[\left(\mathrm{e}^{-\mathrm{i}\mu_H t}-\mathrm{e}^{-\mathrm{i}\mu_L t}\right)-\beta\left(\mathrm{e}^{-\mathrm{i}\mu_H t}+\mathrm{e}^{-\mathrm{i}\mu_L t}\right)\right]. \tag{1.37}$$

Here, we introduce two new quantities: mass difference, $\Delta M \equiv M_H - M_L$, and width difference, $\Delta\Gamma \equiv \Gamma_H - \Gamma_L$. Consequently, we obtain

$$\mu_H - \mu_L = \Delta M + \mathrm{i}\Delta\Gamma. \tag{1.38}$$

We also define the average width of B_H and B_L as

$$\Gamma \equiv \frac{\Gamma_H + \Gamma_L}{2}. \tag{1.39}$$

Now we set $f = \overline{f} = f_{CP}$. Squaring the amplitudes of equations from (1.33) and (1.34), the corresponding decay distributions are obtained:

$$\Gamma_{B\to f_{CP}}(t) = \frac{|a|^2}{2}\mathrm{e}^{-\Gamma t}\Bigg[(1+|\lambda_{CP}|^2)\cosh\frac{\Delta\Gamma}{2}t + 2\mathcal{R}e(\lambda_{CP})\sinh\frac{\Delta\Gamma}{2}t + (1-|\lambda_{CP}|^2)\cos\Delta Mt - 2\mathcal{I}m(\lambda_{CP})\sin\Delta Mt\Bigg], \tag{1.40}$$

$$\Gamma_{\overline{B}\to f_{CP}}(t) = \frac{|a'|^2}{2}\mathrm{e}^{-\Gamma t}\Bigg[(1+|\lambda_{CP}|^{-2})\cosh\frac{\Delta\Gamma}{2}t + 2\mathcal{R}e({\lambda_{CP}}^{-1})\sinh\frac{\Delta\Gamma}{2}t + (1-|\lambda_{CP}|^{-2})\cos\Delta Mt - 2\mathcal{I}m({\lambda_{CP}}^{-1})\sin\Delta Mt\Bigg]. \tag{1.41}$$

We used $a = b'$, $a' = b$, and $\alpha = \beta^{-1}$ derived from equation (1.32), and we defined

$$\lambda_{CP} \equiv \alpha. \tag{1.42}$$

Squares of equations (1.36) and (1.37) yield same result as equations (1.40) and (1.41). In the B^0 meson system, $\Delta\Gamma/\Gamma \sim \mathcal{O}(10^{-3})$ can be explicitly derived from the Standard Model, and therefore, the relation $\mathrm{e}^{(\Gamma+\Delta\Gamma/2)t} \simeq \mathrm{e}^{\Gamma t}$ holds. Consequent approximations of equations (1.40) and (1.41) are:

$$\begin{aligned}
\Gamma_{B\to f_{CP}}(t) &\simeq \frac{|a|^2}{2}\mathrm{e}^{-\Gamma t}\Big[(1+|\lambda_{CP}|^2) + (1-|\lambda_{CP}|^2)\cos\Delta M t \\
&\qquad\qquad - 2\mathcal{I}m(\lambda_{CP})\sin\Delta M t\Big], \qquad (1.43)\\
\Gamma_{\overline{B}\to f_{CP}}(t) &\simeq \frac{|a'|^2}{2}\mathrm{e}^{-\Gamma t}\Big[(1+|\lambda_{CP}|^{-2}) + (1-|\lambda_{CP}|^{-2})\cos\Delta M t \\
&\qquad\qquad - 2\mathcal{I}m({\lambda_{CP}}^{-1})\sin\Delta M t\Big]. \qquad (1.44)
\end{aligned}$$

The CP invariance is violated when the time-dependent decay rate of $B^0 \to f_{CP}$ and its CP conjugation decay are different for any possible t:

$$\Gamma_{B\to f_{CP}}(t) \neq \Gamma_{\overline{B}\to f_{CP}}(t). \qquad (1.45)$$

A time-dependent asymmetry $a_{CP}(t)$ is defined as the normalized decay rate difference:

$$\begin{aligned}
a_{CP}(t) &\equiv \frac{\Gamma_{\overline{B}\to f_{CP}}(t) - \Gamma_{B\to f_{CP}}(t)}{\Gamma_{\overline{B}\to f_{CP}}(t) + \Gamma_{B\to f_{CP}}(t)} \\
&= \frac{(|\lambda_{CP}|^2 - 1)\cos\Delta M t + 2\mathcal{I}m(\lambda_{CP})\sin\Delta M t}{1+|\lambda_{CP}|^2}. \qquad (1.46)
\end{aligned}$$

1.3.3 CP Violation in $B^0 \to J/\psi K_S$

Among all $B^0 \to$ charmonium $+ K_S$ channels, $B^0 \to J/\psi K_S$, where K_S decays into two charged pions, is the most promising decay mode to extract ϕ_1 experimentally because of relatively large branching fractions with small backgrounds. In addition, as we see later in this subsection, the final state has only negligible theoretical uncertainty in ϕ_1 measurement. These are the reasons why this mode has earned the name of "golden mode" (GM). We use $\lambda_{J/\psi K_S}$ instead of λ_{CP} to indicate the decay final state being $J/\psi K_S$ explicitly, hereafter.

$\lambda_{J/\psi K_S}$ can be expressed as

$$\lambda_{J/\psi K_S} = \frac{A(B^0 \to J/\psi K_S)}{A(\overline{B}^0 \to J/\psi K_S)}\cdot\frac{q}{p}. \qquad (1.47)$$

In the following discussion, first we compute $A(B^0 \to J/\psi K_S)/A(\overline{B}^0 \to J/\psi K_S)$, then we compute q/p.

The decay of $B^0 \to J/\psi K_S$ is based on the quark transition $b \to c\bar{c}s$. Contributing Feynman diagrams to $b \to c\bar{c}s$ decays are shown in Figure 1.6, where the left diagram is

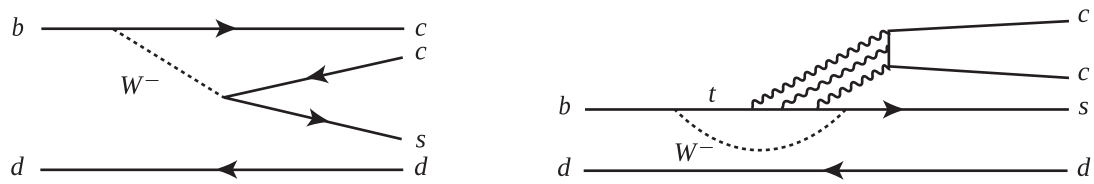


Figure 1.6: Tree (left) and penguin (right) diagrams for $B^0 \to J/\psi K_S$. The penguin contamination is negligible for $\sin 2\phi_1$ measurement.

called "tree diagram" and the right diagram holding one loop is called "(strong) penguin diagram". In the case of electroweak penguin contribution, the gluons are replaced by a Z or a γ. The amplitude of the tree diagram is

$$\langle J/\psi K^0|\mathcal{H}_{\rm t}|B^0\rangle \;=\; V_{cb}^*V_{cs}A^{\rm t}, \tag{1.48}$$

and the amplitude of the penguin diagram is

$$\langle J/\psi K_S|\mathcal{H}_{\rm p}|B^0\rangle \;=\; V_{ub}^*V_{us}A_u^{\rm p} + V_{cb}^*V_{cs}A_c^{\rm p} + V_{tb}^*V_{ts}A_t^{\rm p}. \tag{1.49}$$

Here, $A^{\rm t}$ and $A_i^{\rm p}$ $(i = u, c, t)$ are the amplitudes apart from the explicitly shown CKM factors. Because $\sum_{i=u,c,t} V_{ib}^*V_{is}$ is zero due to the unitarity condition, following equation holds:

$$\langle J/\psi K_S|\mathcal{H}_p|B^0\rangle \;=\; V_{cb}^*V_{cs}(A_c^p - A_t^p) + V_{ub}^*V_{us}(A_u^p - A_t^p). \tag{1.50}$$

According to the Wolfenstein parameterization, the first term has the same weak phase as the tree diagram and the second term has different weak phase from the tree diagram. We can also read from the parameterization that $(V_{ub}^*V_{us})/(V_{cb}^*V_{cs})$ is order of $\mathcal{O}(\lambda^2)$. Thus, the second term is negligible with respect to the first term. Therefore, the penguin diagram possesses the same weak phase as the tree diagram up to very small correction. When there is only one amplitude (or more than but with the same weak phase), $|B^0 \to f_{CP}| = |\overline{B}^0 \to f_{CP}|$ holds. Thus, we can conclude

$$\big|A(B^0 \to J/\psi K_S)\big| = \big|A(\overline{B}^0 \to J/\psi K_S)\big|, \tag{1.51}$$

and

$$\frac{A(B^0 \to J/\psi K_S)}{A(\overline{B}^0 \to J/\psi K_S)} = \frac{V_{cb}V_{cs}^*}{V_{cb}^*V_{cs}}. \tag{1.52}$$

Then we calculate q/p. The dominating quark in the internal loop of the mixing diagrams is t because of its heaviest mass. Neglecting the contribution from u- and c-quarks, the

M_{12} is computed as [19]:

$$M_{12} \simeq \frac{G_{\mathrm{F}}^2 m_W^2}{4\pi^2} \cdot \eta_{\mathrm{QCD}}^{tt} \cdot (V_{td}^* V_{tb})^2 \cdot \left(\frac{m_t}{m_W}\right)^2 \cdot \langle B^0 | (\overline{d_{\mathrm{L}}} \gamma_\mu b_{\mathrm{L}})^2 | \overline{B}^0 \rangle. \tag{1.53}$$

On the other hand, the Γ_{12} is computed as [20]:

$$\Gamma_{12} \simeq \frac{3G_{\mathrm{F}}^2 m_W^2}{8\pi} \cdot (V_{td}^* V_{tb})^2 \cdot \left(\frac{m_b}{m_W}\right)^2 \cdot \langle B^0 | (\overline{d_{\mathrm{L}}} \gamma_\mu b_{\mathrm{L}})^2 | \overline{B}^0 \rangle, \tag{1.54}$$

From above equations, we have

$$\left| \frac{\Gamma_{12}}{M_{12}} \right| \simeq \frac{3\pi m_b^2}{2m_t^2} \sim \mathcal{O}(10^{-2}). \tag{1.55}$$

Using equations (1.25), (1.53), and (1.55), q/p is approximated as

$$\frac{q}{p} \simeq \sqrt{\frac{M_{12}^*}{M_{12}}} = \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} = \mathrm{e}^{-2\mathrm{i}\phi_1}. \tag{1.56}$$

We have to be aware of K^0-$\overline{K}^0$ mixing because we have K_S in the decay final state. Similar discussion as equations (1.26) is available for K^0-$\overline{K}^0$ mixing:

$$\langle K_S | = p_K^* \langle K^0 | - q_K^* \langle \overline{K}^0 |. \tag{1.57}$$

Writing the transition of $K^0 \to K_S$ explicitly, the amplitudes for $B^0 \to J/\psi K_S$ are expressed as

$$\begin{aligned} \langle J/\psi K_S | B^0 \rangle &= \langle K_S | K^0 \rangle \langle J/\psi K^0 | B^0 \rangle = p_K^* V_{cb}^* V_{cs} A, && (1.58) \\ \langle J/\psi K_S | \overline{B}^0 \rangle &= q_K^* V_{cb} V_{cs}^* A. && (1.59) \end{aligned}$$

Here, we used the facts of $CP|B^0\rangle = -|\overline{B}^0\rangle$, $CP|K^0\rangle = -|\overline{K}^0\rangle$, $CP|J/\psi\rangle = |J/\psi\rangle$, and $CP|J/\psi K^0\rangle = |J/\psi \overline{K}^0\rangle$, where the third equation is derived from $J^{CP} = 1^{--}$ for J/ψ, and the last equation is from angular momentum of the system being 1. Therefore,

$$\begin{aligned} \langle J/\psi K_S | \overline{B}^0 \rangle &= -q_K^* \langle J/\psi \overline{K}^0 | (CP)^{-1} (CP)\, \mathcal{H}\, (CP)^{-1} (CP) | \overline{B}^0 \rangle \\ &= q_K^* \langle J/\psi K^0 | \mathcal{H}^{CP} | B^0 \rangle. && (1.60) \end{aligned}$$

$(q/p)_K$ is obtained from the calculation of K^0-$\overline{K}^0$ mixing diagrams similar to Figure 1.5. In this case, using $|V_{cd}^* V_{cs}| \gg |V_{td}^* V_{ts}|$, we obtain the additional factor to $\lambda_{J/\psi K_S}$ as

$$\left(\frac{q}{p}\right)_K \simeq \frac{V_{cs} V_{cd}^*}{V_{cs}^* V_{cd}}. \tag{1.61}$$

Finally, we obtain

$$\begin{aligned}
\lambda_{J/\psi K_S} &= -\frac{A(\overline{B}^0 \to J/\psi K_S)}{A(B^0 \to J/\psi K_S)} \cdot \left(\frac{q}{p}\right) \cdot \left(\frac{q}{p}\right)_K \\
&\simeq -\frac{V_{cb}V_{cs}^*}{V_{cb}^*V_{cs}} \cdot \frac{V_{tb}^*V_{td}}{V_{tb}V_{td}^*} \cdot \frac{V_{cs}V_{cd}^*}{V_{cs}^*V_{cd}} \\
&= -\mathrm{e}^{-2\mathrm{i}\phi_1},
\end{aligned} \tag{1.62}$$

and

$$\mathcal{I}m(\lambda_{J/\psi K_S}) = \sin 2\phi_1. \tag{1.63}$$

The minus sign in front of equation (1.62) is due to the state of $|J/\psi K_S\rangle$ being CP odd. Because of $|B^0 \to J/\psi K_S| = |\overline{B}^0 \to J/\psi K_S|$ and $|q/p|$=1, $|\lambda_{J/\psi K_S}| = 1$ holds. As a result, equations (1.43), (1.44), and (1.46) become much simpler:

$$\Gamma_{B \to f_{CP}}(t) = |a|^2 \mathrm{e}^{-\Gamma t}(1 - \sin 2\phi_1 \cdot \sin \Delta M t), \tag{1.64}$$

$$\Gamma_{\overline{B} \to f_{CP}}(t) = |a|^2 \mathrm{e}^{-\Gamma t}(1 + \sin 2\phi_1 \cdot \sin \Delta M t), \tag{1.65}$$

$$a_{CP}(t) = \sin 2\phi_1 \cdot \sin \Delta M t. \tag{1.66}$$

1.4 Measurement of $\sin 2\phi_1$ at B-Factory

In the B-factory, B mesons are produced from $b\overline{b}$ resonance state of $\Upsilon(4S)$, because the $\Upsilon(4S)$ is the lowest bound state that can decay into two B mesons. In this system, we observe the CP violation in proper-time difference distribution. In the following subsections, first we describe the proper-time difference of two B meson decays produced from $\Upsilon(4S)$. Then we describe the overview to observe the CP violation at the B-factory. Experimental constraints and recent measurement of $\sin 2\phi_1$ are also described.

1.4.1 Proper-Time Difference

$\Upsilon(4S)$ decays into a coherent B^0-$\overline{B}^0$ state with a C odd configuration. Subsequently oscillations take place preserving the C odd configuration: Bose statistics tells us that if one of the mesons is a B^0 at some time, the other one cannot be B^0 at that time, since the state must be odd under exchange of two mesons. Let us consider the decay rate of such pair of B mesons. We label each of the B mesons with its momentum: $\vec{k}$ or $-\vec{k}$. Assuming $\Delta\Gamma = 0$, a state with $B_{\vec{k}}$ at time $t = t_{\vec{k}}$ and $B_{-\vec{k}}$ at time $t = t_{-\vec{k}}$ can be expressed as

$$
\begin{aligned}
|B_{\vec{k}}B_{-\vec{k}}(t_{\vec{k}},t_{-\vec{k}})\rangle &= \frac{1}{\sqrt{2}}\mathrm{e}^{-\frac{\Gamma}{2}(t_{\vec{k}}+t_{-\vec{k}})} \\
&\times \left[\mathrm{i}\sin\frac{\Delta M(t_{-\vec{k}}-t_{\vec{k}})}{2}\left(\frac{p}{q}|B^0_{\vec{k}}B^0_{-\vec{k}}\rangle - \frac{q}{p}|\overline{B}^0_{\vec{k}}\overline{B}^0_{-\vec{k}}\rangle\right)\right. \\
&\left.\quad + \cos\frac{\Delta M(t_{-\vec{k}}-t_{\vec{k}})}{2}\left(|B^0_{\vec{k}}\overline{B}^0_{-\vec{k}}\rangle - |\overline{B}^0_{\vec{k}}B^0_{-\vec{k}}\rangle\right)\right],
\end{aligned}
\tag{1.67}
$$

where $B^0_{\vec{k}}(B^0{}_{-\vec{k}})$ denotes $B_{\vec{k}}(B_{-\vec{k}})$ is B^0 and $\overline{B}^0_{\vec{k}}(\overline{B}^0{}_{-\vec{k}})$ denotes $B_{\vec{k}}(B_{-\vec{k}})$ is $\overline{B}^0$. This equation is obtained from the time evolution of

$$
\frac{1}{\sqrt{2}}\left(|B^0{}_{\vec{k}}(t)\rangle|\overline{B}^0{}_{-\vec{k}}(t)\rangle - |\overline{B}^0{}_{\vec{k}}(t)\rangle|B^0{}_{-\vec{k}}(t)\rangle\right), \tag{1.68}
$$

where the state is chosen to be anti-symmetric due to the C odd state of the system. The decay rates are computed for the case in which one of the particles decays in a "flavor-specific" way, while the other one decays in a "flavor-nonspecific" way, *e.g.*:

$$
\begin{array}{ccccc}
\overline{B}^0 & \rightarrow & \ell^- + X & \not\leftarrow & B^0, \\
\overline{B}^0 & \not\rightarrow & \ell^+ + \overline{X} & \leftarrow & B^0, \\
\overline{B}^0 & \rightarrow & f_{CP} & \leftarrow & B^0.
\end{array}
$$

Using the definitions of $\langle \ell^+ X|B^0\rangle = \langle \ell^-\overline{X}|\overline{B}^0\rangle \equiv A_{SL}$, $\langle f_{CP}|B^0\rangle \equiv a$, and $\langle f_{CP}|\overline{B}^0\rangle \equiv a' = \lambda_{CP}\cdot ap/q$, the decay rates can be calculated as

$$
\begin{aligned}
\left|\langle(\ell^-X)_{\vec{k}},(f_{CP})_{-\vec{k}}|B_{\vec{k}}B_{-\vec{k}}(t_{\vec{k}},t_{-\vec{k}})\rangle\right|^2 &= \frac{1}{4}\mathrm{e}^{-\Gamma(t_{\vec{k}}+t_{-\vec{k}})}|A_{SL}|^2|a|^2 \\
&\times \left[(1+|\lambda_{CP}|^2) + (1-|\lambda_{CP}|^2)\cos\Delta M(t_{-\vec{k}}-t_{\vec{k}})\right. \\
&\left.\quad - 2\mathcal{I}m(\lambda_{CP})\sin\Delta M(t_{-\vec{k}}-t_{\vec{k}})\right],
\end{aligned}
\tag{1.69}
$$

and

$$
\begin{aligned}
\left|\langle(\ell^+\overline{X})_{\vec{k}},(f_{CP})_{-\vec{k}}|B_{\vec{k}}B_{-\vec{k}}(t_{\vec{k}},t_{-\vec{k}})\rangle\right|^2 &= \frac{1}{4}\mathrm{e}^{-\Gamma(t_{\vec{k}}+t_{-\vec{k}})}|A_{SL}|^2|a|^2\left|\frac{p}{q}\right|^2 \\
&\times \left[(1+|\lambda_{CP}|^2) - (1-|\lambda_{CP}|^2)\cos\Delta M(t_{-\vec{k}}-t_{\vec{k}})\right. \\
&\left.\quad + 2\mathcal{I}m(\lambda_{CP})\sin\Delta M(t_{-\vec{k}}-t_{\vec{k}})\right].
\end{aligned}
\tag{1.70}
$$

Approximations of $|\lambda_{CP}| \simeq 1$, and $|q/p| \simeq 1$ for $B^0 \rightarrow J/\psi K_S$ simplify equations (1.69) and (1.70) into:

$$
G_{\ell^\pm}(t_{-\vec{k}},t_{\vec{k}}) \;\propto\; \mathrm{e}^{-\Gamma(t_{\vec{k}}+t_{-\vec{k}})}\left[1 \pm \sin 2\phi_1 \sin\Delta M(t_{-\vec{k}}-t_{\vec{k}})\right], \tag{1.71}
$$

where $G_{\ell^\pm} \equiv |\langle(\ell^\pm X)_{\vec{k}},(f_{CP})_{-\vec{k}}|B_{\vec{k}}B_{-\vec{k}}(t_{\vec{k}},t_{-\vec{k}})\rangle|^2$ and $\sin 2\phi_1 = \mathcal{I}m(\lambda_{CP})$.

Let $t_{-\vec{k}}$ as the time when one of two B meson decays into f_{CP} state, and $t_{\vec{k}}$ as the time when another B meson decays in flavor-specific way, such as semileptonic decay. We label the B meson decaying into CP eigenstate as B_{CP} and the remaining B meson as $B_{\rm tag}$, hereafter. A duration measurement from $B_{\rm tag}$ decay to B_{CP} decay, defined as $\Delta t \equiv t_{-\vec{k}} - t_{\vec{k}}$, provides $\sin 2\phi_1$ instead of the measurements of $t_{-\vec{k}}$ and $t_{\vec{k}}$, as follows. Because we care about neither individual decay times of $t_{\vec{k}}$ nor $t_{-\vec{k}}$, we have to integrate equation (1.71) with respect to $t_{-\vec{k}}$ and $t_{\vec{k}}$ under the constraint of $\Delta t = t_{-\vec{k}} - t_{\vec{k}}$. Using $t_{-\vec{k}} \geq 0$ and $t_{\vec{k}} \geq 0$, we obtain

$$\iint \mathrm{d}t_{-\vec{k}}\,\mathrm{d}t_{\vec{k}}\; G_{\ell^\pm}(t_{-\vec{k}}, t_{\vec{k}})\,\delta(t_{-\vec{k}} - t_{\vec{k}} - \Delta t) \;\propto\; \mathrm{e}^{-\Gamma|\Delta t|}\Big(1 \pm \sin 2\phi_1 \sin \Delta M \Delta t\Big). \quad (1.72)$$

It is worth mentioning what happens if we swap $\vec{k}$ and $-\vec{k}$ in equations (1.69) and (1.70). In this case, "flavor-specific" mode is associated to $B_{-\vec{k}}$, and "flavor-nonspecific" mode is associated to $B_{\vec{k}}$, and therefore the definition of Δt is flipped as $\Delta t \equiv t_{\vec{k}} - t_{-\vec{k}}$. Thus, we have exactly same functions as (1.69) and (1.70).

To summarize, in the C odd system, if we measure the proper-time difference and if we identify the flavor of $B_{\rm tag}$, we can extract $\sin 2\phi_1$ from the Δt distributions.

Calculating a normalization factor for equation (1.72), we obtain a Δt distribution function as

$$f_{\ell^\pm}(\Delta t;\ \sin 2\phi_1) = \frac{1}{2\tau_{B^0}}\mathrm{e}^{-\frac{|\Delta t|}{\tau_{B^0}}}\Big(1 \pm \sin 2\phi_1 \sin \Delta M \Delta t\Big), \quad (1.73)$$

where $\tau_{B^0} \equiv 1/\Gamma$ is a lifetime of B^0 meson. In general, ℓ^+ denotes that flavor-specific B meson was B^0 and ℓ^- denotes that flavor-specific B meson was $\overline{B}^0$. Figure 1.7 (a) shows the proper-time difference distributions for $f_{\ell^+}(\Delta t; \sin 2\phi_1)$ (solid) and $f_{\ell^-}(\Delta t; \sin 2\phi_1)$ (dotted). Inputs of $\sin 2\phi_1$, τ_{B^0}, and ΔM are 0.60, 1.548 ps, and 0.472 ps^{-1}, respectively.

1.4.2 Belle Experiment

In this subsection, we give introduction of the experimental procedure to determine $\sin 2\phi_1$ in the Belle experiment, which is the B-factory experiment at KEK.

In the KEK B-factory, a $B\overline{B}$ meson pair is produced by a decay of $\Upsilon(4S)$, where $\Upsilon(4S)$ is a $b\bar{b}$ resonance state possessing a minimum mass to decay B^0-$\overline{B}^0$ pair. One of the B meson pair is fully reconstructed to identify its decay final state. We reconstruct B meson with J/ψ and K_S mesons. J/ψ and K_S are constructed via $J/\psi \to \ell^+\ell^-$ ($\ell = e, \mu$) and $K_S \to \pi^+\pi^-$ decay, respectively. Efficient determination of the charged particles and

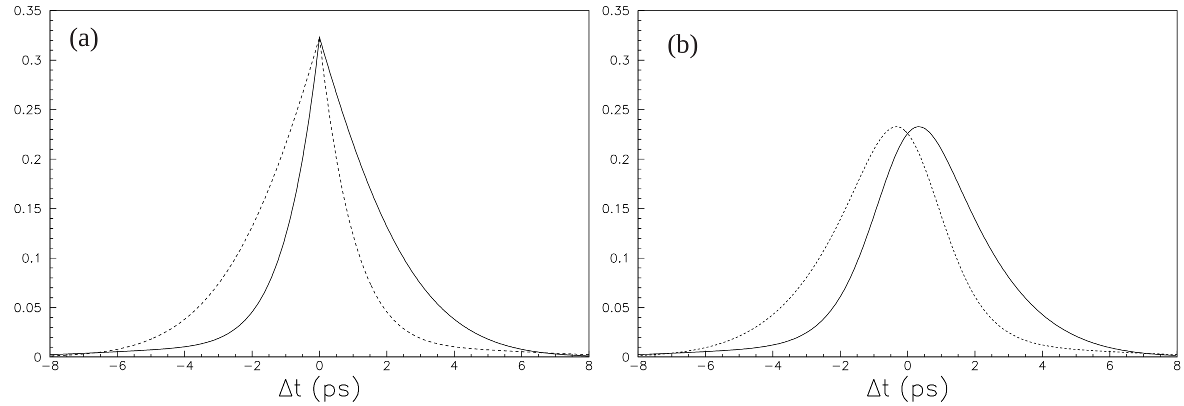


Figure 1.7: (a) Proper-time difference distribution for events with $B_{\rm tag}$ tagged by ℓ^+ (solid) and for events with $B_{\rm tag}$ tagged by ℓ^- (dotted). The input of $\sin 2\phi_1$ is 0.60. The diluted distribution of (a) by a single Gaussian resolution whose standard deviation is $\sigma = \tau_{B^0}/2$ (~ 100 μm) is shown in (b).

good lepton identification capability is essential.

It is necessary to determine flavor of the associated B meson. This process is called "flavor tagging". In addition to the semileptonic decays, the presence of the following particles can be used to tag the flavor of the B mesons: secondary lepton in $b \to c \to \ell^+$ decays, fast pions, which reflects the charge of virtual W in $b \to c + W^-$, slow pions coming from $D^{*\pm}$ whose charge reflects a charge of c, and kaons and Λ from cascade decays of $b \to c \to s$. When the probability to incorrectly assign the flavor of the associated B meson is w, (called wrong tag fraction), the observed Δt distribution becomes

$$\begin{aligned} f_{\ell^\pm}(\Delta t;\, \sin 2\phi_1) &= \frac{1}{2\tau_{B^0}} \mathrm{e}^{-\frac{|\Delta t|}{\tau_{B^0}}} \Big[1 \pm (1-w)\cdot \sin 2\phi_1 \sin \Delta M \Delta t\Big] \\ &\quad + \frac{1}{2\tau_{B^0}} \mathrm{e}^{-\frac{|\Delta t|}{\tau_{B^0}}} \Big[1 \mp w\cdot \sin 2\phi_1 \sin \Delta M \Delta t\Big] \\ &= \frac{1}{2\tau_{B^0}} \mathrm{e}^{-\frac{|\Delta t|}{\tau_{B^0}}} \Big[1 \pm (1-2w)\cdot \sin 2\phi_1 \sin \Delta M \Delta t\Big]. \end{aligned} \tag{1.74}$$

Good particle identification, in particular lepton and kaon identification, is required to minimize w.

The Δt is obtained from the distance of two decay vertices $B_{\rm tag}$ and B_{CP} (Δz) as

$$\Delta t = \frac{\Delta z}{c(\beta\gamma)_B}, \tag{1.75}$$

where $(\beta\gamma)_B$ is a motion of B meson. Because the mass of $\Upsilon(4S)$ is close to a sum of two B^0 masses, two B mesons are produced almost at rest in the $\Upsilon(4S)$ rest frame.

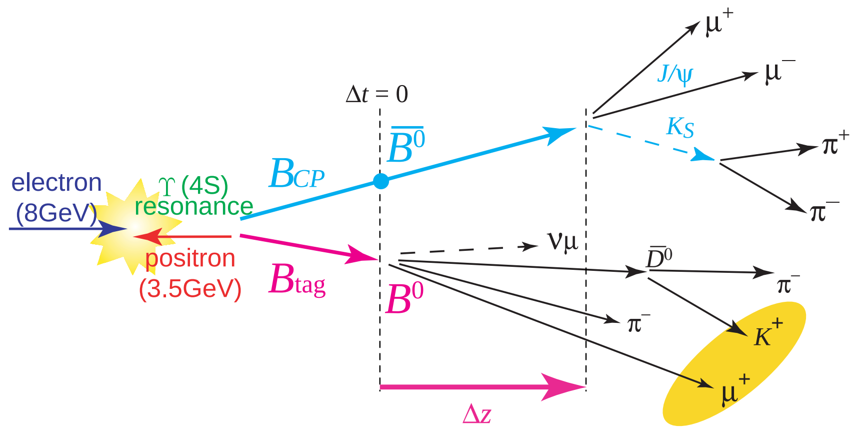


Figure 1.8: Schematic drawing of the vertex reconstruction of two B decay vertices. In the illustration, $B_{\rm tag}$ is assumed $\overline{B}^0$ and B_{CP} is assumed B^0 at the decay time of $B_{\rm tag}$ ($\Delta t = 0$).

Thus, we can approximately assume that $(\beta\gamma)_B$ is common for both B mesons. To make the Δz length sizable, we boost B meson by an asymmetric energy of the e^+e^- collider, 3.5 GeV for e^+ and 8.0 GeV for e^-. The produced $\Upsilon(4S)$ possesses $(\beta\gamma)_\Upsilon = 0.425$. Because B meson pair is produced at the rest frame of the $\Upsilon(4S)$, B mesons also possess $(\beta\gamma)_B = 0.425$. Since B meson lifetime is $1.5 - 1.7$ ps, the produced B mesons run about 200 μm before they decay, which are sizable length by the detector. The decay vertex of B_{CP} is reconstructed by leptons from J/ψ decay and the decay vertex of $B_{\rm tag}$ is reconstructed by all remaining tracks after the $J/\psi K_S$ reconstruction. Figure 1.8 shows the schematic drawing of the vertex reconstruction of two B decay vertices. As equation (1.73) states, accurate measurement of Δz is crucial issue because the integration of equation (1.73) with Δt vanishes $\sin 2\phi_1$. For the precise measurement of the B decay vertex, the Belle is equipped with the silicon vertex detector. Figure 1.7 (b) shows the asymmetry distribution of Δz smeared by Gaussian whose σ is $\tau_{B^0}/2$ (~ 100 μm). The distribution is changed drastically by the detector resolution. Thus, an appropriate understanding of the resolution is one of essential component to measure $\sin 2\phi_1$.

The wrong tagging probability varies event by event according to the $B_{\rm tag}$ decay products. The detector response also varies event by event by multiple scattering and energy loss of the tracks, and/or resolution of each hit on the detector, and so on. To take into account the event-by-event effect, we determine $\sin 2\phi_1$ from the asymmetric Δt distribution by the unbinned-maximum-likelihood method.

1.4.3 Constraints on $\sin 2\phi_1$ and Unitarity Triangle

The review of current experimental constraints is given in this subsection. Those measurements define the preferable area for ϕ_1 by specifying the apex of $(\tilde{\rho}, \tilde{\eta})$ in Figure 1.4 (b).

The entries in the first two rows of the CKM matrix are accessible in so-called direct (tree-level) processes, *i.e.* in weak decays of hadrons containing the corresponding quarks. $|V_{ud}|$ and $|V_{us}|$ are known to an accuracy of better than 1%, $|V_{cb}|$ is known to 5%, and $|V_{cd}|$ and $|V_{cs}|$ are known to about $10-20\%$. Hence, the two Wolfenstein parameters λ and A are rather well determined experimentally:

$$\lambda = |V_{us}| = 0.2205 \pm 0.0018, \qquad A = \left|\frac{V_{cb}}{V_{us}^2}\right| = 0.80 \pm 0.04.$$

On the other hand, $|V_{ub}|$ has an uncertainty of $\sim 30\%$, and the same is true for $|V_{td}|$, which is obtained from B^0-$\overline{B}^0$ mixing. This implies rather significant uncertainty in ρ and η. A more precise determination of these parameters will be done by the B-factory experiments.

To determine the shape of the triangle, one can aim for measurements of the two sides and three angles. So far, experimental information is available only on the sides of the triangle. Using the Wolfenstein parameterization and equation (1.18), the two sides of the unitarity triangle are expressed as

$$R_b \equiv \left|(0,0) \leftrightarrow (\tilde{\rho}, \tilde{\eta})\right| = \sqrt{\tilde{\rho}^2 + \tilde{\eta}^2} = \frac{1}{\lambda}\left(1 - \frac{\lambda^2}{2}\right)\frac{|V_{ub}|}{|V_{cb}|}, \tag{1.76}$$

$$R_t \equiv \left|(1,0) \leftrightarrow (\tilde{\rho}, \tilde{\eta})\right| = \sqrt{(1-\tilde{\rho})^2 + \tilde{\eta}^2} \simeq \frac{1}{A\lambda^3}|V_{td}V_{tb}^*|. \tag{1.77}$$

$|V_{td}V_{tb}^*|$ is accessible through B^0-$\overline{B}^0$ mixing shown in Figure 1.5 by the measurement of mass difference, ΔM. The theoretical prediction by the Standard Model is [19]

$$|V_{td}V_{tb}^*|^2 = \frac{6\pi^2 \Delta M}{G_{\mathrm{F}}^2 m_W^2 \eta_{\mathrm{QCD}}^{tt} B_{B_d} f_{B_d}^2 m_{B_d} S(m_t/m_W)}, \tag{1.78}$$

where η_{QCD}^{tt} is known as QCD correction factor measured to be $\eta_{\mathrm{QCD}}^{tt} = 0.55 \pm 0.01$ [21], and $S(m_t/m_W)$ is a function of the top quark mass. The product, $B_{B_d} f_{B_d}^2$, parameterize the hadronic matrix element of a local four-quark operator between B_d-meson states. Another way to improve the determination of R_t is through a measurement of B_s^0-$\overline{B}_s^0$ mixing,

$$R_t^2 = \frac{f_{B_s}^2 B_{B_s}}{f_{B_d}^2 B_{B_s}} \cdot \frac{m_{B_s}}{m_{B_d}} \cdot \frac{\Delta M_{B_s}}{\Delta M} \cdot \frac{1 - \lambda^2(1-2\tilde{\rho})}{\lambda^2}. \tag{1.79}$$

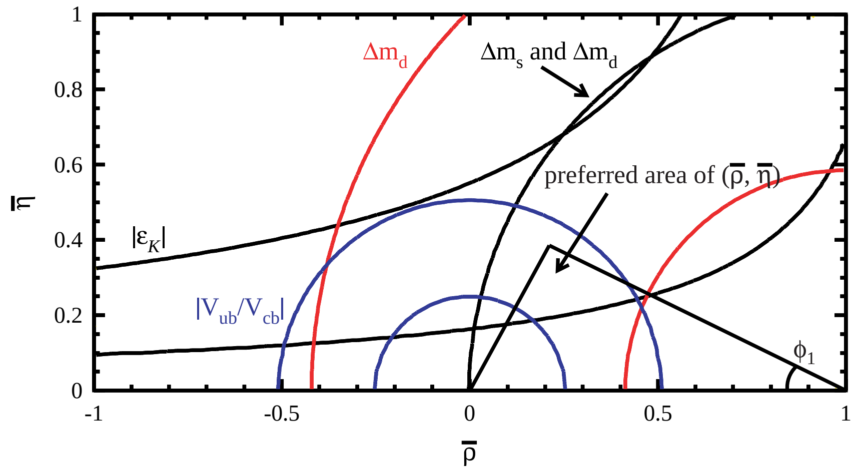


Figure 1.9: Constraints on the apex of the unitarity triangle, $(\tilde{\rho}, \tilde{\eta})$. The superimposed triangle is an example when the apex $(\tilde{\rho}, \tilde{\eta})$ comes within the preferred region.

The advantage of this way to the one from ΔM alone is that the elimination of dependence on m_t and the ratio of f_{B_s}/f_{B_d} can be more precisely determined than each decay constant itself. Presently only a lower limit on ΔM_{B_s} is obtained, and thus it gives upper limit of R_t.

Another constraint is given by K^0-$\overline{K}^0$ mixing parameter ϵ_K. The constraint arising in the $\tilde{\rho}$-$\tilde{\eta}$ plane forms hyperbola, depending on a hadronic parameter B_K.

Figure 1.9 shows the constraint of the apex, $(\tilde{\rho}, \tilde{\eta})$, of the unitarity triangle. An example triangle whose apex is within the area is also superimposed. The preferred region for ϕ_1 is

$$0.47 < \sin 2\phi_1 < 0.89,$$

at 95% confidence level [22], or $\sin 2\phi_1 = 0.70 \pm 0.07$ at 68% confidence level [23].

1.4.4 Previous Measurements of $\sin 2\phi_1$

The first direct measurement of $\sin 2\phi_1$ was presented by the OPAL collaboration in 1998. They selected 24 candidates of $B^0 \to J/\psi K_S$ decay with a purity of $\sim 60\%$ from 4.4×10^6 hadronic Z^0 decays. They obtained [24]

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

The ALEPH collaboration recorded 23 candidates of $B^0 \to J/\psi K_S$ from 4.2×10^6 hadronic Z^0 decays with estimated purity of 71%. The result was [25]

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

The CDF collaboration also reported $\sin 2\phi_1$ value. In 110 pb^{-1} of proton antiproton collision at $\sqrt{s} = 1.8$ TeV, they accumulated 395 ± 31 events of $B^0 \to J/\psi K_S$ candidates, with a signal-to-noise ratio (S/N) of 0.7. Their conclusion was [26]

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

Recently the CDF collaboration updated their result with improvement of analysis technique [27]. The result is: $\sin 2\phi_1(\mathrm{CDF}) = 0.91 \pm 0.32$ (stat) $\pm\, 0.18$(syst), but this is still preliminary.

Chapter 2

Experimental Apparatus

We describe the experimental apparatus for the observation of the CP violation in this chapter. The accelerator, the detector including the trigger system and the data acquisition system, and the computing environment are described in this order.

2.1 KEKB Accelerator

The KEKB is an asymmetric energy collider of 8.0 GeV electron beam and 3.5 GeV positron beam, to produce Lorentz boosted B meson pairs. The center-of-mass energy is 10.5 GeV, where $\Upsilon(4S)$ resides, and $\Upsilon(4S)$ decays into two B mesons, $B^0\overline{B}^0$ or B^+B^- in motion with $(\beta\gamma)_\Upsilon = 0.425$. Created B mesons run about 200 μm before they decay, because their lifetimes are $1.5 - 1.7$ ps. The design luminosity of the accelerator is 10^{34} cm^{-2}s^{-1}.

Figure 2.1 is an illustration of the KEKB accelerator. Electrons and positrons are injected from a linear accelerator (LINAC) into two main rings at Fuji area. The storage rings of electrons (HER) and positrons (LER) are about 3 km long in circumference. They are installed in existent TRISTAN tunnel. The traveled electrons and positrons interact with each other at only one interaction point, Tsukuba area. The crossing angle of two beams at the interaction point is ± 11 mrad in order to avoid parasitic collisions near the interaction point. The Belle detector is installed surrounding this interaction point.

The luminosity (L cm^{-2}s^{-1}) for a flat beam can be expressed as

$$L = 2.2 \times 10^{34}\, \xi\, (1 + r) \left(\frac{EI}{\beta_y^*} \right), \tag{2.1}$$

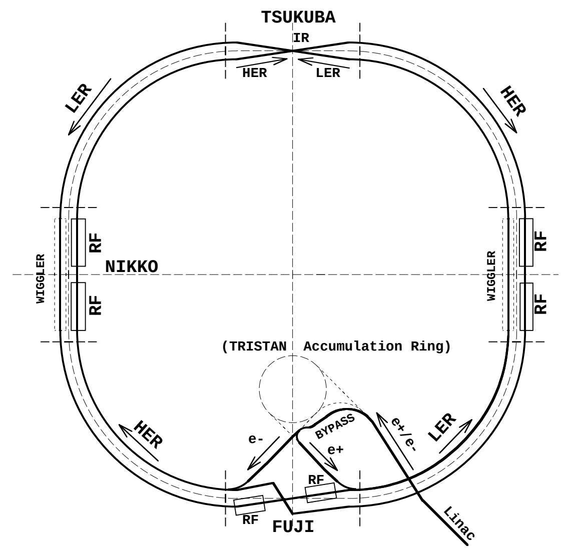


Figure 2.1: KEKB accelerator. Electrons go through high-energy ring (HER) in clock wise, and positrons go through low-energy ring (LER) in anti-clock wise.

where

$$
\begin{aligned}
\xi &= \text{beam-beam tune shift parameter,} \\
& \text{(this parameter reflects the strength of the beam-beam collision),} \\
r &= \text{aspect ratio of the beam shape (1 for a round and 0 for a flat beam),} \\
E &= \text{beam energy in GeV,} \\
I &= \text{circulating current in A, and} \\
\beta_y^* &= \text{vertical beta function at interaction point in cm.}
\end{aligned}
$$

At the design luminosity of 10^{34} cm^{-2}s^{-1}, which corresponds to 10^8 $\Upsilon(4S)$ yield per 10^7 s, 5000 bunches will be injected in each storage ring, where the bunch interval is 2 ns, *i.e.* 60 cm in space. The main parameters of the KEKB are listed in Table 2.1.

The commissioning of the KEKB accelerator started in December 1998. The first beam collision was observed in February, 1999 without the Belle detector at interaction region. The Belle detector was installed at the interaction region in May, 1999. Although a present performance is not good as the design, the achieved performance is highest in the world: the peak luminosity has reached 5.47×10^{33} cm^{-2}s^{-1}, the daily integrated luminosity usually exceeds 200 pb^{-1} (our record is 288 pb^{-1}), and monthly integrated luminosity is recorded to be 6.1 fb^{-1}.

Table 2.1: Main parameters of the KEKB. (†: without wiggles, ‡: with wiggles)

Ring	LER	HER	Units
Particle	e^+	e^-	
Energy (E)	3.5	8.0	GeV
Circumference (C)	3016.26		m
Luminosity ($\mathcal{L}$)	1×10^{34}		cm^{-2}s^{-1}
Crossing angle (θ_x)	$+11$		mrad
Tune shifts (ξ_x/ξ_y)	0.039/0.052		
Beta function at IP (β_x^*/β_y^*)	0.33/0.01		m
Beam current (I)	2.6	1.1	A
Natural bunch length (σ_z)	0.4		cm
Energy spread (σ_E/E)	7.1×10^{-4}	6.7×10^{-4}	
Bunch spacing (s_B)	0.59		m
Particles/bunch (σ_E/E)	3.3×10^{10}	1.4×10^{10}	
Emittance ($\varepsilon_x/\varepsilon_y$)	$1.8 \times 10^{-8}/3.6 \times 10^{-10}$		m
Synchrotron tune (ν_s)	$0.01 - 0.02$		
Betatron tune (ν_x/ν_y)	45.52/45.08	47.52/43.08	
Momentum compaction factor (α_p)	$1 \times 10^{-4} - 2 \times 10^{-4}$		
Energy loss/turn (U_0)	$0.81^\dagger/1.5^\ddagger$	4.8	MeV
RF voltage (V_c)	$5 - 10$	$10 - 20$	MV
RF frequency (f_{RF})	508.887		MHz
Harmonic number (h)	5120		
Longitudinal dumping time (τ_ε)	$43^\dagger/23^\ddagger$	23	ms
Total beam power (P_b)	$2.7^\dagger/4.5^\ddagger$	4.0	MW
Radiation power (P_{SR})	$2.1^\dagger/4.0^\ddagger$	3.8	MW
HOM power (P_{HOM})	0.57	0.15	MW
Bending radius (ρ)	16.3	104.5	m
Length of bending magnet (L_b)	0.915	5.86	m

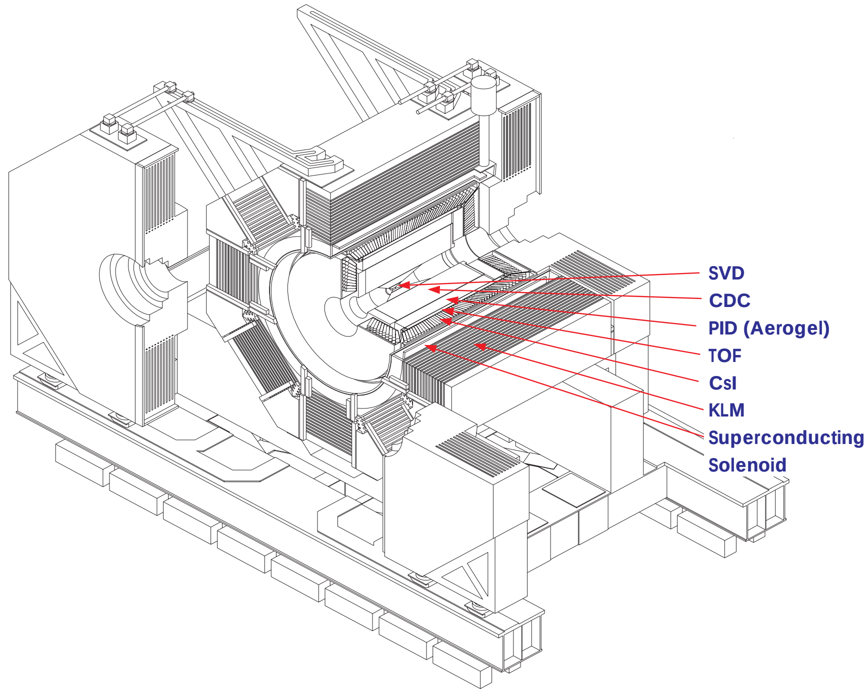


Figure 2.2: Overview of the Belle detector.

2.2 Belle Detector

The Belle detector is installed at the interaction point of electron-positron beams in Tsukuba experimental hall. Figures 2.2 and 2.3 show the cut-off and the side views of the Belle detector.

Because of the asymmetric beam energy, the detector is configured to be asymmetric. It means the detector acceptance is larger in the direction of electrons, which is defined as “forward” direction. The detector is configured with 1.5 T super-conducting solenoid and iron structure surrounding the beams. B meson decay vertices are measured by a silicon vertex detector (SVD) [28] situated just outside a beryllium beam pipe. Charged particle reconstruction is provided by a wire drift chamber, called central drift chamber (CDC) [29]. Particle identification is provided by $\mathrm{d}E/\mathrm{d}x$ measurements in CDC, the aerogel Cherenkov counter (ACC) [30], and time of flight counters (TOF) [31]. Electromagnetic showers are detected in an array of electromagnetic calorimeter (ECL) [32] made of thallium doped CsI located inside the solenoid coil. Muons and K_L mesons are detected by resistive plate counters (KLM) [33] interspersed in the iron return yoke of the magnet.

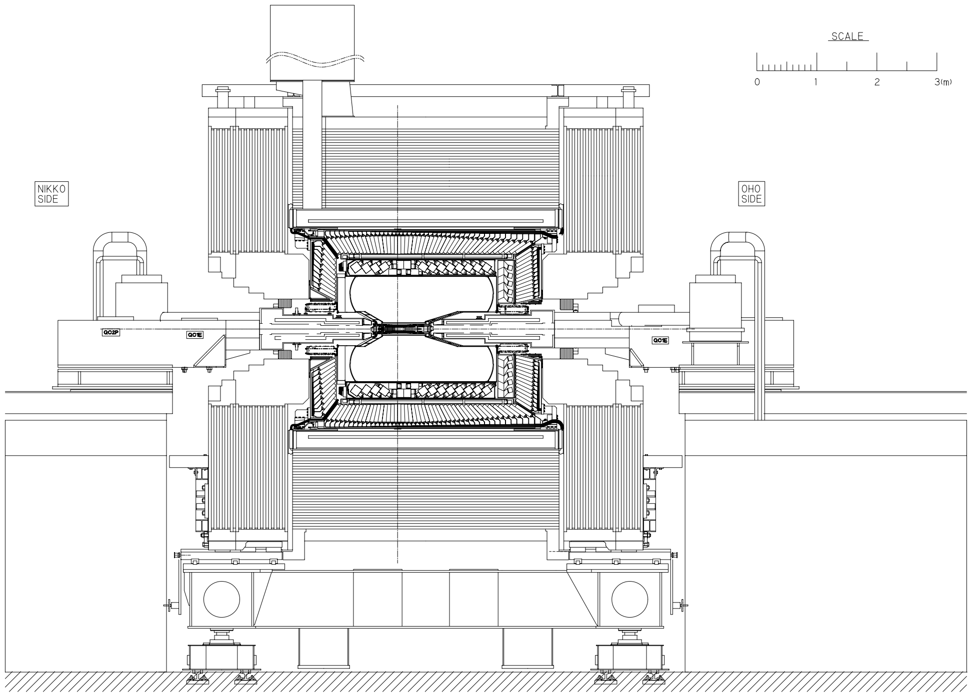


Figure 2.3: Side view of the Belle detector.

The coordinate system of the Belle detector is defined a

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

The performance of each detector is summarized in Table 2.2. In the following subsections, we give a brief description of each detector together with the beam pipe and the iron yoke.

2.2.1 Beam Pipe

Figure 2.4 shows the cross sections of the beam pipe at the interaction region. The beam pipe consists of two cylinders with different radii, the inner one is $r = 20.0$ mm and the outer one is 23.0 mm. Each cylinder have 0.5 mm thickness. The space within two cylinders is filled with chiller-helium gas. The beam pipe is made of beryllium to reduce the multiple Coulomb scattering at the beam-pipe wall.

Table 2.2: Summary of performance parameters of the Belle detector. p and p_T in GeV/c and E in GeV.

Detector	Type	Configuration	Readout	Performance
Beam pipe	Beryllium double-wall	Cylindrical, $r = 20$ mm 0.5 mm (Be) / 2.5 mm (He) / 0.5 mm (Be)		Helium gas chilled
SVD	Double sided Si strip	300 μm-thick, 3 layers $r = 3.00 - 6.05$ cm length $= 22 - 34$ cm	ϕ : 40.96 k z : 40.96 k	$\sigma_z \sim 200\ \mu$m
CDC	Small cell drift chamber	Anode: 50 layers Cathode: 3 layers $r = 8 - 88$ cm, $-79 \leq z \leq 160$ cm	A: 8.4 k C: 1.8 k	$\sigma_{r\text{-}\phi} = 130\ \mu$m $\sigma_z = 200 - 1400\ \mu$m $\sigma_{p_\mathrm{t}}/p_\mathrm{t} = (0.20 p_\mathrm{t} \oplus 0.29)\%$ $\sigma_{\mathrm{d}E/\mathrm{d}x} = 7\%$
ACC	$n = 1.01 - 1.03$ silica aerogel	$\sim 12 \times 12 \times 12$ cm^3 block 960 barrel, 228 end-cap FM-PMT readout	1788	$\mu_\mathrm{eff} \geq 6$ $\pi^\pm/K^\pm : 1.2 < p < 3.5$
TOF	Scintillator	128 ϕ segments $r = 120$ cm, 3 m long	128×2	$\sigma_t = 100$ ps $\pi^\pm/K^\pm : p < 1.2$
TSC	Scintillator	Attached to TOF	64	
ECL	Thallium doped CsI	Towered structure $\sim 5.5 \times 5.5 \times 30$ cm^3 crystals $r = 125 - 162$ cm, $z = -102$ and 196 cm	r : 6624 z_f : 1152 z_b : 960	$\sigma_E/E = (0.066/E$ $\oplus 0.81/E^{1/4} \oplus 1.34)\%$ $\sigma_\mathrm{pos} = 0.27 + 3.4/\sqrt{E}$ $+1.8/E^{1/4}$ mm
Magnet	Super conducting	inner radius = 170 cm		$B = 1.5$ T
KLM	Resistive plate counter	14 layers 5 cm Fe + 4 cm gap 2 RPCs in each gap θ and ϕ strips	θ : 16 k ϕ : 16 k	$\Delta\phi = \Delta\theta$ $= 30$ mrad for K_L σ_t is a few ns
EFC	$Bi_4Ge_3O_{12}$	$2 \times 1.5 \times 12$ cm^3	θ : 5 ϕ : 32	R.m.s. energy resolution 7.3% at 8.0 GeV 5.8% at 3.5 GeV

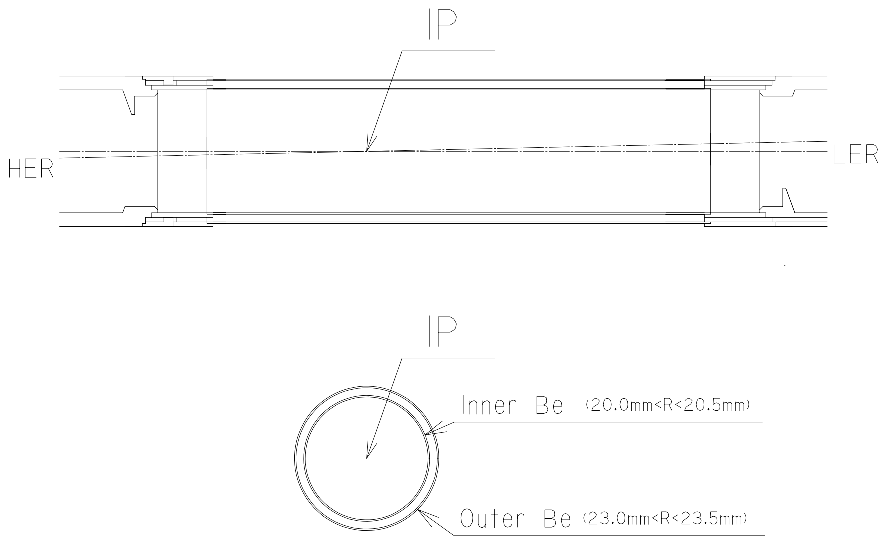


Figure 2.4: Cross section of the beryllium beam pipe at the interaction point.

2.2.2 Silicon Vertex Detector – SVD

It is crucially important to measure the flight length of produced B mesons in the z direction to study the CP violation. The SVD provides information necessary for the precise reconstruction of the decay vertices, where the Δz resolution is required less than $\sim$ 200 μm. In addition, the vertex detector is useful for the identification and measurements of the decay vertices of charmed mesons and tau leptons. The SVD also plays an important role for the track reconstruction. The layout of the SVD is shown in Figure 2.5. The SVD comprises three cylindrical detection layers of "ladders", where a ladder is composed of two, three, or four double-sided silicon strip detectors (DSSDs). The polar angle coverage is from 23° to 140°. The layers consist of 8, 10, and 14 ladders, respectively. The radii of each layer are 3.0 cm, 4.55 cm, and 6.05 cm. In order to achieve a vertex resolution better than 200 μm, good alignment is required. To help accomplish good internal alignment, the SVD is designed to include some overlap regions for DSSDs in adjacent ladders. Each ladder in turn is made from two "half ladders". In the innermost layer, each half ladder has one DSSD. In the second layer, one half ladder has two DSSDs and another half ladder has one DSSD. In the outermost layer all half ladders have two DSSDs. Figure 2.6 shows a schematic view of the DSSD. The DSSD has a size of 57.5×33.5 mm^2 with thickness of 300 μm. One side (n-side) of the DSSD has n^+-strips oriented perpendicular to the beam direction to measure the z coordinate and the other side (p-side) with longitudinal p^+-strips, allows the ϕ coordinate measurement. The bias voltage of 75 V is supplied to the n-side, while grounding the p-side. A charged

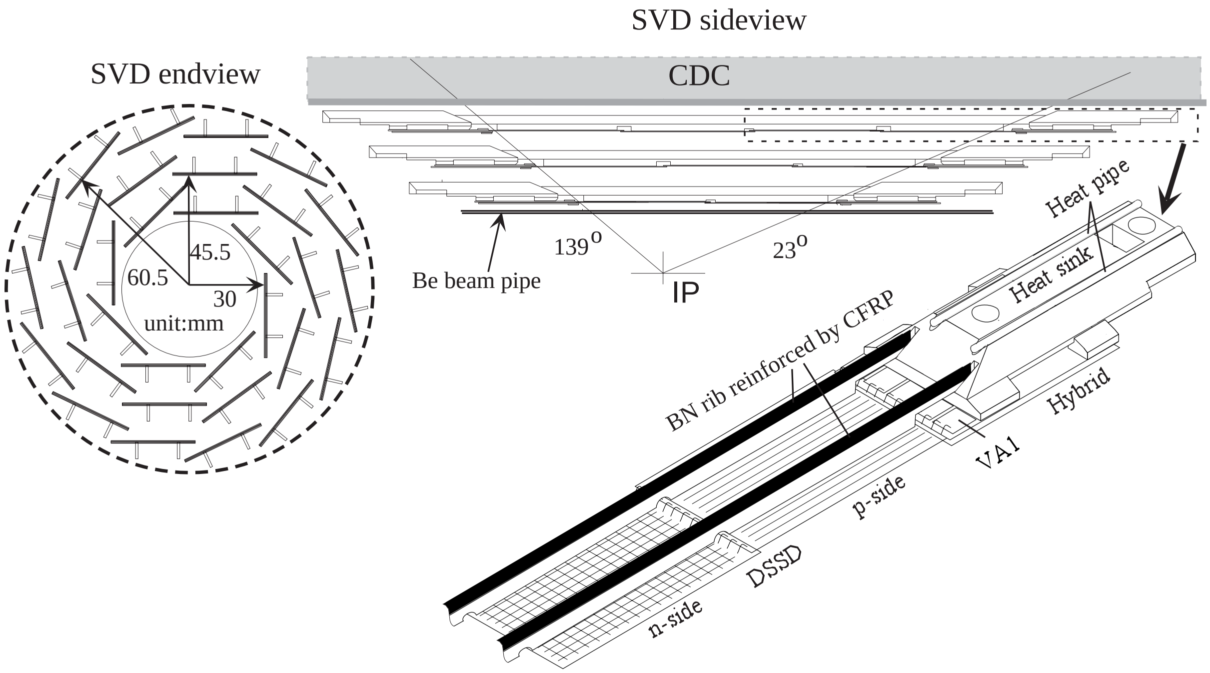


Figure 2.5: Detector configuration of the SVD.

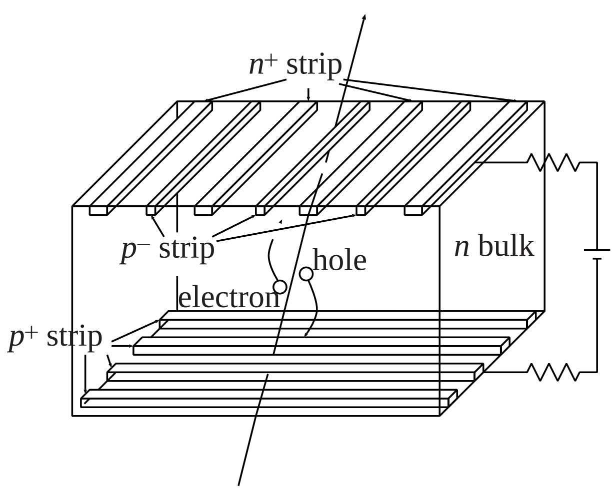


Figure 2.6: Schematic drawing of the DSSD.

particle traversing the depletion region of the n-bulk silicon generates an electron-hole pairs, which drift to each strip and make hit signals. Aluminum readout electrodes for the charge probing on the strips are AC-coupled to the strips with high-resistive poly-silicon. Induced charges on the electrodes by the strip charge are readout as hit signals. The DSSD strip pitch for the p-side (n-side) is 25 (42) μm and the readout pitch is 50 (84) μm. The total number of readout channel is 81920. The front-end readout is performed by 128 channel VA1 chips [34] that include pre-amplifiers, shapers and track-and-hold circuits. The scanned analog signals are transferred to the electronics hut by repeater modules, then digitized by flash ADC modules, where event buffering and zero suppression are performed together with the A/D conversion.

Good signal-to-noise ratio and high strip yields are needed to ensure the efficient matching between tracks detected by the CDC and the signal clusters detected by the SVD. Typical S/N ratio is measured to be 47 for p-strip sides and 19 for n strip sides. The strip yields that are defined as the fraction of channels with $S/N > 10$ are measured to be 98.8%, 96.3%, and 93.5% for innermost, second, and outermost layers, respectively. To evaluate the SVD performance, the matching efficiency between tracks found in the CDC and the SVD, and the impact parameter resolution are measured. The matching efficiency is 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 the r-ϕ and r-z information. Tracks from $K_S \to \pi^+\pi^-$ decays are excluded from this calculation because these tracks do not necessary go through the SVD. Figure 2.7 shows the SVD-CDC track matching efficiency for hadronic events as a function of the date of data taking. The average matching efficiency is measured to be better than 98.7%, although we observe slight degradation after one year operation as a result of the gain loss of VA1 from radiation damage. The momentum and angular dependence of the impact parameter resolution are shown in Figure 2.8 and well represented by following formula: $(19 \oplus 50/p\beta \sin^{3/2}\theta)$ μm for the resolution in the r-ϕ plane and $(36 \oplus 42/p\beta \sin^{5/2}\theta)$ μm for the resolution in the z direction, respectively, where "$\oplus$" indicates a quadratic sum.

2.2.3 Central Drift Chamber – CDC

The primary role of the CDC is the detection of charged particle tracks and the reconstruction of their momenta. The magnetic field in the CDC bends charged particle according to its momentum. The CDC tracks its trajectory and determine its momentum. The physics goals of the experiment require a momentum resolution of $\sigma_{p_\mathrm{t}}/p_\mathrm{t} \sim 0.5\%\sqrt{1+p_\mathrm{t}^2}$ (p_t in GeV/c) for charged particles with $p_\mathrm{t} > 100$ MeV/c in the polar angle region of $17^\circ \leq \theta \leq 150^\circ$. The CDC is also involved in the particle identification using the measurements of ionization loss of a charged particle ($\mathrm{d}E/\mathrm{d}x$). Amount of $\mathrm{d}E/\mathrm{d}x$ depends on $\beta = v/c$ (Bethe-Bloch formula), which enables us to identify the particle species.

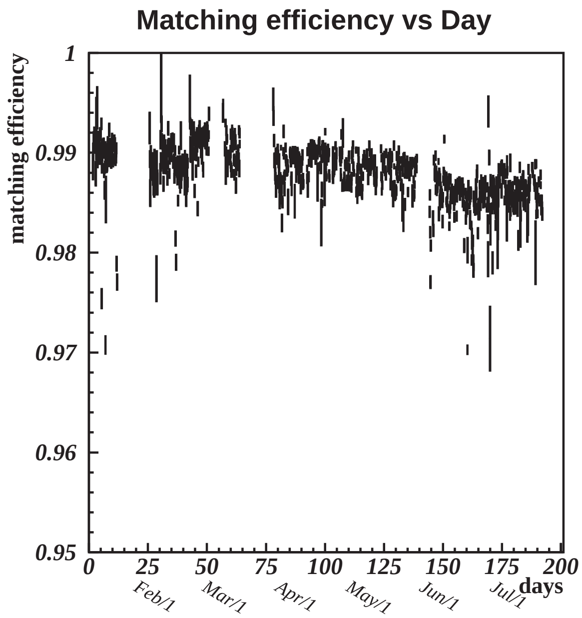


Figure 2.7: SVD-CDC track matching efficiency as a function of the data of the data taking.

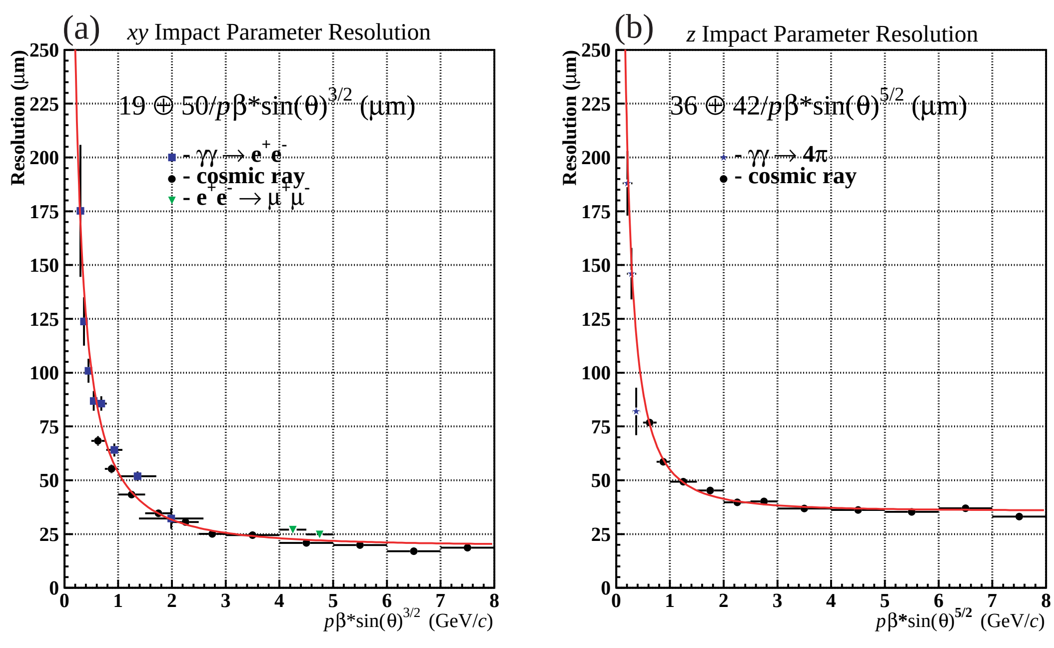


Figure 2.8: Impact parameter resolution by the SVD (a) in the r-ϕ plane and (b) in the z direction.

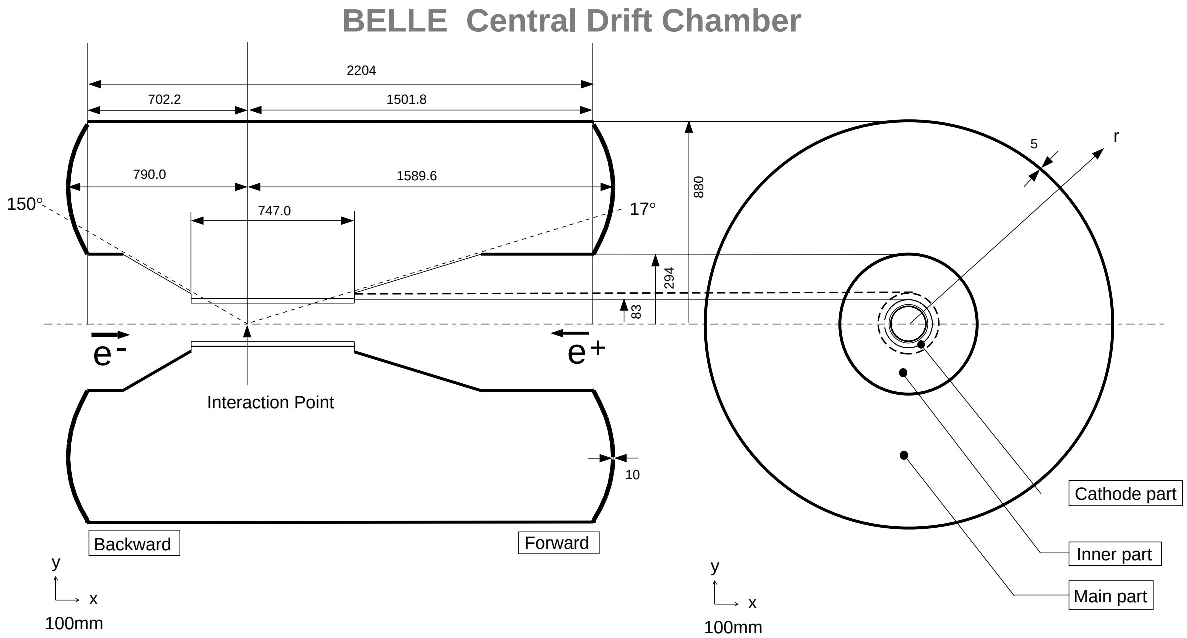


Figure 2.9: Overview of the CDC.

The configuration of the CDC is shown in Figure 2.9. The polar angle coverage of the CDC is from 17^{o} to 150^{o}. The inner and outer radii of the CDC are 8 cm and 88 cm, respectively. The detector coverage in the z direction is from -702.2 mm to 1502.2 mm. The CDC is a small-cell drift chamber containing a total of 50 cylindrical sense-wire layers, with 32 axial-wire layers and 18 stereo-wire layers. The axial wires are configured to be parallel to z axis, while the stereo wires are slanted approximately $\pm - 50$ mrad. The stereo angle enables us to reconstruct the particle trajectory three-dimensionally. The CDC has 8400 drift cells, which are rectangle ($16 \times 17\ \mu\text{m}^2$). One cell consists of one sense wire and eight electric field wires. The sense wires are gilded tungsten of 30 μm in diameter to maximize the drift-electric-field, and the field wires are unplated aluminum of 126 μm in diameter to reduce the material of the chamber. The field wires produce electric fields up to the edge of the cell. The electric field strength at the surface of the aluminum field wires is configured to be less than 20 kV/cm to avoid radiation damage. A mixture of He (50%) and C_2H_6 (50%) is filled in the chamber. The use of the helium minimizes the multiple Coulomb scattering contributing to the momentum resolution. The radiation length and the drift velocity of the He-C_2H_6 mixture are about 640 m and 4 cm/μs, respectively. A charged particle traversing in the CDC produces ionized gas ($\sim$ 100/cm). A charge avalanche is caused by the ionized gas and it drifts to the sense wire with a finite drift-time, then a signal is measured. In the innermost radii, three cathode strip layers are installed to provide the z position measurements of tracks for

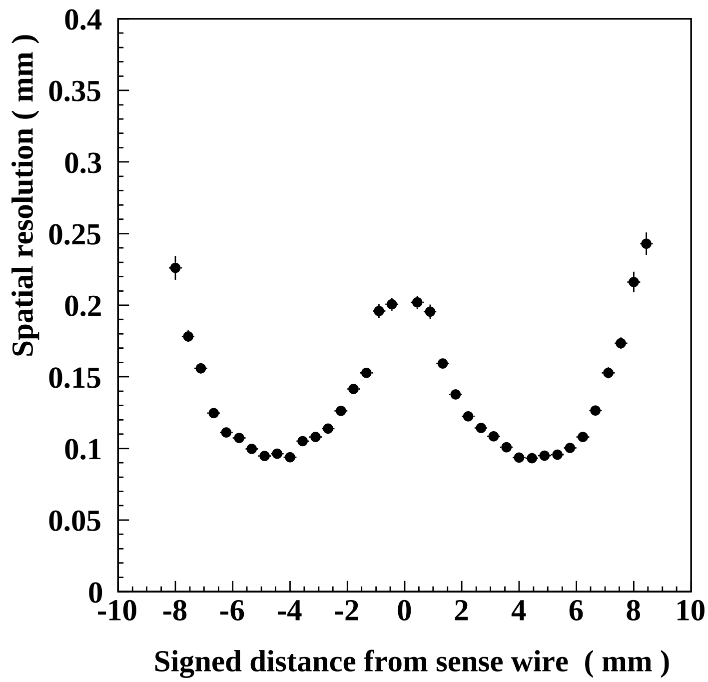


Figure 2.10: CDC spatial resolution as a function of drift distance.

trigger system.

Figure 2.10 shows the spatial resolution as a function of the drift distance. The position resolution is approximately $\sigma_{r\text{-}\phi} = 130\ \mu\text{m}$ in average. The transverse momentum resolution $\sigma_{p_\mathrm{t}}/p_\mathrm{t}$ is $(0.20 p_\mathrm{t} \oplus 0.29)\%$ where p_t is the transverse momentum measured in GeV/c. Figure 2.11 shows the p_t resolution as a function of p_t, with the fitted curve to the resolution function form.

The CDC is involved in the particle identification for the tracks with $p < 0.8$ GeV/c and $p > 2.0$ GeV/c measuring $\mathrm{d}E/\mathrm{d}x$. A scatter plot of the measured $\langle \mathrm{d}E/\mathrm{d}x \rangle$ and the particle momentum is shown in Figure 2.12, together with the expected mean-energy-losses for different particles. The resolution of $\mathrm{d}E/\mathrm{d}x$ is $\sigma_{\mathrm{d}E/\mathrm{d}x} = 6.9\%$ for minimum ionizing pions from K_S decays, as shown in Figure 2.13.

Charged Particle Reconstruction

The charged particle reconstruction is initiated by finding of track-segment-hit-patterns in the CDC. The tracks projected onto the r-ϕ plane are searched for with axial wires, then the hits of the stereo wires are used to determine z positions of the track. Track parameters (momentum and position) of the found track are fitted based on the Kalman filtering technique [35], which minimizes the effects of the multiple Coulomb scattering and non-uniformity of the magnetic field in the CDC in the determination of the track parameters. Then, all of the hit points are connected and fitted to a helix to obtain the

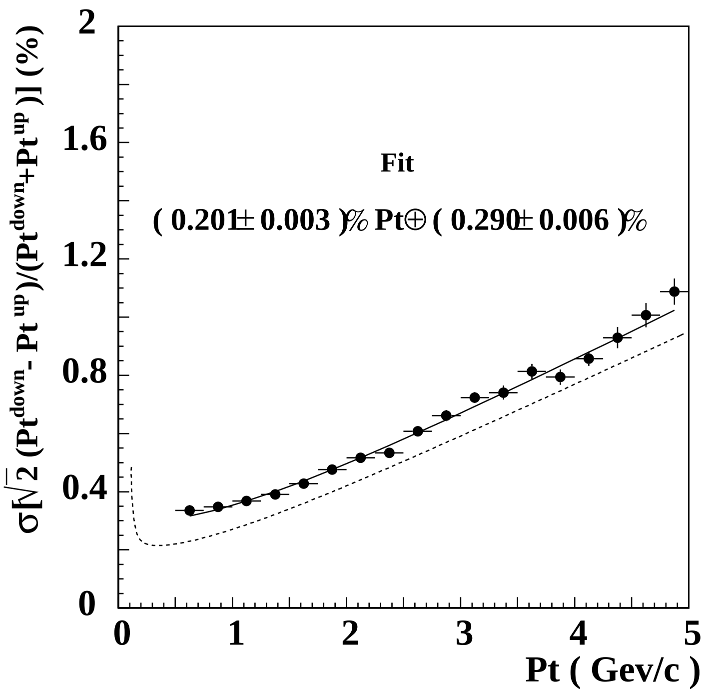


Figure 2.11: p_{t} resolution as a function of p_{t} itself. The solid curve shows the fitted result and the broken curve represents an ideal expectation assuming for $\beta = 1$ particles.

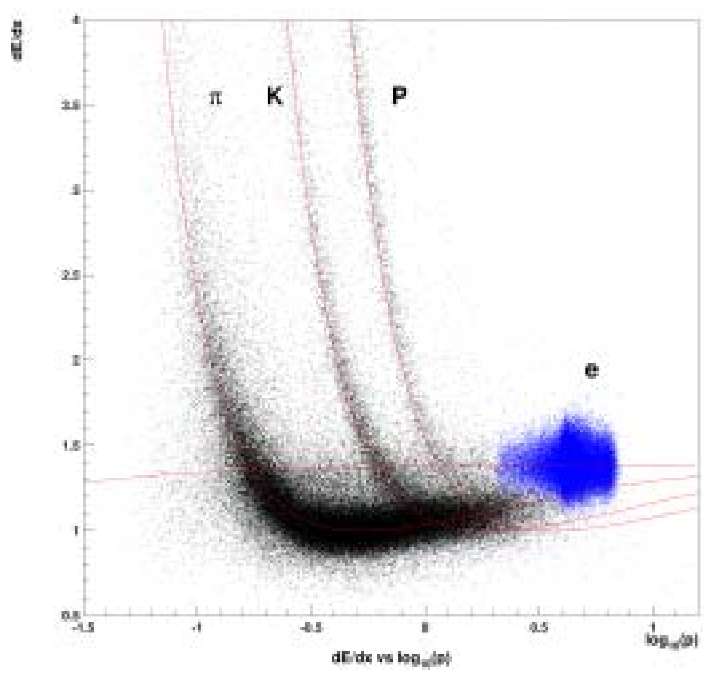


Figure 2.12: Measured dE/dx as a function of particle momentum. Curves represent the expected energy losses for each particle species.

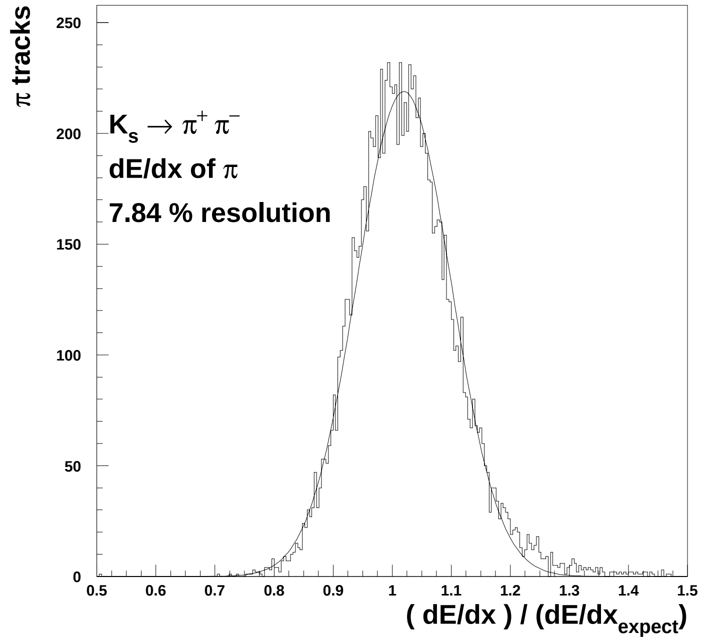


Figure 2.13: Distribution for measured dE/dx divided by expected dE/dx for pions from K_S decays.

particle momentum and position. Finally, the reconstructed charged particle trajectory is extrapolated toward the SVD to be connected to the SVD hits to improve the resolution of the track parameters. The track parameters are computed again with the combination of the hits on the CDC and the SVD.

The track momenta are calibrated by a scaling constant so that reconstructed invariant masses of J/ψ and D^0 become consistent with the world averages. In the calibration procedure of the track momenta, first we construct J/ψ mass from $J/\psi \to \mu^+\mu^-$ decay and D^0 mass from $D^0 \to K^-\pi^+$ decay. The means of the reconstructed mass distributions are compared to the world averages and then track momenta are tuned to reproduce the world averages by the reconstructed masses. The calibrations are made according to each change of the detector configuration. The major source of a fluctuation in the calibration constant comes from the fluctuation of a current supplied to the solenoid magnet. The amount of the correction is $\mathcal{O}(10^{-3})$. The estimated errors of track positions and momenta are also calibrated by a scaling constant. A cosmic ray that penetrates the detector is recognized by the track finding algorithm as two individual tracks. The “two” tracks should possess same track parameters, and the difference of the parameters between the “two” tracks must be Gaussian distribution. The scaling constants are introduced so that the standard deviation of this Gaussian should be equivalent to the error of the track parameters estimated by the Kalman filtering. Typical amounts of the corrections by the scaling constants are $10 - 15\%$.

2.2.4 Aerogel Cherenkov Counter – ACC

The ACC provides fine $\pi^\pm/K^\pm$ separation for the momentum range of $1.2 < p < 3.5$ GeV/c by detection of the Cherenkov light from particle penetrating through silica aerogel radiator. The light emission condition is represented as

$$m < p \cdot \sqrt{n^2 - 1}, \tag{2.2}$$

where m and p are particle mass and momentum, and n is a refractive index. Light mesons such as pions fire the ACC while heavier mesons such as kaons do not, which is the basic concept for the particle identification by the ACC. The $\pi^\pm/K^\pm$ separation and consequent kaon detection are one of essential issues for the CP violation study because a sum of kaon charges in an event gives good information to tag the flavor of B_{tag}. To meet the high $\pi^\pm/K^\pm$ separation performance in the multi-GeV/c momentum ranges, the threshold type Cherenkov counters are required to have refractive index between those of liquid and the solid [30]. The silica aerogel provides a good refractive index for this requirement. The silica aerogel is a porous colloidal form of $(\mathrm{SiO_2})_n$ with more than 95% porosity. It has low density because of the structure, and consequently it has low refractive index. The density and the consequent refractive index is determined according to its chemical production procedure. 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. A global configurations of the ACC are shown in Figure 2.14, and a typical single ACC module is shown in Figure 2.15 (a) and (b) for the barrel and the end-cap ACC, respectively. A single ACC module consists of blocks of silica aerogel contained in 0.2 mm-thick aluminum box. The silica aerogel has different five refractive indices, $n = 1.010$, 1.013, 1.015, 1.020, and 1.028, depending on its polar angle regions in order to obtain good $\pi^\pm/K^\pm$ separation for whole kinematic range of particles from B decays. The Cherenkov light generated in the silica aerogel is fed into one or two fine-mesh photo-multipliers (FM-PMTs) attached to the aerogel radiator modules. The FM-PMTs can operate in the 1.5 T magnetic field. A number of readout channels is 1560 in the barrel ACC and 228 in the end-cap ACC.

2.2.5 Time of Flight Counter – TOF

A time of flight counter (TOF), which is made of plastic scintillation counters, is used to distinguish kaons from pions up to 1.2 GeV/c. In addition to particle identification, the TOF provides fast timing signals for the trigger system to generate gate signals for ADCs and stop signals for TDCs, which are described in 2.4.2. The trigger modules attached to the TOF is called TSC (Thin Trigger Scintillation Counter). The counters measure the

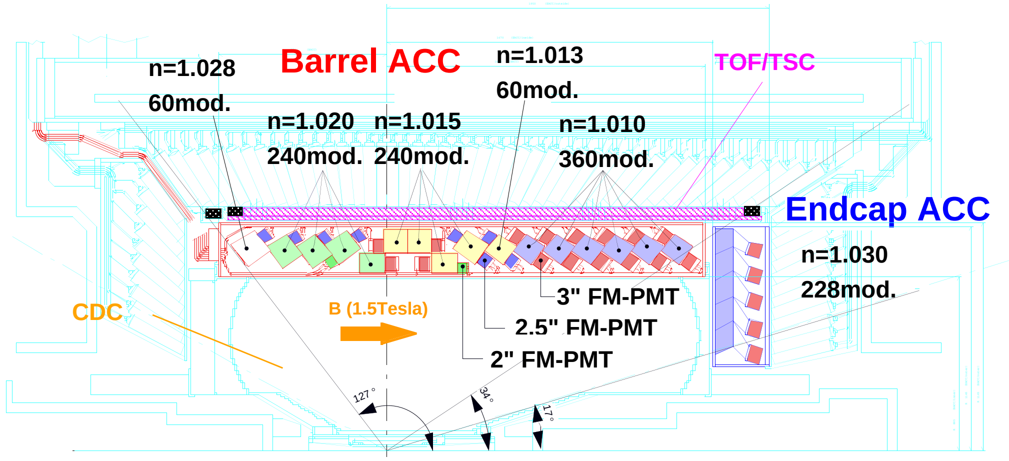


Figure 2.14: Arrangement of the ACC at the central part of the Belle detector.

elapsed time between a collision at the interaction point and the time when the particle hits the TOF layer. For the measured flight time with an appropriate correction (T) a particle mass (m) is represented by

$$m = p \cdot \sqrt{\frac{T^2}{L^2} - 1}, \tag{2.3}$$

where L is flight path length. The detection of p and T gives the particle species. The time of flight at $p = 1.2$ GeV/c is 4.3 ns for $K^{\pm}$ and 4.0 ns for $\pi^{\pm}$, respectively. The time resolution of 100 ps gives $\pi^{\pm}/K^{\pm}$ separation capability in 3σ.

Figure 2.16 is an illustration of one TOF/TSC counter of both ends. One 5 mm-thick TSC layer and one 4 cm-thick TOF counter layer are arrayed cylindrically at the position of $L = 1.2$ m in radius from the interaction point with 1.5 cm gap. The scintillators are wrapped with 45 μm thick polyvinyl film for light tightness and surface protection. A total number of 128 TOF counters are placed in the ϕ sectors, and each counter is viewed by the FM-PMT at both ends. One FM-PMT is glued to each TSC at backward end. A total number of TSC counters is 64. A total number of readout channels is 256 for the TOF and 64 for the TSC.

Figure 2.17 shows the time resolution for forward and backward PMTs of the TOF, separately, and for the weighted average of the time resolution as a function of hit position. For 1.2 m flight path, the time resolution must be better than 100 ps to achieve 3σ $\pi^{\pm}/K^{\pm}$ separation at 1.2 GeV/c. With this accuracy, we have clearly separated distributions of kaon mass from pion mass as shown in Figure 2.18. The result shown meets the

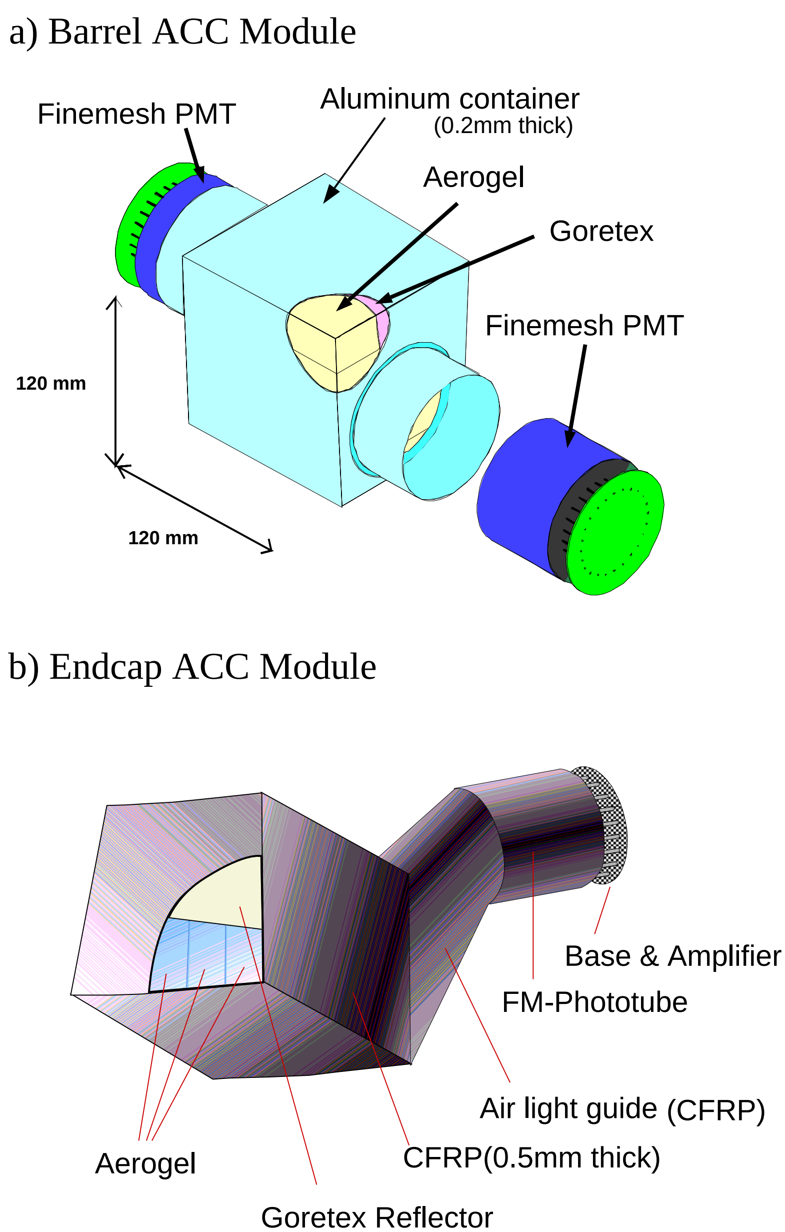


Figure 2.15: Schematic drawing of a typical ACC counter module for (a) barrel and (b) end-cap ACC.

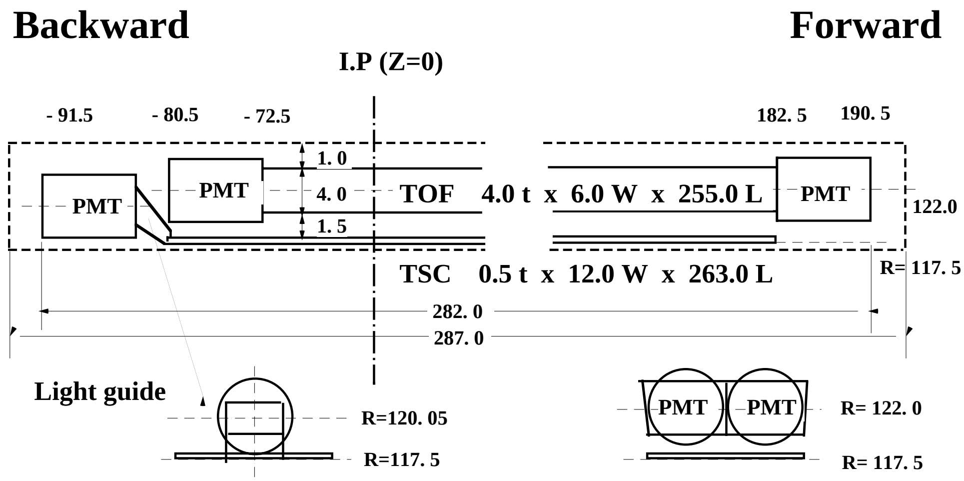


Figure 2.16: Dimensions of a TOF/TSC module.

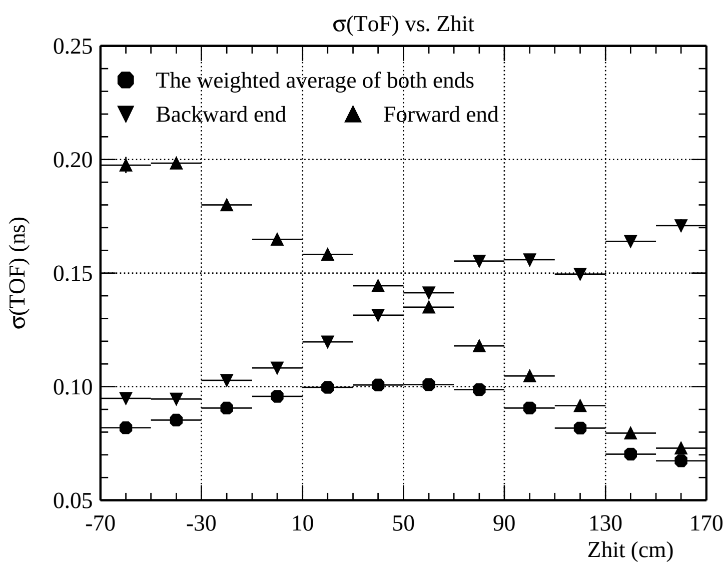


Figure 2.17: Time resolution for μ-pair events measured by the TOF.

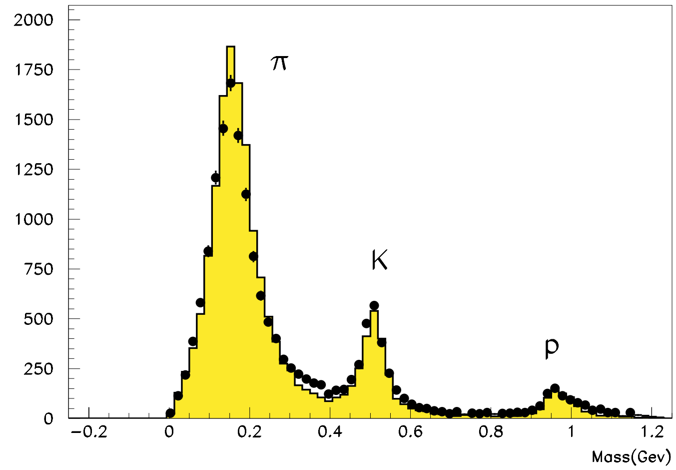


Figure 2.18: Mass distribution from TOF measurements for particle momenta below 1.2 GeV/c. The points are obtained from hadronic events, and histogram is obtained from Monte Carlo prediction by assuming time resolution of 100 ps.

requirements.

2.2.6 $\pi^{\pm}/K^{\pm}$ Separation

The $\pi^{\pm}/K^{\pm}$ separation is an important issue for the flavor tagging. In this subsection, we describe the method to separate kaons from pions. The separation of kaons from pions is based on three nearly independent observables, $\mathrm{d}E/\mathrm{d}x$ measurement by the CDC, the TOF measurement, and the photon emission detected by the ACC. Each of these detector is involved in the particle tagging in different momentum and angular region and the combination of them provides good particle separation performance.
The likelihood measured by the CDC is represented by

$$\mathcal{L}_{\mathrm{d}E/\mathrm{d}x} = \frac{\exp\left(\chi^2/2\right)}{\sqrt{2\pi}\sigma_{\mathrm{d}E/\mathrm{d}x}}, \qquad \chi^2 = \left[\frac{(\mathrm{d}E/\mathrm{d}x)_{\mathrm{mes}} - (\mathrm{d}E/\mathrm{d}x)_{\mathrm{exp}}}{\sigma_{\mathrm{d}E/\mathrm{d}x}}\right]^2, \tag{2.4}$$

where "mes" and "exp" denotes measured and expected $\mathrm{d}E/\mathrm{d}x$, respectively, and $\sigma_{\mathrm{d}E/\mathrm{d}x}$ is an expected resolution of $\mathrm{d}E/\mathrm{d}x$. The TOF χ^2 is calculated by taking the difference between a two-dimensional vector containing the observed times in two ends of the TOF counter and one containing the predicted times; $\Delta = \vec{t_{\mathrm{obs}}} - \vec{t_{\mathrm{prd}}}$. With defining a 2×2 error matrix E, χ^2 is represented as $\chi^2 = \Delta^T E^{-1} \Delta$, and the TOF likelihood is computed as

$$\mathcal{L}_{\mathrm{TOF}} = \frac{\exp\left(\chi^2/2\right)}{\prod_i^{\mathrm{ndf}} \sqrt{2\pi}\sigma_i}. \tag{2.5}$$

In contrast to the $\mathrm{d}E/\mathrm{d}x$ and the TOF counters, the ACC is basically an on-off device, or threshold type device, where the observed signal (number of photo-electrons, N_{pe})

takes zero or distributed according to small-number statistics. However, in an actual measurement, a sub-threshold particle has a small chance to fire the ACC with a few $N_{\rm pe}$ due to mainly three sources: (i) scintillation light from the reflector, (ii) high energy δ-ray, which is an orbital e^-, knocked on from the counter material by an incident particle (because of its light mass, the e^- is easily given high enough momentum to fire the ACC), and (iii) electrical and thermal noise from the front-end electronics and dark current of the FM-PMT. To consider these components, the measured $N_{\rm pe}$ is translated into probabilities, $\mathcal{L}_{\rm ACC}$, for assumed particle species at the measured β. We construct look-up table for $\mathcal{L}_{\rm ACC}$ from the Monte Carlo simulation. Although the ACC is still an on-off device basically, the usage of the $\mathcal{L}_{\rm ACC}$ gives high performance of the $\pi^\pm/K^\pm$ separation.

Figures 2.19 (a) through (c) show the discriminants for kaons and pions that are used to calculate their likelihoods. The Monte Carlo expectations are also shown. Figure 2.19 (a) shows the distributions of the number of photo-electrons for kaons and pions. The distribution is obtained for the ACC whose refractive index is 1.010. Figure 2.19 (b) shows the distribution of "measured" minus "expected" time-of-flight normalized by the expected error in an assumption of kaon track for $p < 1.2$ GeV/c. Figure 2.19 (c-1) and (c-2) show the distribution of $\Delta({\rm d}E/{\rm d}x)/\sigma$ for (c-1) $p < 0.8$ GeV/c and for (c-1) $p < 2.5$ GeV/c . The total likelihood is calculated as a product of above three likelihoods. The kaons are distinguished from pions with the likelihood calculated with each detector information as

$$P(K) = \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi}, \qquad P(\pi) = 1 - P(K). \tag{2.6}$$

The validity of the $\pi^\pm/K^\pm$ identification is demonstrated using charm decay, $D^{*+} \rightarrow D^0\pi^+$, followed by $D^0 \rightarrow K^-\pi^+$. Figure 2.20 shows two-dimensional plots of the likelihood ratio, $P(K)$ and measured momenta for the tagged kaon and pion tracks. Clear separation of kaons and pions up to 4 GeV/c can be observed. The identification efficiency and wrong identification fraction to pion as kaon is shown in Figure 2.21. The likelihood ratio selection, $P(K) > 0.6$, is applied in this figure. For most of the region, the measured kaon efficiency exceeds 80%, while the wrong identification fraction is kept below 10%.

2.2.7 Electromagnetic Calorimeter – ECL

The main purpose of the ECL is the detection of electrons and photons from B meson decays with high efficiency and good energy and position resolutions. In this study, the electron detection is especially an important issue because we reconstruct J/ψ meson from $J/\psi \rightarrow e^+e^-$(+radiative γ) to identify the decay final state of B_{CP}. Moreover, the leptons significantly indicate the flavor of $B_{\rm tag}$.

The electron identification relies primarily on a comparison of the charged particle momentum reconstructed by the CDC and the energy deposit at the ECL clusters. Elec-

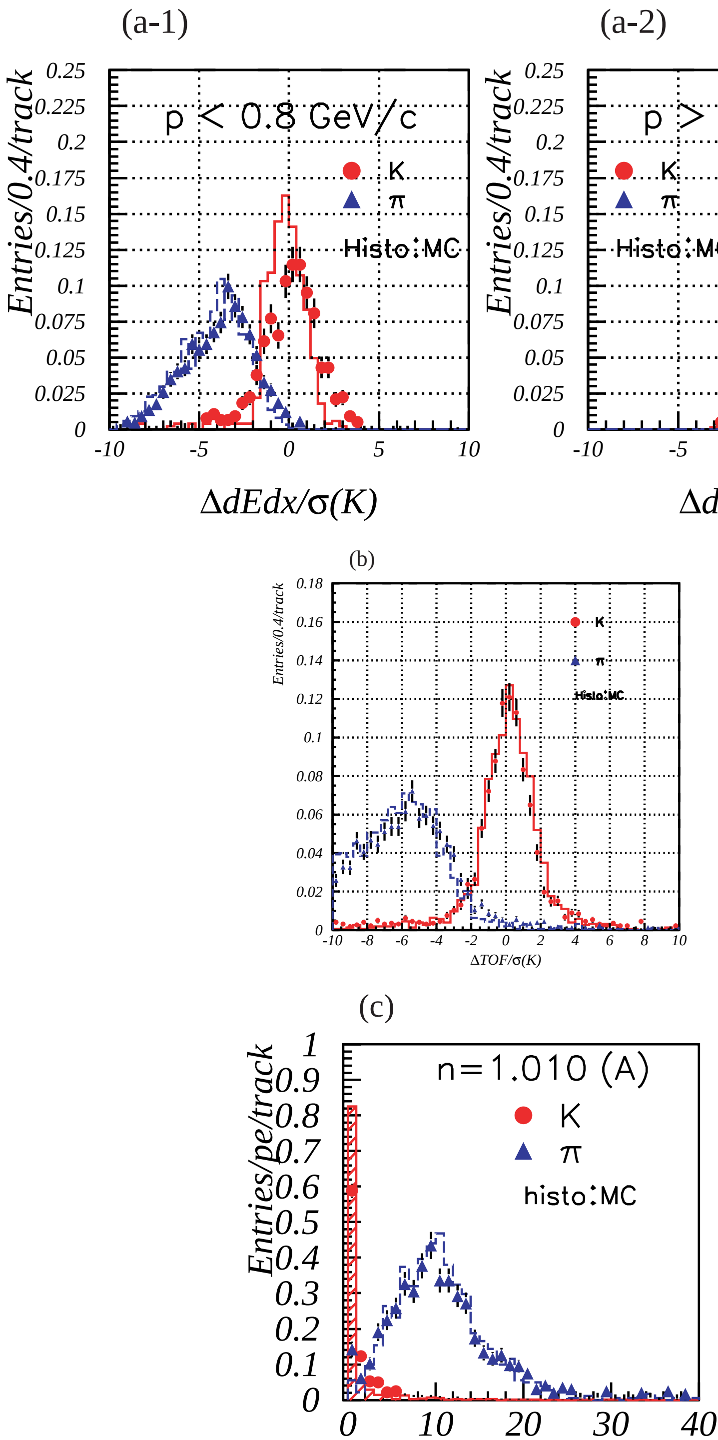


Figure 2.19: Discriminants for the calculation of kaon likelihood ratio. Circles and solid lines are for kaon distributions and the triangles and broken lines are for pion distributions. The points represent the distributions from data, and the histograms represent the ones from the Monte Carlo simulation. (a) dE/dx distributions normalized by the expected error, (b) "measured" minus "expected" time-of-flight normalized by the expected error, and (c) N_{pe} distribution.

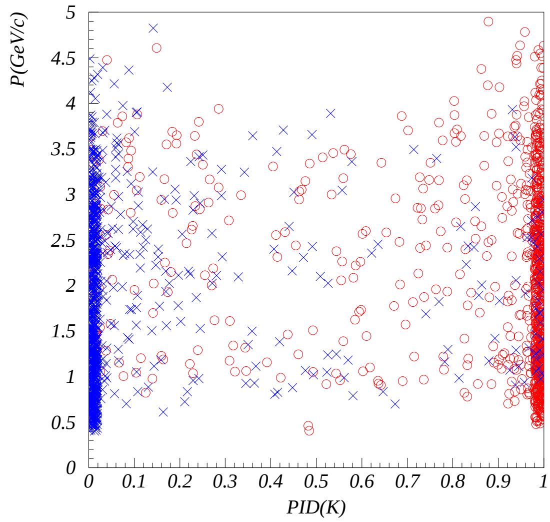


Figure 2.20: Likelihood ratio $P(K)$, versus momenta for daughter tracks from $D^0 \to K^-\pi^+$ decays, tagged by the charge of the slow π^+'s. The open circles correspond to kaons and the cross points to pions.

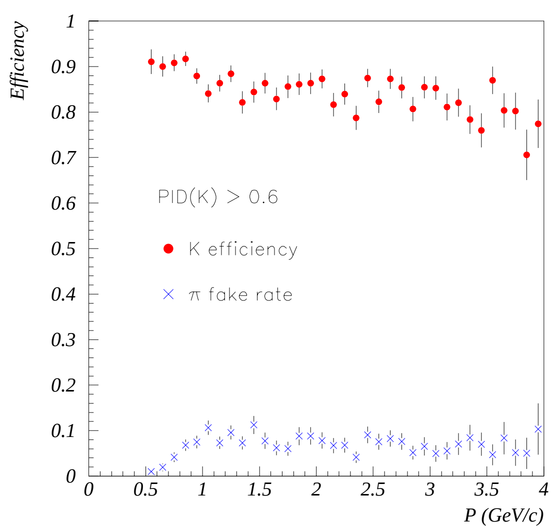


Figure 2.21: Kaon identification efficiency and wrong identification fraction to pion as kaon, measured with $D^{*+} \to D^0(K\pi) + \pi^+$ decays, for the barrel region. The likelihood ratio selection, $P(K) > 0.6$, is applied.

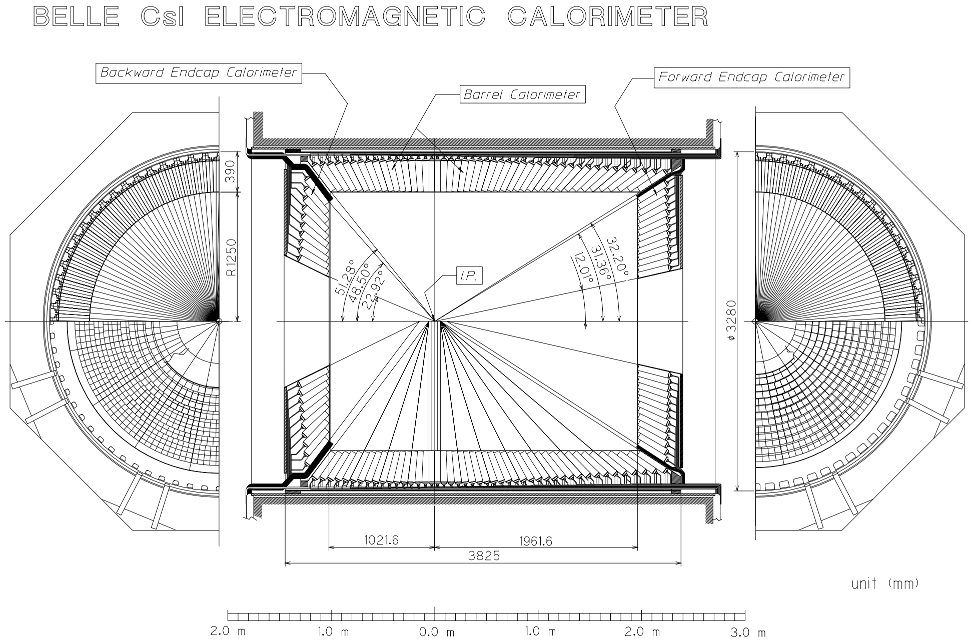


Figure 2.22: Overall configuration of the ECL.

trons deposit most of its energy at the ECL by electromagnetic shower, while $\mu^{\pm}$ or other hadrons deposit small fraction of their energies. Good energy resolution results in better hadron rejection. Since most photons are the end product of cascade B decays especially from π^0, the photons have low energies. It is also important to detect up to about 4 GeV photons for $B \to K^*\gamma$ or $B^0 \to \pi^0\pi^0$ decay studies. Therefore, the ECL is designed to provide good sensitivity of photons from low ($\lesssim$ 500 MeV) to high energy.

The overall configuration of the ECL is shown in Figure 2.22. The ECL consists of 8736 thallium doped CsI crystal counters. CsI(Tℓ) crystals have various nice features such as a large photon yield, weak hygroscopicity, mechanical stability, and moderate price. Each CsI(Tℓ) crystal is 30 cm long, which corresponds to $16.2X_0$, where X_0 is a radiation length. Each crystal has a tower-like shape and is arranged almost to the interaction point. The barrel crystals are located at $r = 1.25$ m, while forward and backward ECLs are located at $z = +1.96$ m and $z = -1.02$ m, respectively. The polar angle coverage of the ECL counters is $17^{\text{o}} < \theta < 150^{\text{o}}$, corresponding to 91% of total solid-angle. The barrel part ECL has a 46-fold segmentation in the ϕ direction and 144-fold segmentation in the θ direction. The ECL at the forward end-cap is segmented into 13-folds in the θ direction while the segmentation in the ϕ direction varies form 48 to 144. The ECL at the backward end-cap is segmented in 10-folds in the ϕ direction. A total readout channels of the ECL is 17472.

The position resolution is measured to be dependent on the photon energy:

$$\sigma_{\text{pos}} = \left(0.27 + \frac{3.4}{E^{1/2}} + \frac{1.8}{E^{1/4}}\right) \text{ mm}, \tag{2.7}$$

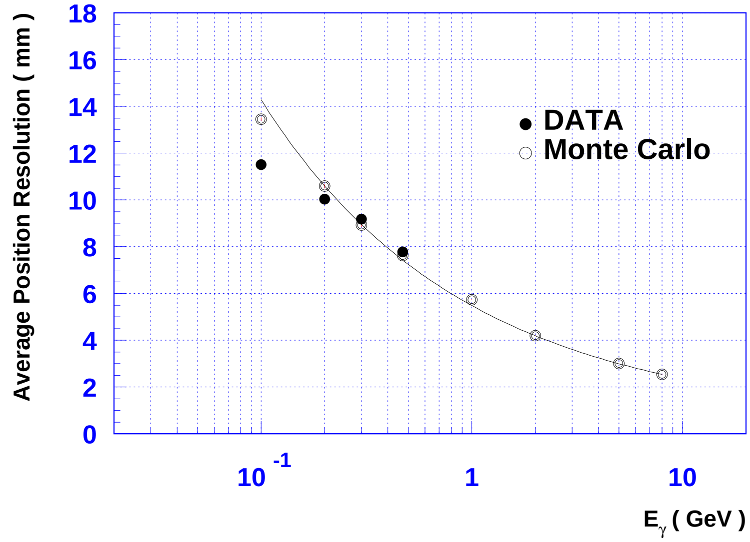


Figure 2.23: Position resolution of the ECL as a function of incident photon energy. The solid curve indicates a fit to equation (2.7).

where E is in GeV. The position resolution is shown in Figure 2.23. The points above 1 GeV are obtained from Monte Carlo simulation. The energy resolution is measured as a function of incident photon energy for 3×3 matrices and for 5×5 matrices of the ECL counters. The nominal resolution is measured to be

$$\frac{\sigma_E}{E} = \left(1.34 \oplus \frac{0.066}{E} \oplus \frac{0.81}{E^{1/4}}\right) \% \tag{2.8}$$

with the study of 3×3 ECL matrices. E is in GeV. The measurement results are shown in Figures 2.24. These performances enable us, for example, to reconstruct $\pi^0 \to \gamma\gamma$ candidate with invariant mass resolution of 4.8 MeV/c^2, and $\eta \to \gamma\gamma$ candidates with the resolution of 12 MeV/c^2.

Photon Reconstruction

Photon reconstruction starts from the ECL clustering. First, “seed” crystal are searched for that have highest energy deposit than any neighboring crystals. Total energies in 3×3 crystal matrices ($E_{3\times3}$) and 5×5 crystal matrices ($E_{5\times5}$) (crystals taken into account has energy deposit at least 500 keV) around the seed are calculated. To be identified as a photon cluster, the crystal surface of the cluster must not be associated to charged particle trajectory. When such cluster has total energy deposit greater than 500 MeV, it is identified as photon. In case of the cluster energy greater than 20 MeV, $E_{3\times3}/E_{5\times5} > 0.75$ is required for the photon identification. When the energy deposit is

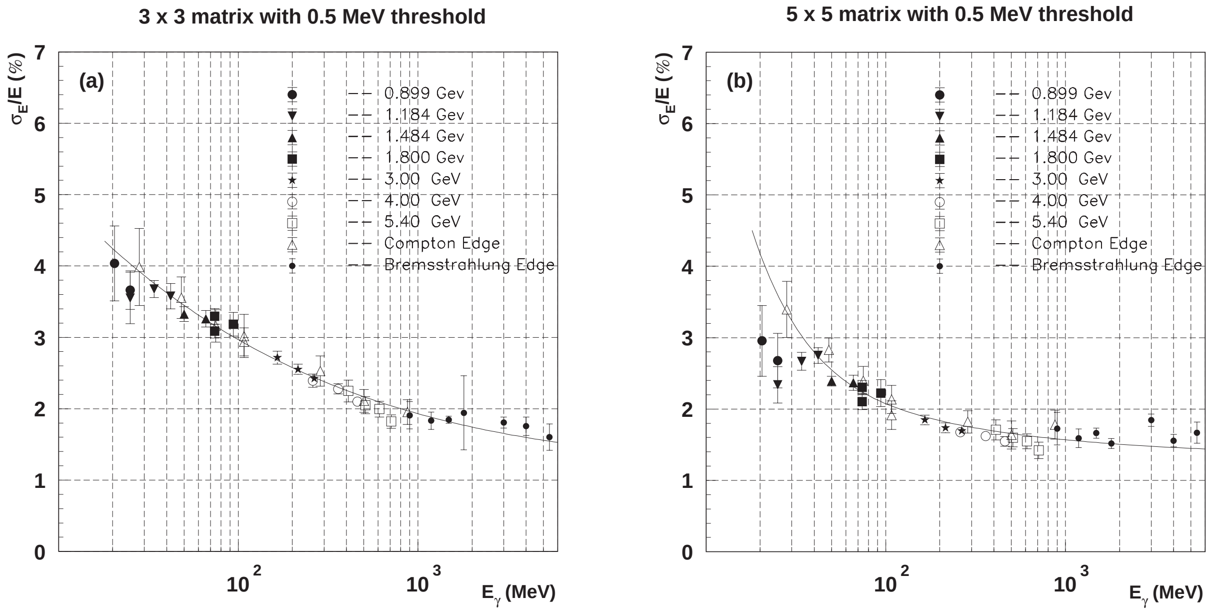


Figure 2.24: ECL energy resolution as a function of incident photon energy for (a) 3×3 and (b) 5×5 matrices. The error bars are the r.m.s. values of four measurements taken with the crystals (3,3), (3,4), (4,3),and (4,4) into the photon beam.

smaller than 20 MeV, the cluster is not considered as a photon hit.

Electron Identification

As we described, most significant signature of electron is an energy deposit in the ECL by the electromagnetic shower associated with a charged track. Since most electron energy is lost by the electromagnetic shower in the ECL, the E/p distribution, where E is ECL cluster energy and p is reconstructed momentum by the CDC, for electrons forms a peak around 1, while that for hadrons does not. In addition, the lateral spread of the electromagnetic shower and the hadronic shower differs because the radiation length of the electron is smaller than the interaction length of the hadron, which is also an significant signature for electron identification. The other remarkable signature is a distinguishable $\mathrm{d}E/\mathrm{d}x$ distribution of electrons from other charged hadrons or muons. The electron identification is established based on mainly these facts. The electron identification watches following discriminants: (i) matching of an extrapolated charged track position and the ECL cluster, (ii) the ratio E/p, (iii) shower shape at the ECL, and (vi) $\mathrm{d}E/\mathrm{d}x$ measured by the CDC. In addition to above requirements, light yield in the ACC is also incorporated for the electron identification to eliminate kaons.

Figure 2.25 (a) shows an E/p distribution of electrons and pions. The histogram of electrons has a peak around $E/p \sim 1$. A tail observed in lower E/p region of electrons comes from an interaction of electrons with materials in front of the ECL. The additive

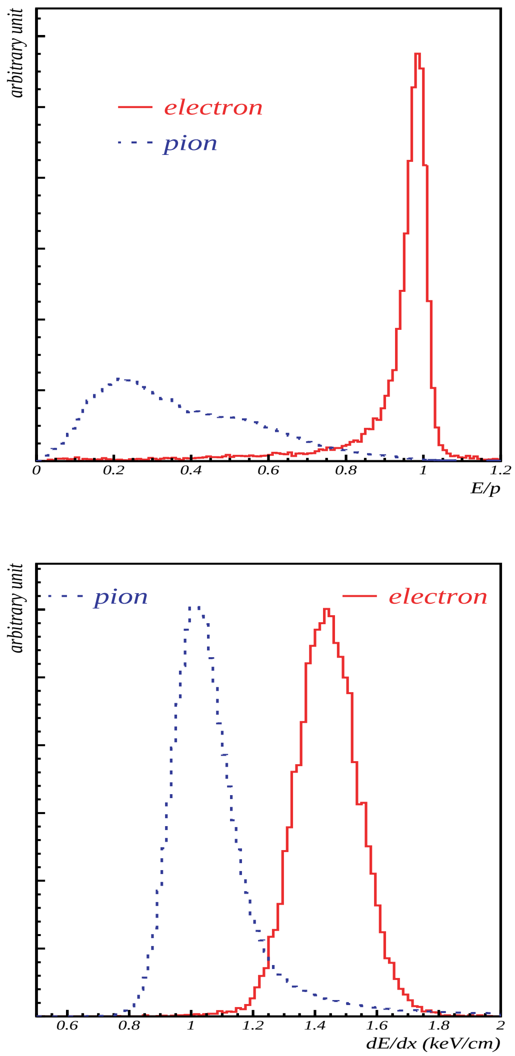


Figure 2.25: Input parameters for the electron identification algorithm. Each distribution is comparing the distributions of electron and pion. (a) and (b) are E/p and $E_{3\times3}/E_{5\times5}$ distributions measured by the ECL, respectively, and (c) is $\mathrm{d}E/\mathrm{d}x$ distribution by the CDC.

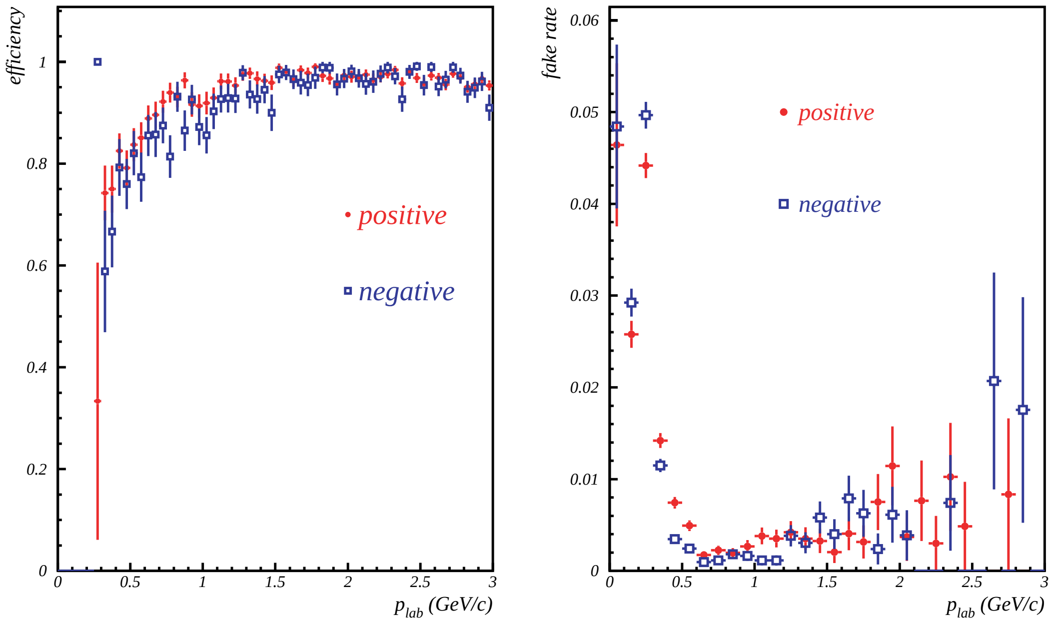


Figure 2.26: (a) Electron identification efficiency with $\mathcal{L}_e > 0.5$ in radiative Bhabha events for the barrel region, and (b) wrong identification fraction of electron to pion as a function of momentum in the laboratory frame.

$p_{\rm lab}$ information gives better electron identification in even lower E/p region. Figure 2.25 (b) shows the distributions of $E_{3\times3}/E_{5\times5}$ of electrons and pions. As we stated, due to the difference of the shower evolution, $E_{3\times3}/E_{5\times5}$ for the electromagnetic shower is larger than that of hadronic interaction shower. Energy deposit for electrons and pions measured by the CDC is shown in Figure 2.25 (c). The dE/dx also enables us to distinguish electrons from hadrons effectively. The likelihood ratio computed with the combination of above parameters are

$$\mathcal{L}_e \equiv \frac{\prod CL_e^i}{\prod CL_e^i + \prod CL_{\rm not-e}^i}, \tag{2.9}$$

where CL_e^i and $CL_{\rm not-e}^i$ denotes the confidence level of the i-th discriminator for electrons and non-electrons, respectively. The electron identification efficiency with $\mathcal{L}_e > 0.5$ for radiative Bhabha events is shown in Figure 2.26 (a). The distribution of the electron identification efficiency and the wrong identification fraction of electrons to pions as a function of particle momentum in the laboratory frame are shown in Figure 2.26 (b). Typical efficiency is measured to be $> 90\%$, and typical wrong identification fraction is measured to be $< 1\%$ for $p_{\rm lab} > 1.0$ GeV/c.

The detailed description of electron identification procedure is given in [36].

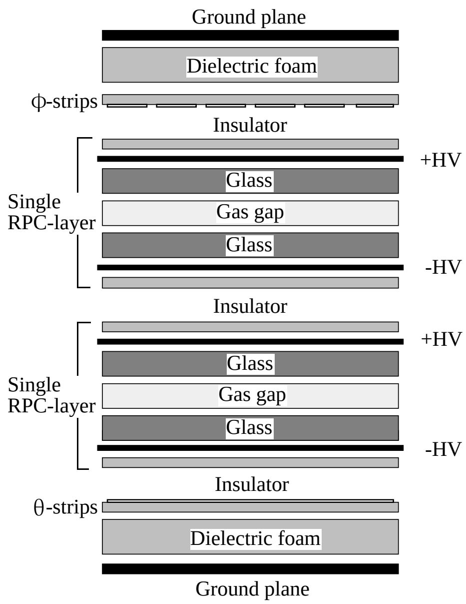


Figure 2.27: Cross section of a KLM superlayer.

2.2.8 Magnetic Field

A charged particle in a magnetic field runs through helical path of which radius corresponds its momentum. To measure particle momentum in the CDC, 1.5 T magnetic field is applied parallel to the beam pipe. The provider of the magnetic field is a superconducting coil consisting of a single layer of niobium-titanium/copper embedded in high purity aluminum stabilizer. The coils are chilled by liquid helium. The return path of the magnetic flux is provided by the iron structure. The iron structure also works as an absorber material for the KLM and a support for all of the detector components.

2.2.9 K_L and Muon Detector – KLM

Outside of the magnetic field, K_L and muon detector (KLM) is installed. The main purpose of the KLM is the detection of muons from $J/\psi \rightarrow \mu^+\mu^-$ decay or K_L from $B^0 \rightarrow J/\psi K_L$ decay, which can be used for CP violation study. The KLM is designed to detect K_L and muons with momenta of $p > 600$ MeV/c. The muons with less than 500 MeV momentum do not reach at the KLM due to energy loss in front of the KLM. The KLM consists of alternating layers of charged particle detectors and 4.7 cm-thick iron plates. There are 15 detector layers and 14 iron layers in the octagonal barrel region and 14 detector layers in each of forward and backward end-cap region. Figure 2.27 shows a cross section of the KLM super-layer. The polar angle coverage of the KLM is $20^{\circ} < \theta < 155^{\circ}$. The detection of charged particles is provided by glass-electrode

resistive plate counters (RPCs) [37]. The RPC has two electrodes with high resistivity of $\geq 10^{10}\ \Omega\cdot$cm separated by gas-filled gap that consist of 30% argon, 8% butane, and 62% freon. In the streamer mode, 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 the time of the ionization. The strips are roughly 5 cm wide and configured in the θ and ϕ directions. Total number of the readout channels is 37984.
K_L's are identified by the detection of K_L hadronic interactions in the ECL, iron yoke, or the KLM itself. It is not possible to measure K_L energy with high accuracy because of large fluctuation in the size of the shower, but possible to measure the direction of K_L momentum. Since $B^0 \rightarrow J/\psi K_L$ is two-body decay and J/ψ can be fully reconstructed, the measurement of K_L energy is not essential for the reconstruction of the B^0 candidate decaying into this mode. The K_L's positions are detected with the accuracy of $\Delta\phi = \Delta\theta = 30$ mrad, while timing resolution is few ns. Charged pions and kaons are attenuated by hadronic interactions in the ECL and other absorbers, while muons penetrate the iron plates of the KLM much farther with few deflections by Coulomb scattering. This enables us to detect muons by the KLM. The muon detection efficiency is better than 90% above 1.5 GeV/c momentum.

Muon Identification

We identify muons by the facts that due to small interactions muons are more penetrating than pions that are dominating background for muon identification. The extrapolated tracks from the CDC is associated to the hits on the KLM. Around the extrapolated track trajectory, number of penetrated iron plates, called N_{meas}, are counted. The number of crossing iron plates with the track trajectory, N_{cros}, is also computed geometrically. Because of the penetrative characteristic of the muons, $\Delta N \equiv N_{\mathrm{cros}} - N_{\mathrm{meas}}$ for muons becomes smaller number than hadrons. Figure 2.28 (a) shows the ΔN distribution for muons and pions obtained from Monte Carlo simulation. In addition, we use the fact that due to small interactions, muons are not scattered by the KLM material during traversing the KLM. Therefore, the hit positions on the KLM RPCs can be assumed not to be deviated from the geometrical expectation by the CDC track trajectory and the KLM RPC positions. If we extrapolate the CDC track by assuming the particle being muon, and if the real species of the particle is pion, the deviation of the real hit positions from the expected hit positions becomes large. We represent such deviations by the reduced χ^2. The distributions of the reduced χ^2 for muons and pions are shown in Figure 2.28 (b). The muon identification is performed with computing likelihood ratio being muon, as well as the electron identification. Figure 2.29 shows the distribution of likelihood ratio for muons and pions. Figure 2.30 (a) shows the muon detection efficiency.

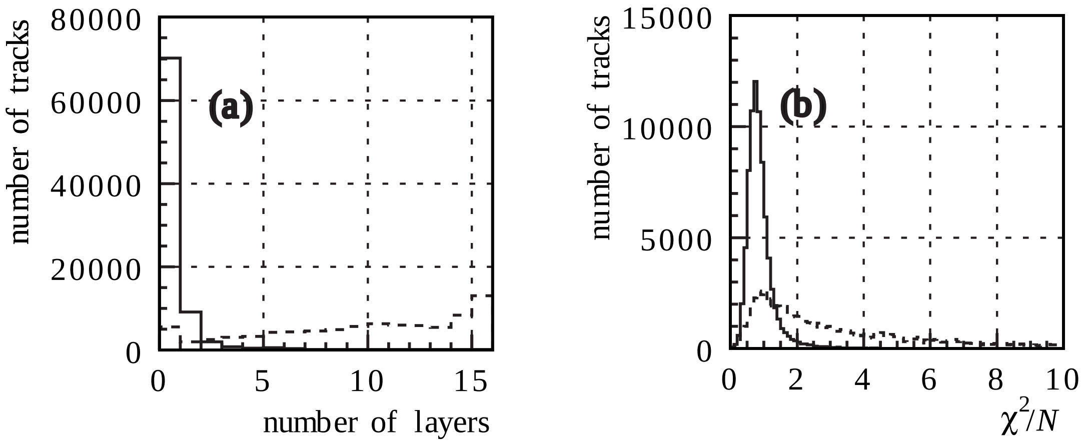


Figure 2.28: (a) ΔN distribution and (b) the reduced χ^2 of the transverse deviations between computed and measured points. Solid lines and broken lines represent the distributions for muons and for pions, respectively.

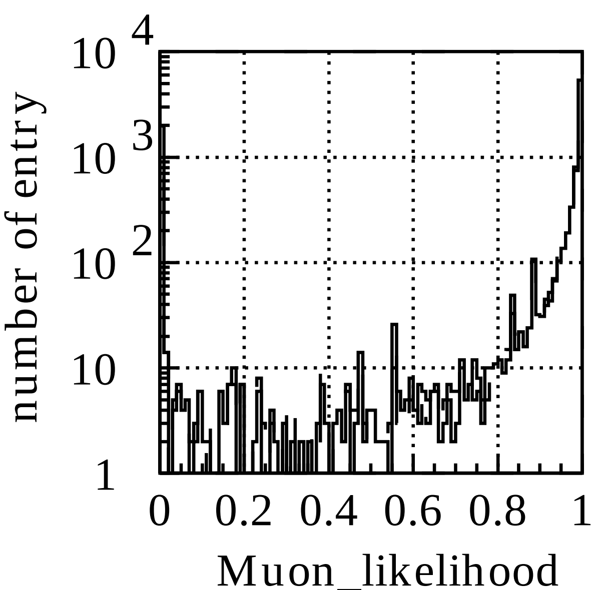


Figure 2.29: Likelihood distribution for muon identification.

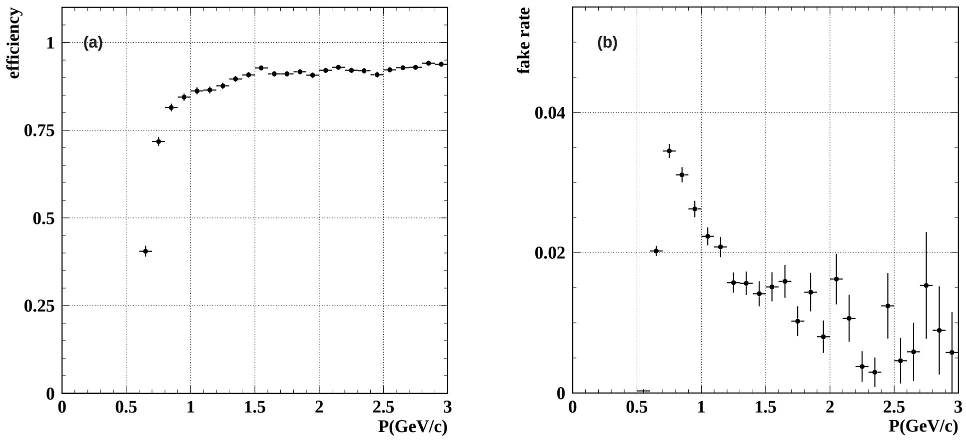


Figure 2.30: Detection efficiency and the wrong identification fraction of muons to pions as a function of the muon momentum in the KLM.

Typical efficiency for $p_{\rm lab} > 1$ GeV/c is better than 90%. Wrong identification fraction of muons to pions is measured with $K_S \to \pi^+\pi^-$ decay. The fraction is smaller than 2% for $p_{\rm lab} > 1$ GeV/c. It is shown in Figure 2.30 (b). The details of the muon identification is given in [38].

2.2.10 Extreme Forward Calorimeter – EFC

In order to improve the experimental sensitivity to some physics processes such as $B \to \tau\nu$, the extreme forward calorimeter (EFC) is needed to extend the ECL coverage, $17^{\rm o} < \theta < 150^{\rm o}$. The EFC covers the angular range from 6.4° to 11.5° in the forward direction (electron direction), and from 163.3° to 171.2° in the backward direction. Since the EFC is placed in the very high radiation-level area by photons or electrons due to synchrotron radiation and spent electrons, radiation hardness is required at Mrad level. We adopt a crystal calorimeter made of $\rm Bi_4Ge_3O_{12}$ as the EFC, which satisfies the radiation hardness and provides good energy resolution for electrons and photons, $(0.3 - 1.0)\%/\sqrt{E({\rm GeV})}$, with reasonable costs. Both forward and backward EFC consist of $\rm Bi_4Ge_3O_{12}$ crystals segmented into 5 regions in the θ direction and 32 regions in the ϕ direction in order to provide better position resolution. Typical cross-section of a crystal is about 2×2 cm^2 with $12X_0$ for forward and $10.5X_0$ in backward, where X_0 is the radiation length.

2.3 Trigger

The role of the trigger system is to determine to record or discard the hit signals on each sub-detector. A data taking procedure to record the signals is initiated by the trigger signal. The event rates for both physics processes and backgrounds are estimated

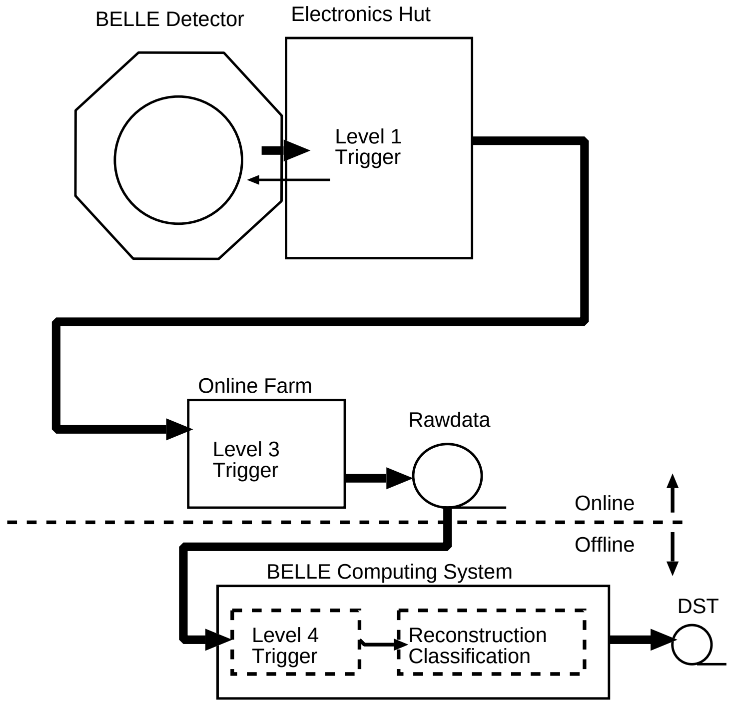


Figure 2.31: Overview of the Belle trigger system.

to be $\sim$ 100 Hz, respectively. The total event rate is estimated to be about 200 Hz at the goal luminosity of 10^{34} cm^{-2}s^{-1}. The components of the event rates for the physics processes are listed in Table 2.3, together with the background trigger rate. Besides the electrical noise, we have many background sources associated with the beams: spent (or off-momentum) electrons ($e^{\pm}$'s) or/and photons produced by Bremsstrahlung of beam particles with residual gas, Coulomb-scattered $e^{\pm}$s by residual gas atoms, $e^{\pm}$s from the radiative Bhabha scattering at the interaction point, and the synchrotron radiation. Among these backgrounds, the spent e^- background from the high-energy beam is measured to be the severest.

The trigger system consists of the level-1 hardware trigger [39] and the level-3 software trigger [40], where the latter one runs on the online computer farm. In addition, the Belle trigger system is equipped with the level-4 trigger [41], which can perform more elaborated reduction of background events on the offline computer farm. This multi-level trigger system is designed to keep up with an expected increase of beam background due to the upgrade of the KEKB performance. The overview of the Belle trigger system is shown in Figure 2.31. We describe each trigger level in this section.

Figure 2.32 shows the schematic view of the level-1 trigger. The CDC, TOF, ECL, KLM, and EFC generate trigger signals, which are fed to global decision logic (GDL)

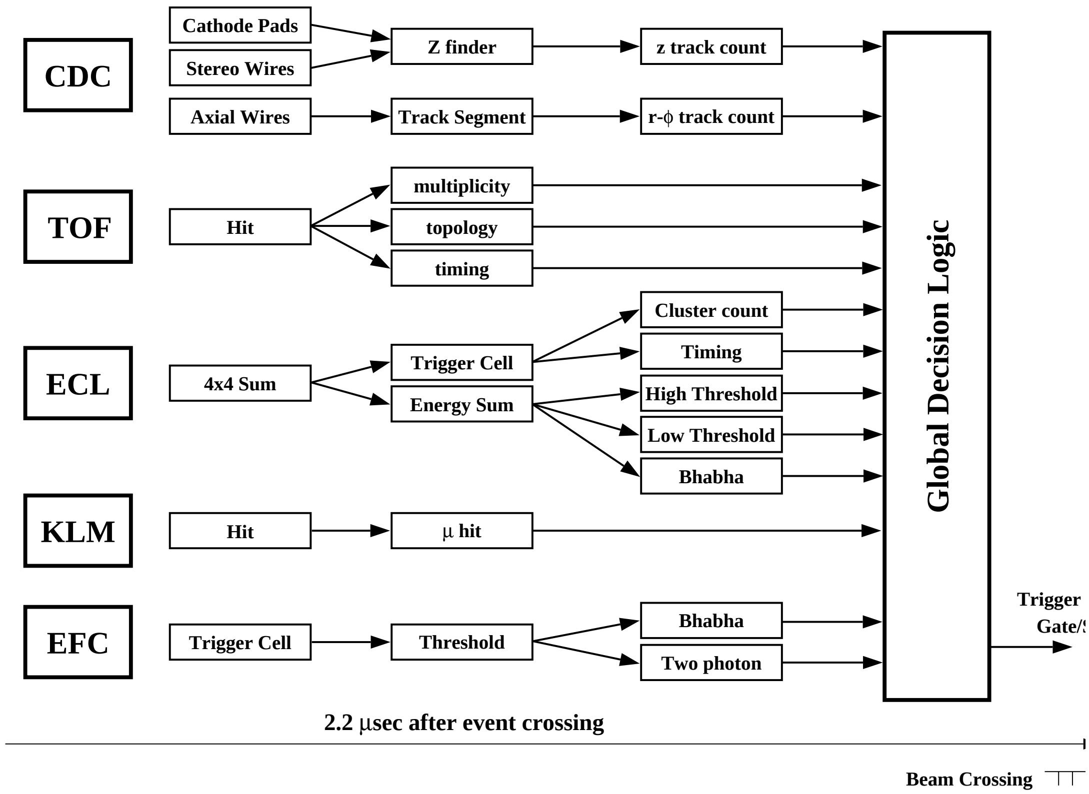


Figure 2.32: Level-1 trigger system for the Belle detector.

to determine whether the event should be stored or discarded. The sub-detector trigger systems can be categorized into two types: one is a track trigger and the other is an energy trigger. The CDC provides r-ϕ and r-z track trigger signals. The TOF trigger system provides an event timing signal and information of the hit multiplicity and topology. The ECL generates trigger signals based on the total deposit energy and the cluster counting. Two photon events as well as Bhabha events are tagged by the ECL triggers. The KLM trigger gives additional information on muons. Because of the high beam current, high beam backgrounds are anticipated. Since the background rates are sensitive to the actual accelerator condition, it is difficult to make reliable estimation. Therefore, the trigger conditions should be flexible enough so that background rates are kept within tolerance of the data acquisition system, while the efficiency of physics events is kept high. The trigger condition installed in the GDL is programmable to realize the flexibility. The GDL makes final decision within 2.2 μs after beam crossing time. The primary timing is generated by TOF trigger that has a time jitter less than 10 ns time jitter. Once the GDL trigger is generated, the trigger signal is distributed to each DAQ sub-system, which is described in the next section.

The aim of the level-3 trigger is the reduction of the storage usage. It runs on the online

Table 2.3: Total cross section and trigger rates with $\mathcal{L} = 10^{34}$ cm^{-2}s^{-1} from various physics processes at $\Upsilon(4S)$. † : values pre-scaled by a factor 1/100, ‡ : restricted condition of $p_t \geq 300$ MeV/c.

Physics process	Cross section (nb)	Rate (Hz)
$\Upsilon(4S) \to B\overline{B}$	1.2	12
Hadron production from $e^+e^- \to q\overline{q}$	2.8	28
$\mu^+\mu^-$	0.8	8
$\tau^+\tau^-$	0.8	8
Bhabha ($\theta_{\mathrm{lab}} \geq 17^{\mathrm{o}}$)	44	$4.4^\dagger$
$\gamma\gamma(\theta_{\mathrm{lab}} \geq 17^{\mathrm{o}})$	2.4	$0.24^\dagger$
2γ process ($\theta_{\mathrm{lab}} \geq 17^{\mathrm{o}}$, $p_t \geq 100$ GeV/c)	~ 15	$\sim 35^\ddagger$
Total	~ 67	~ 96
Backgrounds	–	~ 100

computer farm, which outputs the data to be stored on the tape recorder. The description of the online computer farm is given in Section 2.4.3. The event filtering is based on the number of reconstructed tracks and the total momentum in the r-ϕ plane obtained by a fast tracking program. The charged tracks with $p_{\mathrm{t}} > 350$ GeV/c are reconstructed form the GDL inputs using a memory look-up method. Δz can also be measured up to the resolution of about 9 mm using drift time information. The reduction factor of the events is about 2.

The level-4 trigger running on the offline computer farm is installed to reduce the CPU consumption of the analysis by 10%. The required reduction factor by the trigger is $5-10$ without causing sizable inefficiencies for all physics events. The level-4 trigger reconstruct tracks using the CDC hit signals. It rejects 78% of triggered events, while retaining very close to 100% of B decay events.

2.4 Data Acquisition – DAQ

The goal performance for the data acquisition system (DAQ) is to catch up 500 Hz trigger rate, while keeping a dead-time fraction less than 10% so that we can catch up with the improvement of the KEKB performance. In order to achieve this requirements, the distributed-parallel system has been devised. The global design of the system is shown in Figure 2.33. The entire system is segmented into 12 DAQ sub-systems running in parallel, where each of them is comprised in a VME-crate. Because of relatively large data size to other sub-detectors, SVD DAQ consists of four DAQ sub-systems and CDC

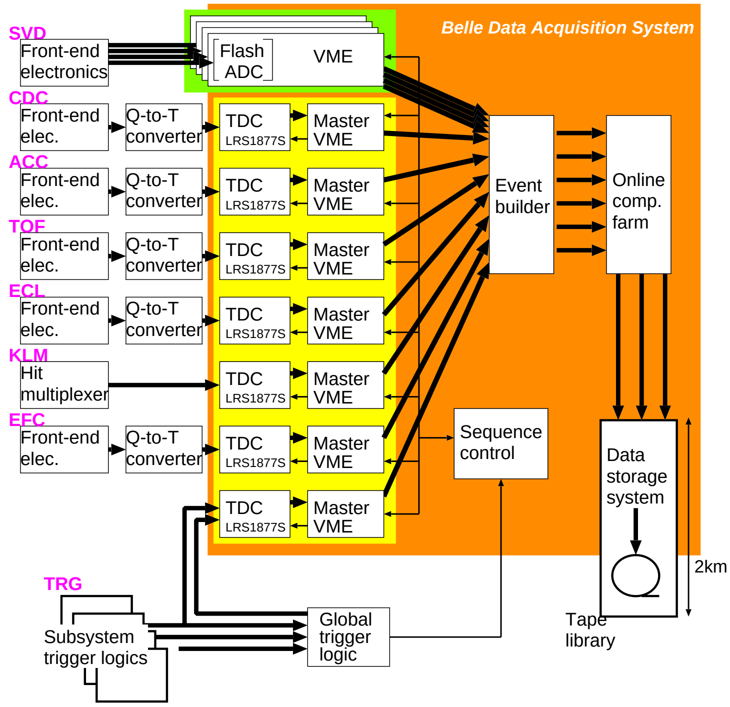


Figure 2.33: Schematic view of the Belle DAQ system.

DAQ consists of two DAQ sub-systems. Once the trigger signal is generated by the GDL, it is notified to the DAQ sub-systems by timing distribute modules (TDMs) through a sequence controller (SEQ). The data acquisition for the triggered event is started by the TDM signal. The analog hit signal is digitized by Q-to-T technique except for the SVD and by A/D converter for the SVD. Digitized data from each sub-system are read by the DAQ sub-system, and then transferred to the event builder to be combined into a single event record [42]. The event builder converts "detector-by-detector" parallel data to an "event-by-event" data stream. The output from the event builder is transferred to the online computer farm. At the online computer farm, data formatting, calibration, event reconstruction, and event filtering known as level-3 trigger is performed. The online computer farm also performs data sampling to be fed to the data-quality-monitoring stream, so that experimentalists can verify the quality of data flowing. As the last procedure, the online computer farm dumps the data to mass storage system, located at the computer center connected with optical fibers. A typical data size of a hadronic event by $B\overline{B}$ or $q\overline{q}$ production is measured to be about 30 kB, which corresponds to the maximum data transfer rate of 15 MB/s.

Figure 2.34 shows the total dead-time of the DAQ system as a function of trigger rate measured in real beam runs. At the typical trigger rate, 200 Hz, we achieve the dead time less than 4%. The overall error rate of the DAQ flow is measured to be less than

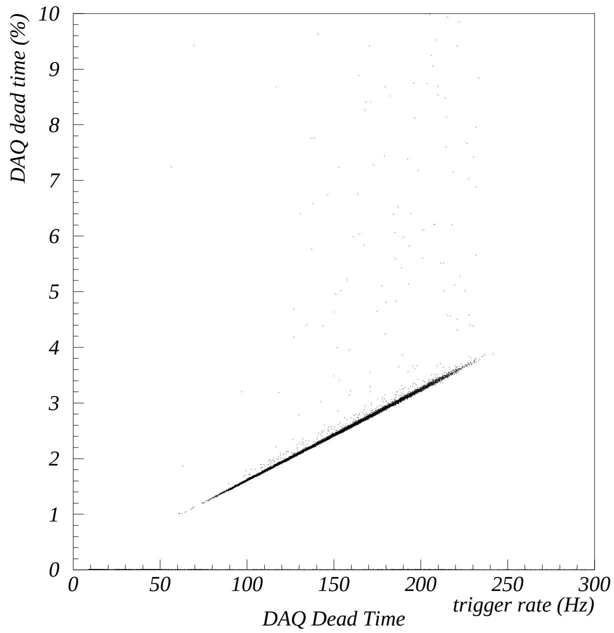


Figure 2.34: Dead time of the DAQ as a function of the trigger rate.

0.05%. The major source of the dead time is considered the data transfer from readout system to the event builder.

2.4.1 Trigger Timing Distribution

A signal of the final trigger decision by the GDL is first fed to a sequence controller (SEQ). The SEQ ensures the correct timing handshaking by detecting "busy" signal from each DAQ sub-system. It also monitors the trigger rate and the dead time. The TDMs are located in the VME-crate for the control of each DAQ sub-system. The TDM receives the trigger signals from the SEQ, and distributes them to the DAQ modules in the VME-crate in which the TDM resides, and hence from the detector readout side, it is seen as if triggers are generated from the TDM. Upon the signal from the TDMs, the DAQ sub-system initiates the data acquisition procedure. During the data processing, the DAQ sub-systems raise "busy" flag for the SEQ through the TDM until they finish the signal digitization so that the next trigger is not issued.

2.4.2 Signal Digitization

The hit information of each detector is digitized using a charge-to-time (Q-to-T) converter except for the SVD. The Q-to-T converts charge amount to the time interval by storing the charge to a capacitor and discharging it with a constant rate. A pulse, whose width corresponds to the signal amplitude, is generated with the timing of the leading and trailing edges corresponding to the start and stop of discharging. By digitizing the timing

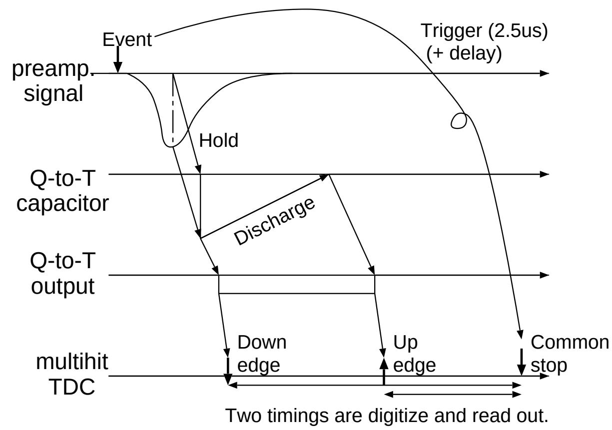


Figure 2.35: Principle of the Q-to-T technique.

of the leading and trailing edge of the pulse with respect to “a common stop timing”, we can determine both the timing and the amplitude of the input signal, which has great advantage for the CDC data acquisition. The distributed trigger signal from the TDM is regarded as the “common stop timing”. Figure 2.35 shows the principle of the Q-to-T technique. It is necessarily for particle tracking by the CDC to measure both charge left on sense wires and drift time of the initial ionization clusters to the wire. The Q-to-T technique realizes both detections with single probe. These timings are digitized with time digitization module (TDC) equipped with LIFO. Upon the common stop timing, the TDC stores the digitized data to its memory. The data on the TDC LIFO are read by VME modules and then transferred to the event builder, as illustrated in Figure 2.36.

The SVD has 81920 readout channels and it is not realistic solution to perform the A/D conversion channel by channel in parallel. The charge information is read out by 32 intelligent flash ADC modules (HALNYs) with embedded four DSPs in each, which performs data sparsification, zero suppression, and data compression. As a resulting data size is compressed to $\sim$ 10 kB. Since the single event size is still larger than the capacity of one event-builder-transmitter, the SVD readout system is divided into four sub-systems. In each sub-system, signals are digitized by the HALNYs, and transferred to the event builder in parallel. The SVD readout sequence is synchronized and initiated by an external timing modules (TTMs). TTMs receive a trigger signal from the TSC and distribute it to the VA1s. TTMs also wait for the decision from the GDL notified via TDM, so that they can initiate A/D conversion of the scanned data. The readout software from the flash ADCs is combined with the event-builder-transmitter software that runs on a SPARC VME CPU to minimize the overhead of copying data. The block diagram of the SVD DAQ is shown in Figure 2.37.

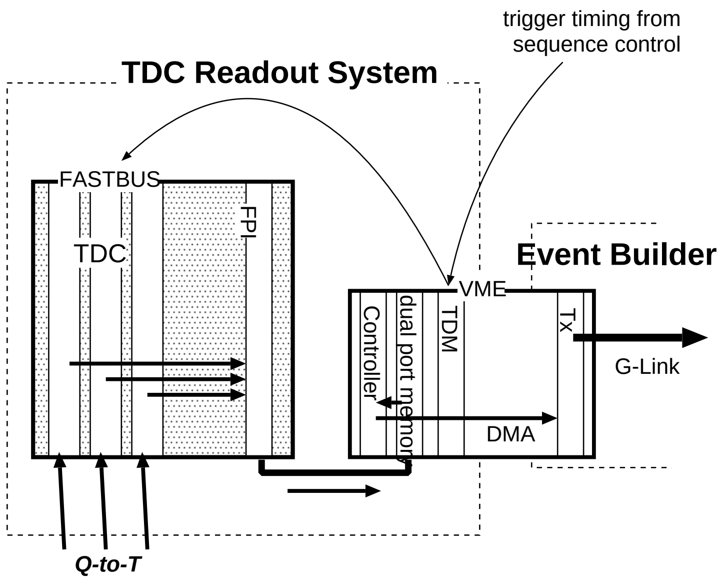


Figure 2.36: Illustration of the TDC readout system.

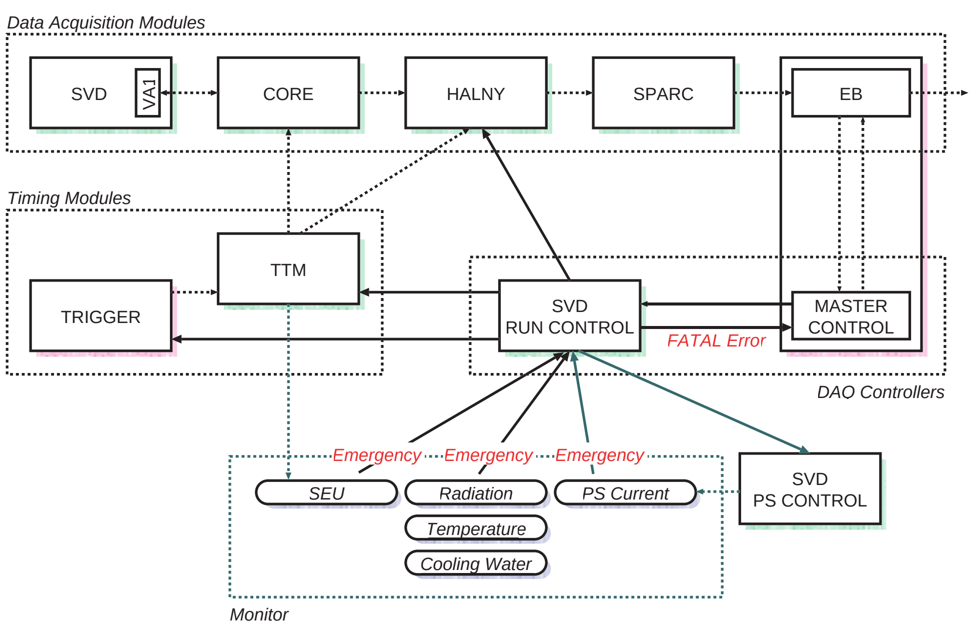


Figure 2.37: Block diagram of the SVD DAQ system. The global control is made through run control software, where the detail is described in Section 2.4.4.

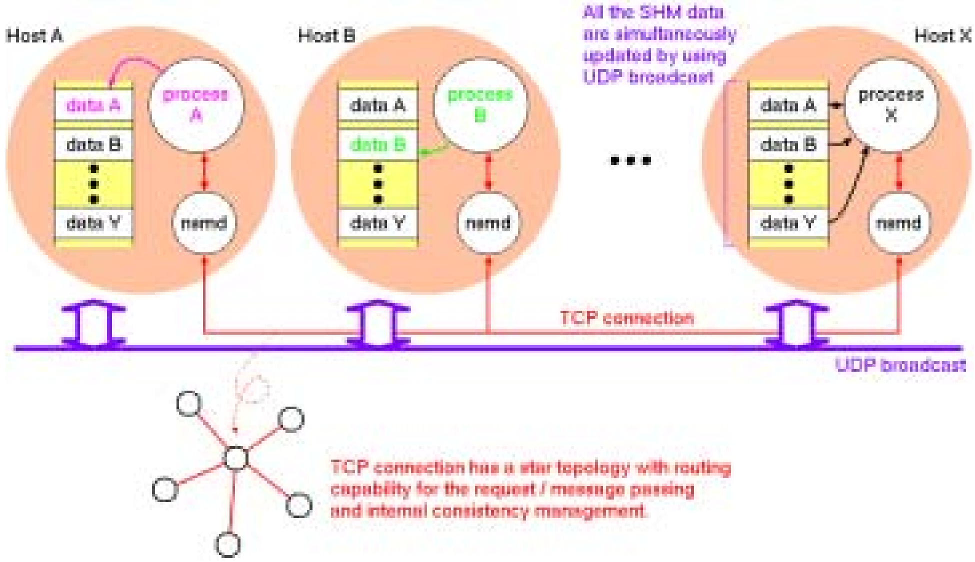


Figure 2.38: Architecture of the NSM.

2.4.3 Event Building and Online Computer Farm

The detector signals from VME readout modules and SPARC readout modules are combined to a single event by the event builder. The data are sent via 1.2 Gbps network from the modules to the event builder. The event builder is configured to be 12×6. The hit signals from up to 12 DAQ sub-systems are built up to single event then transferred to one of six online computers. The role of the online computer farm is to format an event data into an offline event format, to proceed level-3 trigger, and to output the data to the storage system. The online computer farm consists of six VME crates, where each contains a PowerPC based control processor. A fraction of events are sampled by the online computer farm and sent to a PC server via Fast Ethernet for the real-time monitoring of data quality and event display. The sampling rate is typically 20 Hz.

2.4.4 Run Control

The control of all the Belle DAQ components is done based on the Ethernet or Fast Ethernet connection. Each detector sub-system has its own DAQ management software communicating with the central DAQ manager and with the detector-environment-monitoring system. The communication software, called NSM (Network Shared Memory), provides two functionalities. One is data image sharing between any DAQ management software, and the other is message passing to the run management softwares, such as start and stop queue of runs from central DAQ manager, or run-halt signal due to emergency situation of the detectors from the monitoring system. Figure 2.38 shows an architecture of the NSM.

2.5 Offline Computing

Collected data by the Belle detector are analyzed at the offline computer farm. It is also an important task for the offline computer farm to make Monte Carlo simulation study. Required processing power of the offline computers amounts to 15000 SPECint95. Because this computational power cannot be achieved by a single CPU, we developed parallel processing scheme by multi-CPUs. We chose the Symmetric Multi Processor (SMP) architecture as a platform. The total storage capability of the offline computer farm is ~ 10 TB.

We developed own data processing framework, BASF/FPDA [43][44] (Belle AnalysiS Framework/Framework for Parallel Data Analysis), which are especially suitable for both the production of data summary tape (DST) and physics analysis. The recorded events are sequentially scanned by one process. Each bulk of around 10 events is distributed to another process. The number of processes are equal to the number of embedded CPUs on the SMP machine. The distributed bulks of the events are processed on CPUs in parallel, then returned to an another process to be stored onto a tape, a hard disk, or other storage media. The branching output path for the HBOOK [45] is also equipped. With the framework running on the offline computer farm, the production speed of the DST is $\sim 1\ \mathrm{fb}^{-1}$ per day.

The processing framework also provides a scheme of Monte Carlo generation. The bulks of seeds of a random numbers instead of events themselves are distributed to generator processes running on CPUs. A decay simulator is a program which generates particles tracing the decay chains according to a given manuscripts. The initial state of the particle generator is typically chosen from $\Upsilon(4S)$ or $q\overline{q}$. The decay simulator used by the Belle analysts is QQ98 originally developed by the CLEO group [46], and is modified by the Belle group [47]. The GEANT [48] is used as a full simulator of the detector performance in the Belle. It was originally developed at CERN. The GEANT simulator takes the decay simulation generated by QQ98 and simulates hit records on each detector associated to the generated particles. The output from the full simulator is stored in the same format as the real data so that one can use same analysis programs for both real and simulated data samples.

Chapter 3

Event Reconstruction

In this chapter, we describe the procedure of the event reconstruction to measure the CP asymmetry. We use $B^0 \to J/\psi K_S$ as the decay final state of the CP eigenstate in this study. First, we select hadronic events from the data samples accumulated with the Belle. Then we reconstruct J/ψ and K_S mesons, and finally we fully reconstruct B^0 mesons from them. For the reconstructed events, we apply the flavor tagging procedure to identify the associated B meson's flavor. We also reconstruct the proper-time difference from decay vertices of B^0 and $\overline{B}^0$ mesons.

3.1 Event Sample

Figures 3.1 show an integrated luminosity per day and a total integrated luminosity since the start of the Belle experiment. We use data accumulated since January 2000. The total triggered luminosity is 29.1 fb^{-1} at the $\Upsilon(4S)$ resonance, corresponding to 31.3 million $B\overline{B}$ pairs. The branching fraction of each decay is listed in Table 3.1 [49]. The expected number of B mesons decaying into $B^0 \to J/\psi K_S$ on this data sample is ~ 1100.

Table 3.1: Branching fractions for each decay mode. †: sum of B^0 and $\overline{B}^0$.

Mode	Branching ratio
$\Upsilon(4S) \to B^0\overline{B}^0$	$> 48\%$
$B^0, \overline{B}^0 \to J/\psi K_S$	$(4.5 \pm 0.6) \times 10^{-4} \times 2^\dagger$
$J/\psi \to \ell^+\ell^-$	$(11.81 \pm 0.14)\%$
$K_S \to \pi^+\pi^-$	$(68.61 \pm 0.28)\%$
Total	$(3.6 \pm 0.1) \times 10^{-5}$

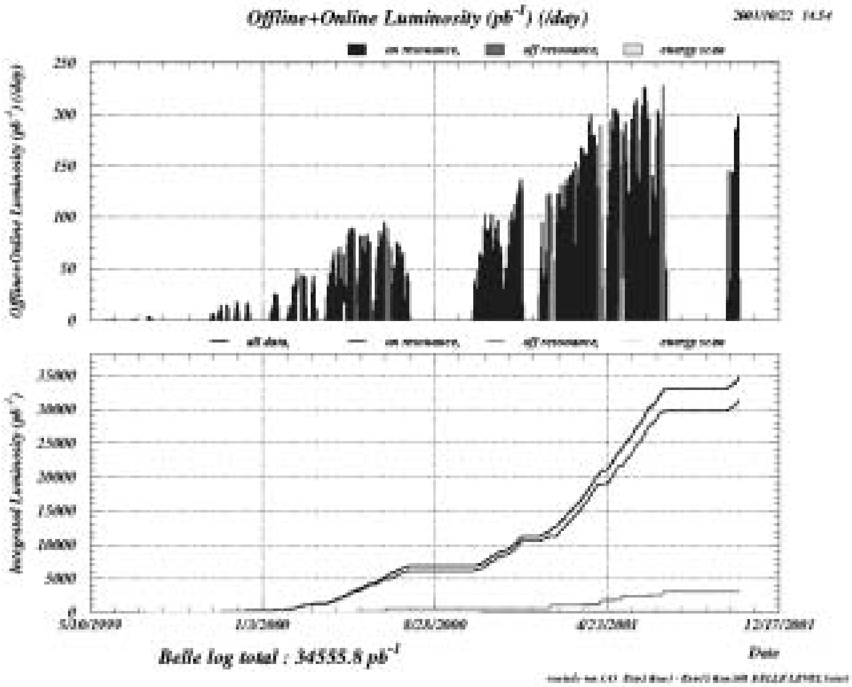


Figure 3.1: Integrated luminosity per day (top) and total integrated luminosity (bottom) since June 1999.

3.2 $B\overline{B}$ Event Selection

We select hadronic events from the data sample to eliminate the QED events, $\mu^+\mu^-$ process, two-photon process, and beam backgrounds. Here, "QED events" is referring to Bhabha or radiative Bhabha processes. The applied primary selection criteria follows:

- Events with only two tracks or less can be considered to come from the QED event. Even if they come from some hadronic processes, we cannot utilize them because of few tracks. Therefore, at least three "good" tracks are required, where a "good" track is defined by $|\delta r| < 2.0$ cm and $|\delta z| < 4.0$ cm measured from an nominal interaction point, and the transverse momentum of the "good" track must be greater than 100 MeV/c. The definitions of the "goodness" is loose enough not to discard physics events of any interest.

- At least two "good" clusters must be detected by the ECL within the volume of $-0.7 < \cos\theta < 0.9$. The "good" cluster is defined as the cluster with energy deposit of 100 MeV or larger. This is required because most clusters have a very shallow angle in the QED events.

- Total visible energy ($E_{\rm vis}$) that is a sum of good track momenta and "good" photon energies in an event, should be greater than 20% of the $\Upsilon(4S)$ energy. The "good" photon is defined as the "good" ECL energy clusters which is not associated to the charged particle trajectory reconstructed by the CDC.

- To suppress the QED events, the fraction of the total energy deposit in the ECL in an event must be between 10% and 80%. The QED events have larger energy deposit in the ECL, around 100%.
- The sum of the z components of the momentum of good tracks and good photons is required to be balanced around zero to eliminate the beam backgrounds. The absolute quantity of the sum must be smaller than a half of the $\Upsilon(4S)$ energy.
- The "event-primary-vertex", which is formed by all good tracks, must be between 1.5 cm and 3.5 cm from the interaction point in the r-ϕ and r-z planes, respectively. This is also required to reject the beam background events, because the background events are not originated from the interaction region.

From the Monte Carlo simulation, we find the above selection criteria retains more than 99% of $B\overline{B}$ events, while keeping the contamination of non-hadronic components smaller than 5%.

To suppress continuum $e^+e^- \to q\overline{q}$ process, we applied selection of $R_2 \equiv H_2/H_0 \leq 0.5$, where H_0 and H_2 are defined as zeroth and second order Fox-Wolfram moments [50]. R_2 is explicitly defined as

$$R_2 \equiv \frac{\sum_i^N \sum_j^N \left[|\vec{p_i}||\vec{p_j}| \cdot (3\cos^2\phi_{ij} - 1) \right]}{2\sum_i^N \sum_j^N \left[|\vec{p_i}||\vec{p_j}| \right]}, \tag{3.1}$$

where $\vec{p_i}$ is a momentum vector of i-th particle, ϕ_{ij} is an angle between $\vec{p_i}$ and $\vec{p_j}$ in the cms, and N is a total number of particles in the event. In jet-like events such as $e^+e^- \to q\overline{q}$, most $\cos\phi_{ij}$ tend to get $\sim \pm 1$ and the numerator and the denominator in equation (3.1) becomes similar number. R_2 becomes 1 in the limits of narrowest two-jet event. On the other hand, spherical events like B decays becomes $R_2 \simeq 0$. The efficiency loss due to R_2 selection is estimated to be smaller than 0.5% [51]. The R_2 distribution for the $B\overline{B}$ events and the continuum events are shown in Figure 3.2.

3.3 $B^0 \to J/\psi K_S$ Reconstruction

The reconstruction of $B^0 \to J/\psi K_S$ is described step by step in this section. J/ψ is reconstructed from two identified leptons and K_S is reconstructed from two charged

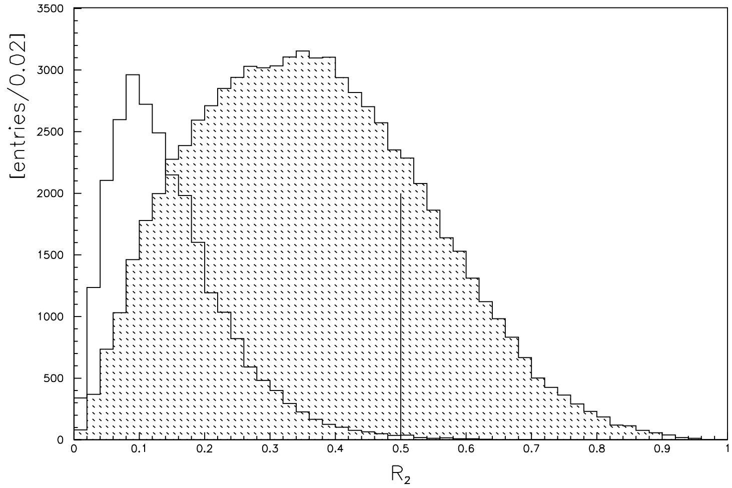


Figure 3.2: R_2 distributions. The white histogram represents $B\overline{B}$ events and the hatched histogram represents continuum events. The distributions are obtained from Monte Carlo.

particles. These intermediate mesons are combined to reconstruct B^0. In the following subsections, we describe the J/ψ, K_S, and B^0 reconstructions.

3.3.1 Reconstruction of J/ψ

J/ψ is reconstructed using the decay modes going into two leptons, $J/\psi \to \mu^+\mu^-$ and $J/\psi \to e^+e^-$.

J/ψ reconstruction using $J/\psi \to \mu^+\mu^-$ decay starts from searching for oppositely charged particles. Since reconstruction of $B^0 \to J/\psi K_S$ mode is free from background, we require loose muon identification to gain the reconstruction efficiency. At least one muon is required to have the muon likelihood ratio greater than 0.1. Remaining muon is required to satisfy much looser criterion that the energy deposit in the ECL be consistent with minimum ionizing, *i.e.* between 100 MeV and 300 MeV. The invariant mass of $\mu^+\mu^-$ pair is calculated with the assumption that charged particles have the muon mass ($M_\mu^\pm$ = 105.7 MeV/c^2):

$$M^2_{\mu^+\mu^-} = \left(\sqrt{M^2_{\mu^+} + |\vec{p_{\mu^+}}|^2} + \sqrt{M^2_{\mu^-} + |\vec{p_{\mu^-}}|^2}\right)^2 - |\vec{p_{\mu^+}} + \vec{p_{\mu^-}}|^2. \tag{3.2}$$

The invariant mass, $M_{\mu^+\mu^-}$, must be between $M_{J/\psi} - 60$ MeV/c^2 and $M_{J/\psi} + 36$ MeV/c^2, which corresponds to -5 and $+3.5$ times the standard deviation (σ) apart from the average J/ψ mass ($M_{J/\psi}$ = 3096.87 MeV/c^2), respectively. The invariant mass window is asymmetric due to missing energy carried out by radiative photon though the effect is much smaller than e^+e^- case. Figure 3.3 (top) shows an invariant mass distribution of $M_{\mu^+\mu^-}$, where arrows indicate the edges of the mass window.

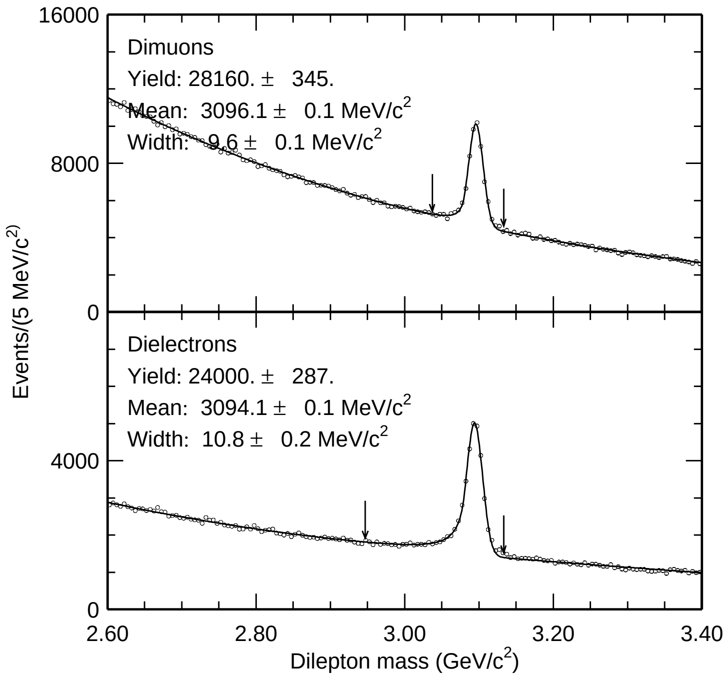


Figure 3.3: Invariant mass distributions for $J/\psi \to \mu^+\mu^-$ (top) and $J/\psi \to e^+e^-$ (bottom). Arrows indicate the invariant mass ranges.

J/ψ reconstruction with another leptonic decay, $J/\psi \to e^+e^-$, also uses oppositely charged electron-position pairs. As well as $J/\psi \to \mu^+\mu^-$ case, we apply loose requirements for the electron identification. At least one electron is required to have electron likelihood greater than 0.1. Remaining electron is required to have one of electron probabilities calculated with the dE/dx or E/p information being greater than 0.5. In the $J/\psi \to e^+e^-$ decay, radiative photon carries out the energy, and thus the reconstructed J/ψ mass distribution has tail in lower mass region. To take into account the radiative photon energy, all photons are searched for in the cone within ± 0.05 rad from the electron (or positron) track direction measured at the interaction point of the track (Figure 3.4). When such photons are found, their energies are added to the associated electron (position). The invariant mass of e^+e^- is calculated as

$$M^2_{e^+e^-} = \left(\sqrt{M^2_{e^+} + |\vec{p_{e^+}}|^2} + \sqrt{M^2_{e^-} + |\vec{p_{e^-}}|^2}\right)^2 - |\vec{p_{e^+}} + \vec{p_{e^-}}|^2, \tag{3.3}$$

where $\vec{p_{e^+}}$ and $\vec{p_{e^-}}$ are the corrected by the four-momentum of the radiative photons when such photons are found. Although the correction is applied, the invariant mass distribution has broad tail in the lower mass region, because photon is absorbed by the material in front of the ECL. The invariant mass of e^+e^- pair must be between $M_{J/\psi} - 150$ MeV/c^2 and $M_{J/\psi} + 36$ MeV/c^2, which corresponds to the -15σ and $+3.5\sigma$ range. Figure 3.3 (bottom) shows an invariant mass distribution of $M_{e^+e^-}$.

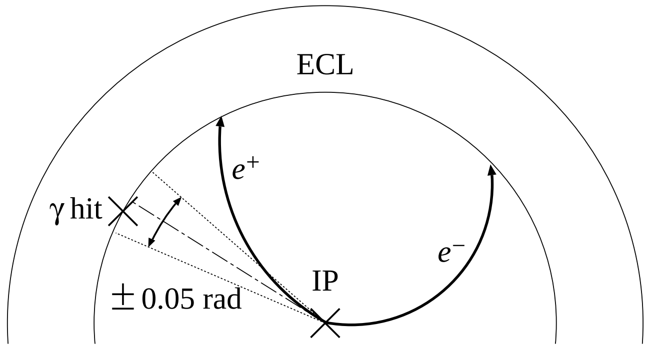


Figure 3.4: Illustration of search region for photons on the ECL for $J/\psi \to e^+e^-$ decay.

The J/ψ momentum in the center-of-mass frame must be below 2 GeV/c. After the invariant the mass selection, we calibrate the lepton momenta under the constraint that the invariant mass of $\ell^+ + \ell^-$ becomes the known J/ψ mass. It improves the reconstructed B^0 momentum resolution.

3.3.2 Reconstruction of K_S

K_S is reconstructed from oppositely charged two pions. Because K_S can be reconstructed with small backgrounds, particle identification for pion tracks are not performed to keep high efficiency. The initial decay vertex is reconstructed as the middle point of two helices at the crossing point in the r-ϕ plane. Then the parameters for the two helices are recomputed at the initial decay vertex while taking into account the effect of the energy loss and multiple scattering. The vertex position is reconstructed using improved helices. When the vertex reconstruction succeeds, the initial decay vertex is replaced with the newly reconstructed vertex at which the invariant mass of K_S candidate is calculated. Since K_S has relatively long lifetime than other mesons. it traverses around 7.5 cm in the r-ϕ plane in the Belle experiment. We reject the K_S candidates with short flight length. When none of two charged tracks has the associated SVD hits, the ϕ coordinate of the $\pi^+\pi^-$ vertex point is required to agree with the three-momentum direction of the $\pi^+\pi^-$ candidate within 0.1 rad. When one has SVD hits and the other does not, at least one of the distances from the track to the interaction point must be larger than 250 μm in the r-ϕ plane to suppress tracks from the interaction points[1]. The invariant mass of $\pi^+\pi^-$, $M_{\pi^+\pi^-}$, is calculated assuming the average pion mass, $M_{\pi^\pm} = 139.6$ MeV/c^2, for

[1] In case that both tracks have associated SVD hits, we require only common decay vertex for two tracks (distance of the two tracks z direction should be smaller than 1.0 cm) rather than the flight length.

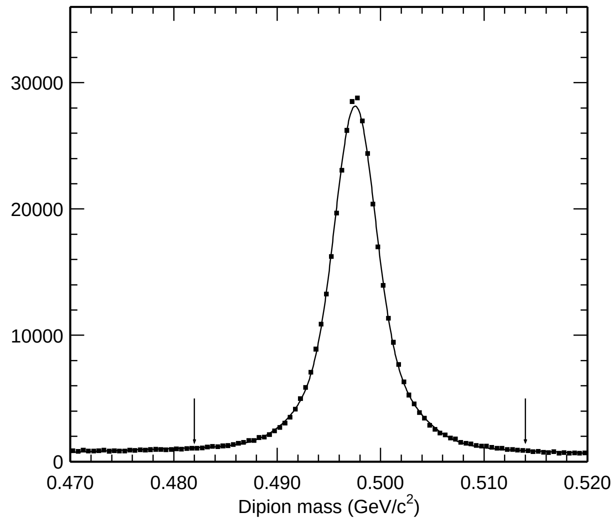


Figure 3.5: Invariant mass distribution of $K_S \to \pi^+\pi^-$. The arrows indicate the invariant mass range.

the charged tracks:

$$M^2_{\pi^+\pi^-} = \left(\sqrt{M^2_{\pi^+} + |\vec{p_{\pi^+}}|^2} + \sqrt{M^2_{\pi^-} + |\vec{p_{\pi^-}}|^2}\right)^2 - |\vec{p_{\pi^+}} + \vec{p_{\pi^-}}|^2. \tag{3.4}$$

The invariant mass should be between 482 and 514 MeV/c^2, which corresponds to three standard deviation away from the world average of the K_S mass. The selection listed above retains 99.7% of K_S signals. Figure 3.5 shows the invariant mass distribution of K_S candidates. Arrows in the Figure represent the mass-window edge. The detail of the K_S reconstruction is described in [52].

3.3.3 Reconstruction of B^0

B^0 is reconstructed by the combination of the reconstructed J/ψ and K_S. The energy and the momentum of the reconstructed B^0 candidate are calculated as

$$E_B = E_{J/\psi} + E_{K_S}, \qquad \vec{p_B} = \vec{p_{J/\psi}} + \vec{p_{K_S}}. \tag{3.5}$$

For the B^0 reconstruction, we calculate the energy difference, ΔE, and the beam-energy constraint mass, $M_{\rm bc}$. The energy difference is defined as $\Delta E \equiv E_B^{\rm cms} - E_{\rm beam}^{\rm cms}$ and the beam-energy constraint mass is defined as $M_{\rm bc} \equiv \sqrt{(E_{\rm beam}^{\rm cms})^2 - |p_B^{\rm cms}|^2}$, where $E_{\rm beam}^{\rm cms}$ denotes the beam energy in the center-of-mass frame of $\Upsilon(4S)$, and $E_B^{\rm cms}$ and $p_B^{\rm cms}$ are the energy and the momentum of the reconstructed B^0 in the same frame. The scatter plot of $M_{\rm bc}$ and ΔE is shown in Figure 3.6 together with the projections onto each axis. B^0 candidates are selected by requiring $5.2694 < M_{\rm bc} < 5.2894$ GeV/c^2 corresponding to $\pm 3.5\sigma$ from average B^0 mass ($M_{B^0} = 5.2794$ GeV/c^2), and $|\Delta E| < 40$ MeV corresponding

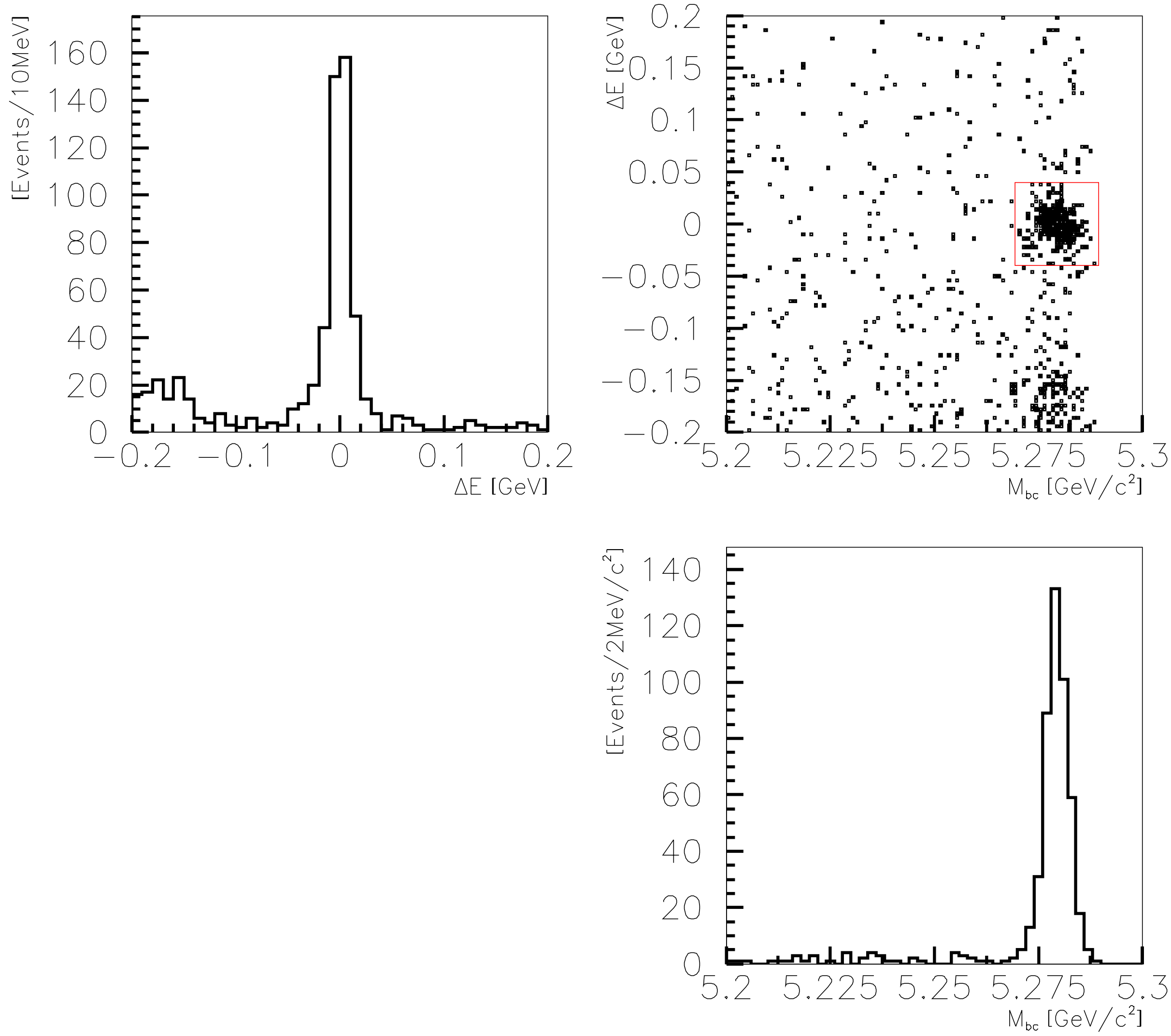


Figure 3.6: Scatter plot of ΔE versus $M_{\rm bc}$. The box in the upper right figure represents the signal region. The upper left figure represents the projections onto ΔE axis with $5.2694 < M_{\rm bc} < 5.2894$ MeV/c^2. The lower figure represents the projections onto $M_{\rm bc}$ axis with $|\Delta E| < 40$ MeV.

to 4σ. When one event contains more than one B^0 candidate, the candidate with the least χ^2 among them is selected, where χ^2 is defined using ΔE and $M_{\rm bc}$ as

$$\chi^2 \equiv \left(\frac{\Delta E}{\sigma_{\Delta E}}\right)^2 + \left(\frac{M_{\rm bc}}{\sigma_{M_{\rm bc}}}\right)^2 \tag{3.6}$$

In the signal region, we have 457 candidates of $B^0 \to J/\psi K_S$. A crowd of events is seen in lower right area of the ΔE-$M_{\rm bc}$ scattered plot. The crowd is a contamination from $B^0 \to J/\psi K^{*0}$ decay. The events of $B^0 \to J/\psi K^{*0}(K_S\pi^0)$ are reconstructed as $B^0 \to J/\psi K_S$ by missing π^0. The number of the backgrounds in the signal region is estimated by a two-dimensional fit on the ΔE-$M_{\rm bc}$ distribution. The background fraction in the signal region is estimated to be 0.038 ± 0.010. The detailed estimation of the signal fraction is described in Section 4.2.1.

3.4 Flavor Tagging

To measure the CP asymmetry, the flavor of $B_{\rm tag}$ must be determined. We discuss the flavor tag procedure in this section.

Following processes are used to determine the flavor of the $B_{\rm tag}$: primary lepton in the semileptonic decays from $b \to c\ell^-\overline{\nu}_\ell$, secondary lepton in $b \to c \to \ell^+$ decays, fast pions, which reflects the charge of virtual W in $b \to c + W^-$, slow pions coming from $D^{*\pm}$ whose charge reflects a charge of c, and kaons and Λ from cascade decays of $b \to c \to s$. All of remaining tracks after the $B^0 \to J/\psi K_S$ reconstruction, except for poorly reconstructed tracks such that whose impact parameters with respect to the interaction point are greater than 10 cm in z direction or greater than 2 cm in the r-ϕ plane, are used to determine the flavor.

Even high quality of flavor tagging capability, we have to consider wrong identifications of the flavor. Let the determined flavor be q, where $+1$ is assigned for events with $B_{\rm tag} = B^0$ and *vice versa*, and the wrong tag fraction be w $(0 \le w \le 1)$. As we discussed in Section 1.4.2, the Δt distribution is represented with q and w as

$$f_{\ell^\pm}(\Delta t;\ \sin 2\phi_1) = \frac{1}{2\tau_{B^0}} \mathrm{e}^{-\frac{|\Delta t|}{\tau_{B^0}}} \Big[1 + (1-2w)\cdot q \cdot \sin 2\phi_1 \cdot \sin \Delta M \Delta t\Big]. \tag{3.7}$$

We always observe the CP violation through "$(1-2w)\cdot\sin 2\phi_1$" instead $\sin 2\phi_1$, and thus the determination of precise w is one of essential issues in the measurement of $\sin 2\phi_1$.

The efficiency of the flavor tagging must be discussed with care. We have to consider both the tagging efficiency and wrong tagging probability. Even if we can achieve high efficiency by some method, it might be meaningless if most events have $w = 0.5$, because they do not carry flavor information. We have to quote the effectiveness to discuss the flavor tagging performance. First, we define a overall (naive) flavor tagging efficiency, ε, by

$$\varepsilon = \frac{\text{Number of flavor-identified events}}{N}, \tag{3.8}$$

where N is a number of total events. As we have seen in equation (3.7), the measured asymmetry, a'_{CP}, is $(1-2w)\cdot\sin 2\phi_1$. Therefore, the significance of a'_{CP} is represented as

$$\text{significance} = \frac{a'_{CP}}{\delta(a'_{CP})} \propto (1-2w)\cdot\sqrt{\varepsilon N}. \tag{3.9}$$

Thus, the "effective tagging efficiency", ε_{eff}, can be defined as

$$\varepsilon_{\text{eff}} \ \equiv \ \varepsilon(1-2w)^2. \tag{3.10}$$

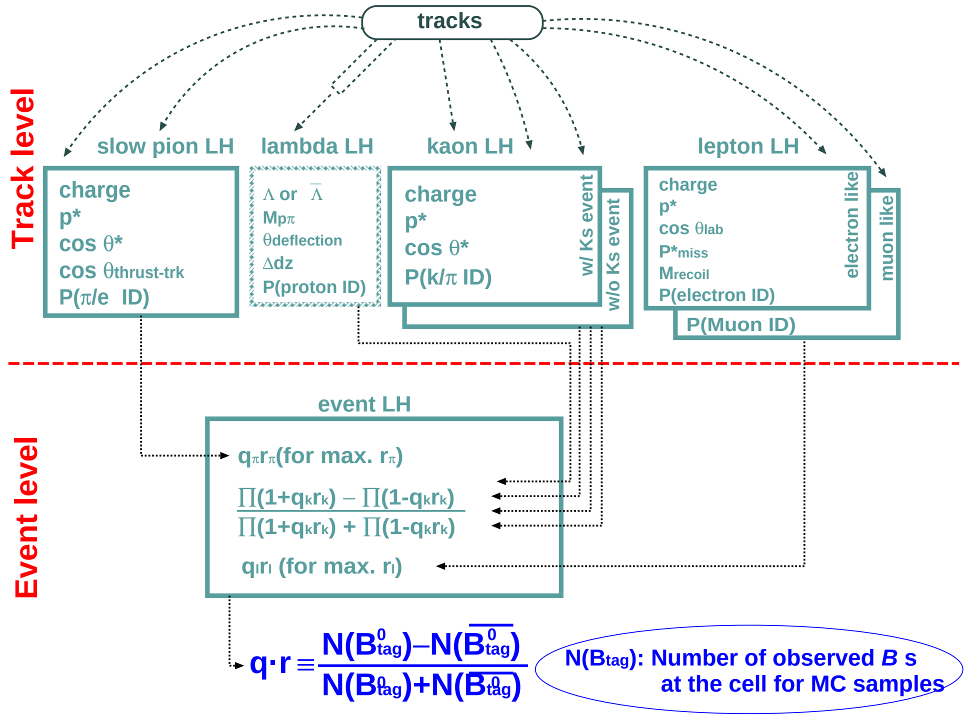
In the following subsections we describe the flavor tagging method, then we discuss the estimation of the wrong tag fraction using flavor specific decays in real data, and we determine the effective tagging efficiency.

3.4.1 Flavor Tagging Method

The flavor of B_{tag} is determined from the charges of the tracks in B_{tag} decay. The flavor tagging consists of two stages, called layers: track-layer and event layer. The track-layer has four categories in parallel with different purpose: lepton like track, slow pion like track, Λ like track, and kaon like track. To determine the flavor, we use multi-dimensional likelihood method to utilize as much discriminants as possible [53]. In the track-layer, each track is given "$r\cdot q$" value, where r is a flavor tagging quality defined as $r \equiv 1-2w$. For every track category, we construct a look-up table of the multi-dimensional likelihood containing "$r \cdot q$" values by combining several discriminants defined for each category based on the Monte Carlo simulation. The second layer is the event layer. The event layer collects the $r\cdot q$ information from the track-layers and determines the final $r\cdot q$ value. Figure 3.7 illustrates the schematic view of the flavor tagging method. Because of the likelihood approach, we can obtain w in event by event. We utilize the event-by-event w to change the weight of the event for the $\sin 2\phi_1$ determination.

Track Layer

Each of the tracks from B_{tag} is classified into one of the track categories. Among B_{tag} decay products, a track pairs that form K_S invariant mass and a track consistent with photon conversion, whose invariant mass is smaller than 100 MeV/c^2 and transverse momentum is larger than 1.5 GeV/c, are never used for the flavor tagging.



$$q\cdot r \equiv \frac{N(B^0_{tag}) - N(\overline{B^0_{tag}})}{N(B^0_{tag}) + N(\overline{B^0_{tag}})}$$

Figure 3.7: Schematic view of the multi-dimensional likelihood method for the flavor tagging.

- Lepton category

  The lepton category is intended to accept the high and middle momentum leptons together with the fast pions. All tracks that come to this category is required to have the cms momentum (p^*) larger than 400 MeV/c. This category consists of two sub-categories according to its lepton flavor. When the likelihood ratio of $e^\pm/K^\pm$ is greater than 0.8, the track is categorized to electron sub-category. The track with $p > 400$ MeV/c and with larger likelihood ratio of $\mu^\pm/K^\pm$ than 0.95 is passed to muon sub-category. Because the lepton identification quality depends on the track polar angle ($\theta_{\rm lab}$), this quantity is also examined. The fast leptons are considered to come from the semileptonic B decays and the middle momentum leptons are considered to come from the semileptonic D decays. We discriminate fast leptons from middle momentum leptons by the p^*, recoil mass ($M_{\rm recoil}$), and missing momentum in the cms ($p^*_{\rm miss}$), where the recoil mass is defined as the invariant mass made up by all $B_{\rm tag}$ tracks except for this "lepton".

  We construct the look-up table of the multi-dimensional likelihood with the p^*, the lepton likelihood, $\cos\theta_{\rm lab}$, $M_{\rm recoil}$, $p^*_{\rm miss}$, and lepton charge. The numbers of bins for each discriminant are 11, 4, 6, 10, 6, and 2, respectively. Thus, the total number of bins of the look-up table is $11 \times 4 \times 6 \times 10 \times 6 \times 2 = 31680$. For each of 31680 bins, we assign the $r \cdot q$ value obtained as

$$r \cdot q = \frac{N_{B^0} - N_{\overline{B}^0}}{N_{B^0} + N_{\overline{B}^0}}, \tag{3.11}$$

where N_{B^0} and $N_{\overline{B}^0}$ are numbers of B^0 and $\overline{B}^0$ in the bin estimated from the Monte Carlo simulation. The look-up tables of other categories are constructed in the similar ways.

- Slow-pion category
  The slow-pion category is intended to accept the charged particles from $D^{*\pm}$ decays. When p^* is smaller than 250 MeV/c and it is not positively identified as kaon, it is classified to the slow-pion category. The major contamination for this category is other pions than $D^{*\pm}$ decay. Electrons that come from photon conversion are also the background of this category, although the such electrons are eliminated by the invariant mass selection of $M_{e^+e^-} < 100$ MeV/c^2. In the slow pion category, we use following discriminant variables: momentum and polar angle in the laboratory frame ($p_{\rm lab}$ and $\theta_{\rm lab}$), an angle between the thrust axis of $B_{\rm tag}$ decay tracks and the candidate slow pion momentum in the cms ($\cos\alpha_{\rm thr}$), likelihood ratio of $e^\pm/\pi^\pm$ based on dE/dx measurement, and the track charge. Since the direction of slow pion from $D^{*\pm}$ follows the direction of the $D^{*\pm}$, it also tends to follows the direction of thrust axis. Therefore, indirect p^* taken from $p_{\rm lab}$ and $\theta_{\rm lab}$ and $\cos\alpha_{\rm thr}$ enables us to discriminate the slow pions from $D^{*\pm}$ decay and other pions. The measurements of dE/dx is used to effectively separate the slow pions from those electrons.

- Λ category
  We reconstruct Λ from $\Lambda \to p\pi^-$ decay. If a track forms Λ with other track, it is passed to the Λ track category. The discriminant variables are: an invariant mass of reconstructed Λ, a mismatch in the z direction of two tracks at Λ vertex (δz), an angle difference between Λ momentum vector and the direction to decay vertex point of Λ from the interaction point, proton identification quality, and its flavor (Λ or $\overline{\Lambda}$). δz indicates whether two tracks are originated from common Λ vertex. Because Λ lifetime is rather long ($c\tau = 7.89$ cm), the smaller angle between the reconstructed Λ direction and the flight direction indicates more possibility of Λ.

- Kaon category
  The remaining tracks that are not classified to any of above categories and are not positively identified as proton, are passed to kaon track category. The kaon category is intended to take kaons and some fast pions. The kaon track category is divided into two sub-categories according to the existence of K_S in the event. Because K_S can also carry out the strangeness from the cascade decay of $b \to c \to s$, the correlation of kaon charge and $B_{\rm tag}$ flavor becomes much weaker when K_S exists in

a event. The discriminant variables used for this category are: p^*, $\theta_{\rm lab}$, likelihood ratio of $K^\pm/\pi^\pm$ and the track charge. First, three variables are expected to separate kaons from pions.

Event Layer

The event layer gathers the information on the flavor from each of the track categories, and determines the final $r\cdot q$ value. The event layer uses following discriminants: best "$r\cdot q$" from the lepton category ($(r\cdot q)_\ell$), best "$r\cdot q$" from the slow-pion category ($(r\cdot q)_{\pi_{\rm s}}$), and "$r\cdot q$" by the product of all outputs from Λ and kaon categories ($(r\cdot q)_K$). Because the number of lepton indicating the $B_{\rm tag}$ flavor is expected to be one in a single event, only best "$r\cdot q$" value from the lepton category is taken. It is also true for the slow-pion category. Unlike lepton category or slow-pion category, the signature of the $B_{\rm tag}$ flavor show up in a sum of the strangeness in a event. Therefore, the third discriminant is calculated by the product for all tracks in this category together with outputs from the Λ category. The product is expressed as

$$(r\cdot q)_K = \frac{\prod_i (1 + r_i\cdot q_i) - \prod_i (1 - r_i\cdot q_i)}{\prod_i (1 + r_i\cdot q_i) + \prod_i (1 - r_i\cdot q_i)}. \tag{3.12}$$

The final $r\cdot q$ is also obtained by referring the multi-dimensional look-up table. The numbers of bins are 25, 19, and 35 for $(r\cdot q)_\ell$, $(r\cdot q)_{\pi_{\rm s}}$, and $(r\cdot q)_K$, respectively. The content "$r\cdot q$" in the look-up table are obtained in the same way as described in the lepton category, although it is obtained using independent Monte Carlo sample of those used to construct the look-up tables for the track categories.

The overall efficiency that is defined as the fraction of $r \neq 0$ events against total input events is 99.6% in the Monte Carlo simulation.

3.4.2 Measurement of Wrong Tagging Probability

We test consistency between the r obtained from the Monte Carlo simulation and r estimated from flavor specific B decays in the real data. We use r from the Monte Carlo simulation to categorize the tagged events. The event is classified into six quality levels according to the computed r values: $0 < r \leq 0.25$, $0.25 < r \leq 0.5$, $0.5 < r \leq 0.625$, $0.625 < r \leq 0.75$, $0.75 < r \leq 0.875$, and $0.875 < r \leq 1$. For each r regions, we estimate $r_{\rm data}$ from the real data of following decay modes: $B^0 \to D^{*-}\ell^+\nu$, $B^0 \to D^-\pi^+$, $B^0 \to D^{*-}\pi^+$, and $D^{*-}\rho^+$ in the real data, where $D^0 \to K^-\pi^+$, $D^0 \to K^-\pi^+\pi^0$, and $D^0 \to K^-\pi^+\pi^+\pi^-$. These decays are flavor-specific and we can know their flavors independently of the flavor tagging method. The detailed reconstruction procedures for these modes are

Table 3.2: Combined wrong tag fractions, w_ℓ, from the semileptonic and hadronic B decays, and the event fractions, f_ℓ, obtained from $J/\psi K_S$ Monte Carlo simulation, for each r interval. The statistical errors dominate over the w uncertainties.

ℓ	r_ℓ	w_ℓ	stat.	syst.	f_ℓ
1	0.000 – 0.250	0.465	± 0.008	$^{+0.005}_{-0.004}$	0.405
2	0.250 – 0.500	0.352	± 0.013	$^{+0.008}_{-0.006}$	0.149
3	0.500 – 0.625	0.253	± 0.016	$^{+0.013}_{-0.025}$	0.081
4	0.625 – 0.750	0.176	± 0.014	$^{+0.017}_{-0.009}$	0.099
5	0.750 – 0.875	0.110	± 0.013	$^{+0.018}_{-0.008}$	0.123
6	0.875 – 1.000	0.041	$^{+0.009}_{-0.008}$	± 0.006	0.140

described in Appendix A. Since we know the flavors of both B mesons for these decays, we can observe the time evolution of the neutral B-meson pair with the opposite flavor (OF) and the same flavor (SF) that are given by

$$\mathcal{P}_{\rm OF}(\Delta t) \propto 1 + (1-2w)\cos(\Delta M \Delta t), \tag{3.13}$$

$$\mathcal{P}_{\rm SF}(\Delta t) \propto 1 - (1-2w)\cos(\Delta M \Delta t). \tag{3.14}$$

The time dependent OF-SF asymmetry is

$$A_{\rm mix} \equiv \frac{\rm OF - SF}{\rm OF + SF} = (1-2w)\cos(\Delta M \Delta t). \tag{3.15}$$

We can obtain the $w_{\rm data}$ by measuring the amplitude of the OF-SF asymmetry. The detailed estimation procedure of $w_{\rm data}$ is described in Appendix B. The summary for the wrong tag fraction is listed in Table 3.2. The uncertainties for w_ℓ ($w_{\rm data}$ for ℓ-th r region) are dominated by the statistics. Figure 3.8 shows the estimated dilution factor $(1-2w_{\rm data})$ obtained from real data as a function of the average $(1-2w)$ in each r-category obtained from the Monte Carlo simulation. As shown in the figure, the measured $(1-2w_{\rm data})$ tend to be slightly smaller than the values based on the Monte Carlo simulation. The overall tendency, however, agrees well and the slight difference does not introduce a systematic error in the measurement of $\sin 2\phi_1$ because we use measured w_ℓ values. The event fraction assigned with $r \neq 0$ is estimated to be 99.7% for real data and the effective tag efficiency, $\varepsilon_{\rm eff}$, is measured to be $\varepsilon_{\rm eff} = (27.0 \pm 0.8({\rm stat})^{+0.6}_{-0.9}({\rm syst}))\%$ that is defined as

$$\varepsilon_{\rm eff} = \sum_{\ell}^{\rm level} f_\ell (1-2w_\ell)^2, \tag{3.16}$$

where ℓ denotes the level of flavor tagging quality, and f_ℓ is the efficiency for each level, estimated with $B^0 \to J/\psi K_S$ signal Monte Carlo. The estimated f_ℓ is listed in Table 3.2.

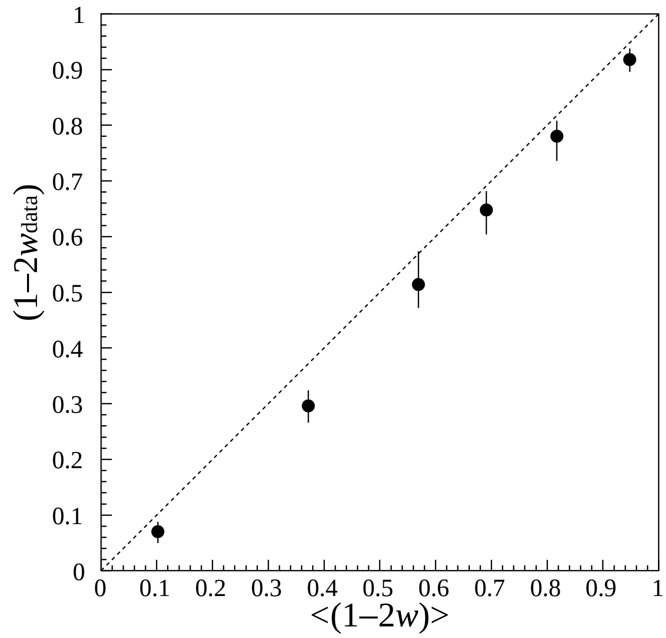


Figure 3.8: Measured dilution r_{data} as a function of Monte Carlo simulated $\langle r \rangle$ from the algorithm.

We check for a possible bias in the flavor tagging by measuring the effective tagging efficiency for B^0 and $\overline{B}^0$ self-tagged samples separately. We find no statistically significant difference.

Four event candidates of $B^0 \to J/\psi K_S$ are lost because their flavors cannot be determined. Number of remaining B^0 candidates is 451.

3.5 Reconstruction of Proper-Time Difference

It is a crucial issue to measure proper-time difference of B^0-$\overline{B}^0$ pair, Δt, throughout this analysis, whereas we can neither know the generation time of B^0-$\overline{B}^0$ pair nor the flight lengths of them. We can only measure the distance of two decay vertices, Δz. We defined Δz as the decay vertex of B decayed into $J/\psi K_S$ (B_{CP}) minus one of B decayed into unspecified decay mode (B_{tag}):

$$\Delta z \equiv z_{CP} - z_{\text{tag}}. \tag{3.17}$$

Since the mass of $\Upsilon(4S)$ is quite close to a sum of two B^0 masses, B^0-$\overline{B}^0$ pair is produced almost at rest in the center-of-mass system of $\Upsilon(4S)$. Therefore, in the laboratory frame, both B^0 and $\overline{B}^0$ are given almost same momentum. In addition, they are boosted in z direction, and the motion in the r-ϕ plane is approximately negligible. Approximating

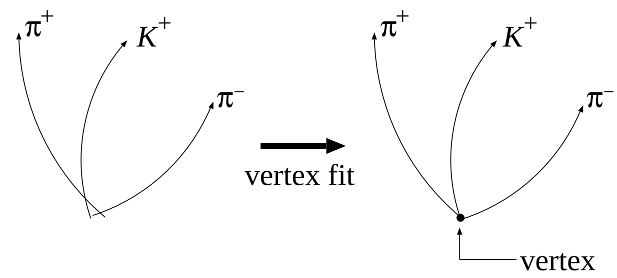


Figure 3.9: Schematic drawing of the vertex reconstruction.

the motion of B as $(\beta\gamma)_\Upsilon$, the proper-time difference can be calculated as

$$\Delta t \simeq \frac{\Delta z}{(\beta\gamma)_\Upsilon}. \tag{3.18}$$

The decay vertices of B_{CP} and $B_{\rm tag}$ are obtained from the kinematic fit based on the least χ^2 method with Lagrange multiplier technique [54]. The constraint for the Lagrange multiplier is that the trajectories of charged tracks must draw helices in a magnetic field according to their momenta and charges. A schematic drawing of the fit is shown in Figure 3.9. Track positions and momenta are tuned according to measurement errors so that all tracks pass a certain point. The point is regarded as the production vertex of the tracks. The estimated error of the reconstructed vertex is propagated from the errors of the track parameter determination, *i.e.* momentum and position. Therefore, even there is large displacement of the tracks from the estimated vertex position, the estimated error of the vertex can still be small when the track parameters are determined well. The displacement reflects onto χ^2 of the vertex reconstruction. The χ^2 is defined as

$$\chi^2 \equiv \vec{\lambda}^{\rm T} V_D^{-1} \vec{\lambda}, \tag{3.19}$$

where $\vec{\lambda}$ is difference of track parameters between before and after the vertex reconstruction, and V_D is error matrix of $\vec{\lambda}$.

The decay vertex of B_{CP} is replaced by the decay vertex of J/ψ because of its negligibly short flight length. The decay vertex of $B_{\rm tag}$ is reconstructed using all remaining tracks after the reconstruction of $B_{CP} \to J/\psi K_S$. We have to be aware that $B_{\rm tag}$ may decay into charmed mesons. The lifetimes of D^0 and D^+ are 413 ± 3 fs and 1051 ± 13 fs, respectively, which are long enough to be detected by the Belle detector. Because the charmed mesons typically have positive motion for z direction, the reconstructed vertex of $B_{\rm tag}$ might be shifted forward, which causes negative shift in Δz and Δt from the true values. To make the vertex resolution better and to make the reconstruction efficiency higher, we require the B_{CP} and $B_{\rm tag}$ decay vertices should be consistent to the nominal B decay position in the r-ϕ plane. The nominal decay position is represented by the beam interaction point convoluted with the nominal flight length of B meson.

In the following subsections, first we discuss the reconstruction of B decay position including the interaction point construction. Then we describe the decay vertex reconstructions of B_{CP} and $B_{\rm tag}$.

3.5.1 Reconstruction of *B* Decay Position

We reconstruct the B_{CP} and $B_{\rm tag}$ vertices with a constraint to B decay position to improve the reconstructed vertex resolution. In addition to that, the vertices can be reconstructed even only single track by obtaining the crossing point of the track and the B decay point distribution. Although the resolution of single-track vertices is much worse than that of multiple-track vertices, the statistical errors of $\sin 2\phi_1$ becomes better. The B decay position is obtained by the interaction point of e^+e^- collision convoluted with the distribution of the B decay point. In the following paragraphs, we describe the reconstruction of interaction point, then we describe the effect of B flight length to the kinematic vertex reconstruction. In the end of this subsection, we describe the performance of the constraint to the B decay point.

The accelerator condition changes injection by injection of the beams, and the interaction points also shifts injection by injection. Moreover, the positions move even during the run. Thus, we construct the interaction points periodically through the runs, and we update the interaction point profile every 60000-events when one run holds greater than 60000 events. The interaction points of e^+e^- collisions are obtained for each event in three-dimensional coordinates by the vertex reconstruction with all reconstructed tracks in the event. The distributions of the interaction points are summarized injection by injection, and then fitted to rotated three-dimensional single Gaussian. The distribution of the reconstructed interaction point can be considered to consist of interaction point fluctuation and the resolution of the vertex reconstruction. However, the fluctuation in y ($\sigma_y^{\rm IP}$) that is directly measured by a beam position monitor is much better than the vertex resolution ($\sigma_y^{\rm vtx}$). Therefore, the reconstructed width of the interaction point distribution in y (σ_y) can be considered $\sigma_y^{\rm vtx}$ approximately. The interaction point distribution in x ($\sigma_x^{\rm IP}$) is recalculated as $\left(\sigma_x^{\rm IP}\right)^2 = \sigma_x^2 - \sigma_y^2$ with assuming $\sigma_x^{\rm vtx} \simeq \sigma_y^{\rm vtx}$. Since the interaction point distribution in z is much worse than the vertex resolution, we do not make special treatments. Typical fluctuations of the interaction point are 100 μm in x, 5 μm in y (accelerator information), and $3 - 4$ mm in z. The rotational angle approximately corresponds to the half of the beam crossing angle, 11 mrad (Figure 3.10). Figures 3.11 show distributions of interaction points. The mean values of the position distributions are regarded as the interaction point.

The effect of B meson lifetime is also taken into account for the kinematic vertex fit

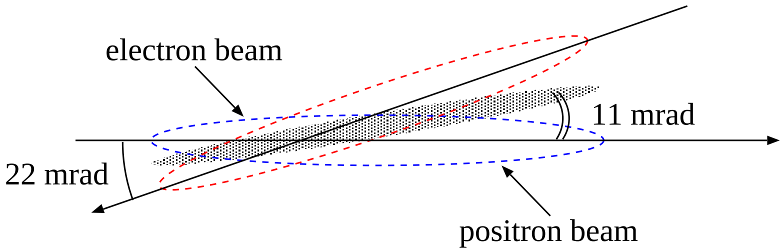


Figure 3.10: Beam crossing angle, ±11 mrad, reflects onto the rotation of Gaussian in 11 mrad. The rotated gray ellipse represents the interaction point profile.

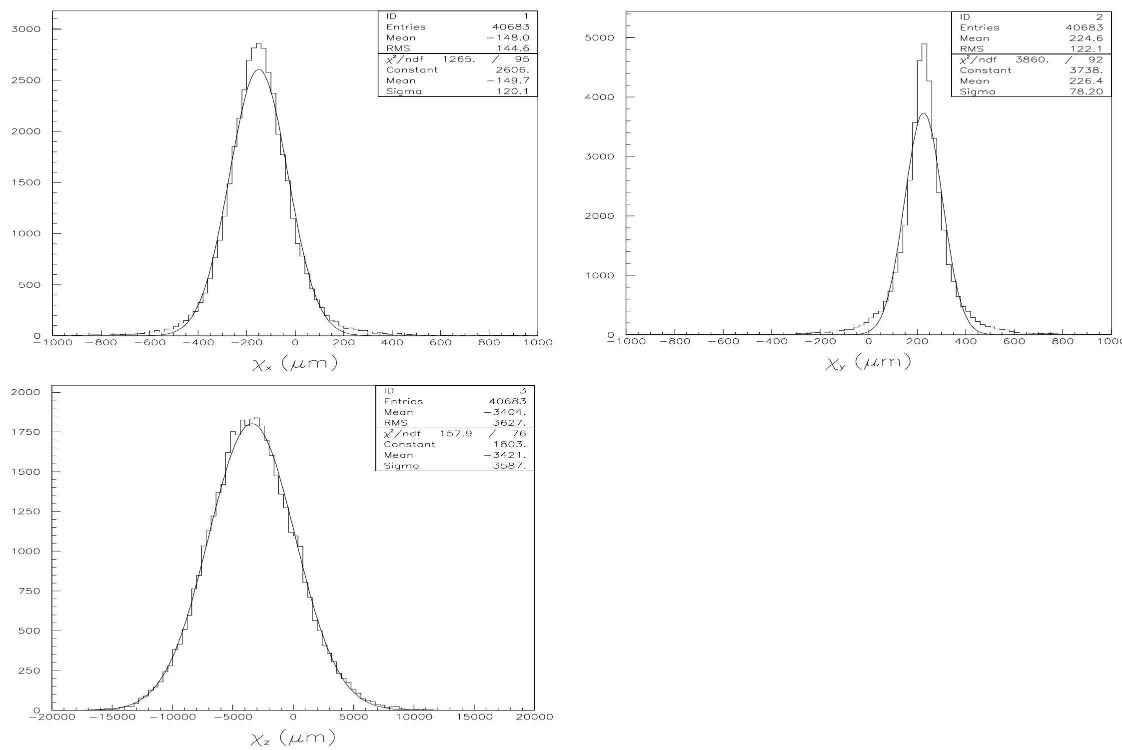


Figure 3.11: Interaction point distribution in x, y, and z coordinates obtained from real data. The mean of the distribution in x, y, and z represents the nominal interaction point. The standard deviations of the interaction point distributions are calculated as described in the text.

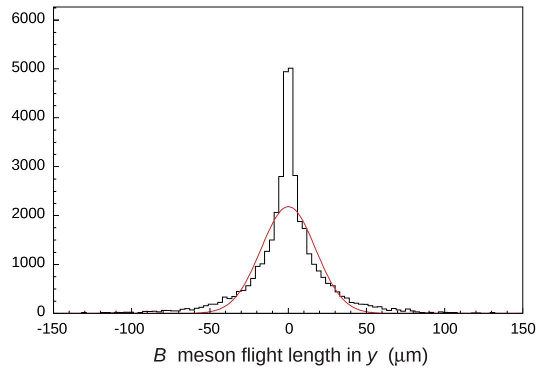


Figure 3.12: Distribution for B meson flight length along with y direction. A fit to a single Gaussian is also superimposed.

procedure. Since the interaction point is well determined in y direction, we evaluate the B meson flight length in y direction obtained from Monte Carlo simulation. Figure 3.12 shows the distribution of B meson flight length along with y direction. A fit to single Gaussian is superimposed. The r.m.s. of the distribution is 25 μm and the σ of the fitted Gaussian is 17 μm. We express the "smearing" due to the B meson flight as a single Gaussian with the standard deviation of the average of the two values, $(25 + 17)/2 = 21$ μm.

Δz resolution is improved by the B decay point constraint from 182 μm to 148 μm for multiple track vertices. We salvage 29% of reconstructed B^0 candidate that have only single track in B_{CP} side or $B_{\rm tag}$ side or in both. The Δz resolution for the salvaged events is typically 273 μm. The reconstruction of Δz, *i.e.* z_{CP} and $z_{\rm tag}$ is described in following subsections.

3.5.2 Vertex Reconstruction of B_{CP}

The vertex of B_{CP} is reconstructed using one or two leptons from J/ψ decays. We require the tracks to have the associated hits on the SVD so that we can have accurate vertex. At least one hit and two hits must be observed in the r-ϕ and r-z planes, respectively. The B decay point constraint is applied in the vertex reconstruction. The reduced χ^2 of the vertex fit must be smaller than 15 to reject poorly reconstructed vertices. Figure 3.13 shows the distribution for $(z_{CP}^{\rm rec} - z_{CP}^{\rm gen})$, defined as reconstructed B_{CP} decay vertex minus generated B_{CP} decay vertex. Typical resolution is estimated to be ~ 85 μm, and the vertex reconstruction efficiency is estimated to be 95%.

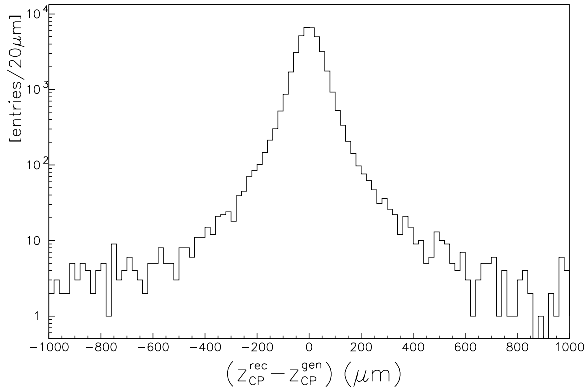


Figure 3.13: Distribution for $(z_{CP}^{\rm rec} - z_{CP}^{\rm gen})$. The distribution is obtained from $B^0 \to J/\psi K_S$ signal Monte Carlo.

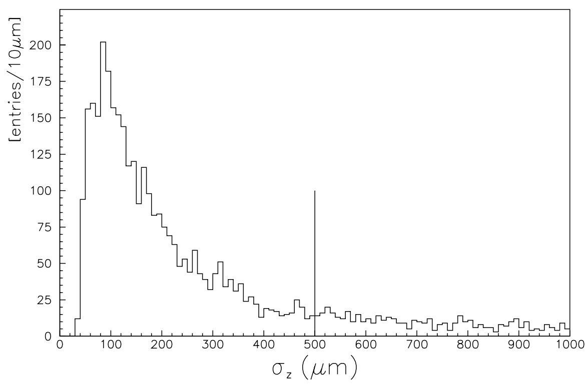


Figure 3.14: Distribution for σ_z obtained from $B^+ \to J/\psi K^+$ decay in real data within the signal region (same definition as $B^0 \to J/\psi K_S$ decay). The usage of charged B events is due to the low statistics of $B^0 \to J/\psi K_S$ decay, and the less significance of the distribution in tail region. The vertical lines indicate the selection criterion.

3.5.3 Vertex Reconstruction of $B_{\rm tag}$

The reconstruction of $B_{\rm tag}$ is much complicated than that of B_{CP} [55]. The major reason is that the $B_{\rm tag}$ side tracks may contain the decay product of mesons with finite lifetime such as charmed mesons or K_S. The $B_{\rm tag}$ vertex is reconstructed with all tracks in the event except for the particles used for B_{CP} reconstruction, although we discard some badly reconstructed tracks or non-primary tracks to improve the $z_{\rm tag}$ resolution. The track elimination is as follows. First, the tracks are required to have at least one hit in the r-ϕ plane and two hits in the r-z plane on the SVD as well as the vertex reconstruction of B_{CP}. It is also required that the track positions must be measured more accurate than 500 μm in z direction (σ_z). Figure 3.14 shows the distribution of σ_z. Then the tracks that are associated to the decay products of K_S whose invariant mass is within

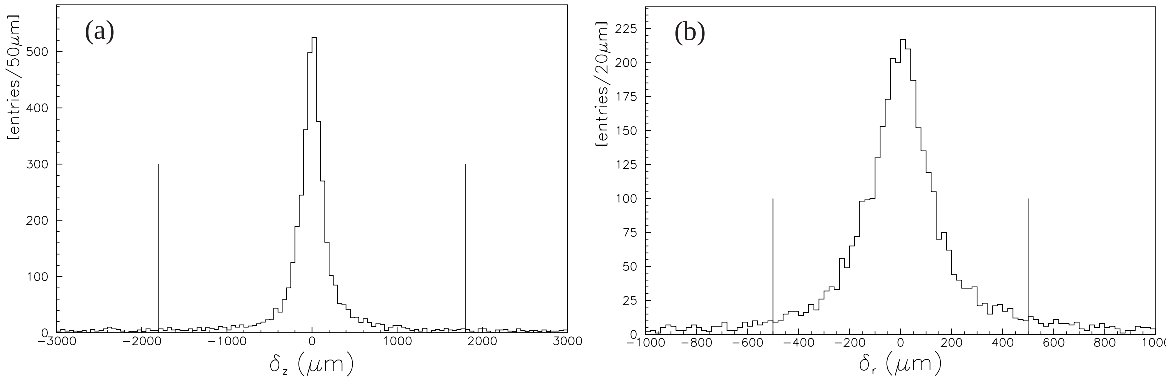


Figure 3.15: Distributions for (a) δ_z and (b) δ_r. The distributions are obtained from $B^+ \to J/\psi K^+$ decay in real data. The vertical lines indicate the thresholds.

$|M_{\pi^+\pi^-} - M_{K_S}| < 15$ MeV/c^2 are removed. Moreover, the suspicious tracks coming from K_S decays are also removed by the following selection criteria:

- The track position measured from J/ψ decay vertex in the r-ϕ plane should satisfy $\delta_r < 500$ μm.
- The track position measured from J/ψ decay vertex in the z direction should satisfy $|\delta_z| < 1800$ μm.

Figures 3.15 show the distributions of δ_r and δ_z. We also eliminate bad tracks during the kinematic vertex fit procedure. The vertex reconstruction is iterated until the reduce χ^2 of the vertex reconstruction becomes smaller than 20. If the reduced χ^2 is worse than 20, we discard one track that has the least contribution to the reconstructed vertex in terms of χ^2 and repeat the vertex fit. In the procedure, we always keep a track that tagged the $B_{\rm tag}$ flavor, because such lepton is confidently originated from $B_{\rm tag}$. The $B_{\rm tag}$ vertex is reconstructed with an application of B decay point constraint as well as z_{CP} reconstruction. When the vertex reconstruction with last remaining track with B decay point constraint still fails (the reduced χ^2 is worse than 20), the event is discarded. All tracks are assumed to have pion mass approximately, which yields only negligible discrepancy between the usage of pion mass and the usage of tagged species masses.

The efficiency loss due to the track selection on σ_z, δ_z, and δ_r is measured to be smaller than 4%. Figure 3.16 shows the comparison of $(z_{\rm tag}^{\rm rec} - z_{\rm tag}^{\rm gen})$, with and without track selection. The distributions are obtained from $B^0 \to J/\psi K_S$ signal Monte Carlo. The solid histogram represents the case with track selection and the broken histogram represents the without selection case. Only the tail part of the distribution effectively lost by the track selection while keeping the main part unchanged. The r.m.s.'s for the cases with and without selection are 153 μm and 162 μm, respectively. The overall reconstruction

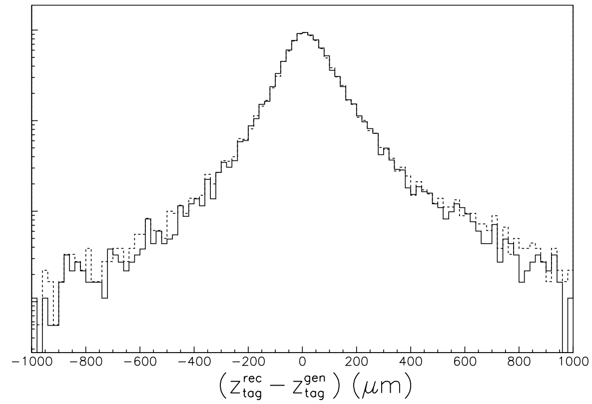


Figure 3.16: Comparison of the $z_{\rm tag}^{\rm rec} - z_{\rm tag}^{\rm gen}$ distributions for the cases with track selection (solid line) and without track selection (broken line). The distributions are obtained from $B^0 \to \psi K_S$ signal Monte Carlo. The r.m.s. of the distributions improved from 162 μm to 153 μm by the track selection.

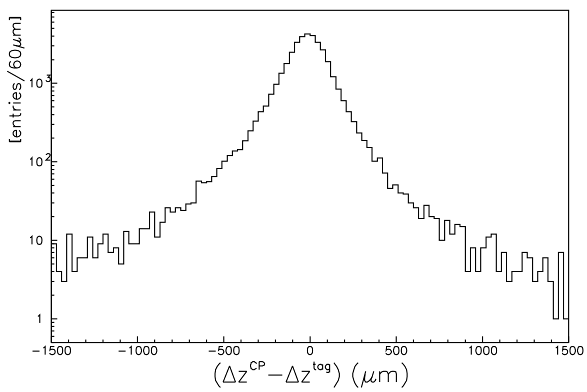


Figure 3.17: Distribution for $(\Delta z^{\rm rec} - \Delta z^{\rm gen})$ obtained from $B^0 \to J/\psi K_S$ signal Monte Carlo.

efficiency of $B_{\rm tag}$ is 93%.

The distribution for $(\Delta z^{\rm rec} - \Delta z^{\rm gen})$ is shown in Figure 3.17. The Δz resolution is studied in Section 4.2.2. Figure 3.18 (a) shows the distribution for $\Delta t \simeq \Delta z/(c\beta\gamma)$ obtained from reconstructed $B^0 \to J/\psi K_S$ candidates in the real data. Figure 3.18 (b) shows the Δt distribution from the data separated according to the $B_{\rm tag}$ flavor. $q = +1$ events are represented by the white points and $q = -1$ events are represented by the black points, respectively. We obtain $\sin 2\phi_1$ value from the asymmetry in this distribution.

In the reconstruction of proper-time difference, we lost 26 events in the vertex reconstruction of B_{CP}. More 12 events are discarded due to the poorness of χ^2/n. The vertices

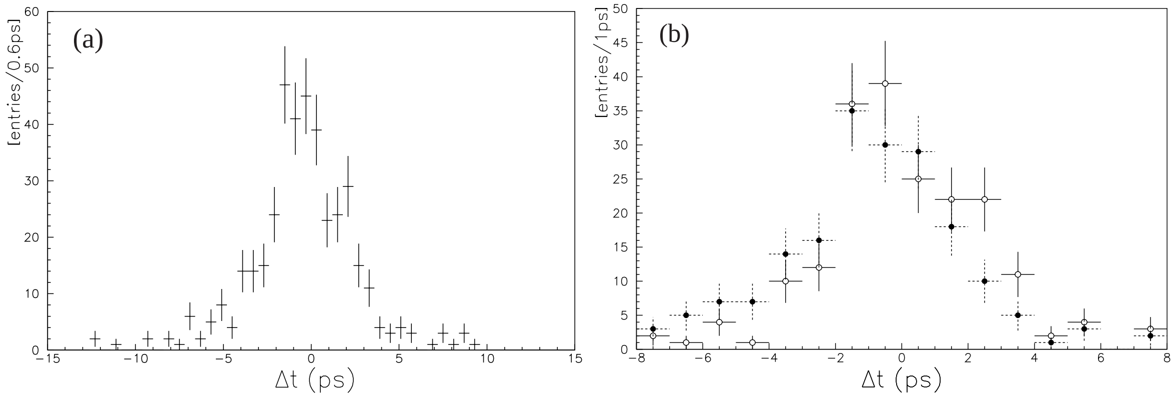


Figure 3.18: (a) Distribution for $\Delta t \simeq \Delta z/(c\beta\gamma)$ and (b) Δt distribution separated with tagged flavor. The distributions are obtained from $B^0 \to J/\psi K_S$ events in real data.

of B_{tag} requires the reconstructed vertex of B_{CP} for the track rejection. After the B_{CP} reconstruction, we perform the vertex reconstruction of B_{tag}. A number of the final B^0 candidates with Δt measurement accompanied with the flavor information is 387.

Chapter 4

Determination of $\sin 2\phi_1$

In this chapter, we discuss the determination of $\sin 2\phi_1$ with reconstructed proper-time differences of 387 $B^0 \to J/\psi K_S$ candidates. The $\sin 2\phi_1$ value is obtained from the fit to the probability density function of Δt on the obtained Δt distribution with unbinned-maximum-likelihood method, which is a common technique for the parameter determination. We describe the technique in Section 4.1. The determination of the probability density function of the Δt distribution is given in Section 4.2. In the section, we discuss a event-by-event signal probability, a Δt resolution, and the treatment of the background components. The result of the $\sin 2\phi_1$ determination is presented in Section 4.3. For the obtained $\sin 2\phi_1$ value, we perform an examination of systematic dependence on the parameters for the $\sin 2\phi_1$ determination in Section 4.4. We also discuss the validity of the determination procedure of $\sin 2\phi_1$ in Section 4.5.

4.1 Maximum-Likelihood Method

The unbinned-maximum-likelihood method [56] is effective to determine a parameter when a number of the events is small. In addition, we can take into account event-by-event effect to the probability density function with the unbinned-maximum-likelihood method. Since we define the probability density function of Δt in event by event, as we describe in Section 4.2, the unbinned-maximum-likelihood is also effective from this point. This method is based on the idea that the parameters manifest themselves through "most probable" path among several paths labeled by parameters. Let the normalized probability density function be $\mathcal{P}(\Delta t;\ \sin 2\phi_1)$. We construct the likelihood function, $\mathcal{L}(\sin 2\phi_1)$, with $\mathcal{P}(\Delta t;\ \sin 2\phi_1)$ as

$$\mathcal{L}(\sin 2\phi_1) \equiv \prod_{i=1}^{N} \mathcal{P}(\Delta t_i;\ \sin 2\phi_1), \qquad \ln\Big[\mathcal{L}(\sin 2\phi_1)\Big] = \sum_{i=1}^{N} \ln\Big[\mathcal{P}(\Delta t_i;\ \sin 2\phi_1)\Big], \quad (4.1)$$

where Δt_i is an i-th measurement of Δt, and N is a number of event, 387. The likelihood function is the joint probability density function of getting a particular experimental result, Δt_1, Δt_2, ..., Δt_N, assuming $\mathcal{P}(\Delta t;\ \sin 2\phi_1)$ is the true distribution function. The relative probabilities of $\sin 2\phi_1$ can be displayed as a plot of $\mathcal{L}(\sin 2\phi_1)$ as a function of $\sin 2\phi_1$. The most probable $\sin 2\phi_1$, $\sin 2\phi_1^*$, is considered to make the likelihood maximum. The spread of $\sin 2\phi_1$ around $\sin 2\phi_1^*$ is an accuracy of the determination of $\sin 2\phi_1 = \sin 2\phi_1{}^*$. This method is called maximum-likelihood method.

In general, the likelihood function will be close to the Gaussian. Let $\sin 2\phi_1$ be $\sin 2\phi_1 \equiv \delta(\sin 2\phi_1) + \sin 2\phi_1^*$. Under the approximation of the Gaussian shape, the likelihood function can be expressed as

$$\ln\Big[\mathcal{L}\big(\delta(\sin 2\phi_1) + \sin 2\phi_1^*\big)\Big] = -\frac{\delta(\sin 2\phi_1)^2}{2\sigma^2} + \ln\Big[\mathcal{L}(\sin 2\phi_1^*)\Big] \tag{4.2}$$

The error of $\sin 2\phi_1$ is estimated to be $\delta(\sin 2\phi_1)$ that satisfies

$$\ln\Big[\mathcal{L}\big(\delta(\sin 2\phi_1) + \sin 2\phi_1^*\big)\Big] = \ln\Big[\mathcal{L}(\sin 2\phi_1^*)\Big] - \frac{1}{2}. \tag{4.3}$$

In this study, we use MINUIT routine [57] to search for the most likely $\sin 2\phi_1$ that makes the likelihood maximum. The original purpose of the MINUIT is to find minimum χ^2 with iteration of free parameters, $\Delta_1, \Delta_2, ...$:

$$\chi^2 = \frac{\Delta_1^2}{\sigma_1^2} + \frac{\Delta_2^2}{\sigma_2^2} + ... \tag{4.4}$$

When minimum χ^2 is $(\chi^2)_{\rm mim}$, the estimated errors of the free parameters are defined as parameter sets that gives $\chi^2 = (\chi^2)_{\rm mim} + 1$. Thus, the MINUIT searches for the parameter sets that gives $\chi^2 = (\chi^2)_{\rm mim} + 1$ to return the estimated errors. To utilize the MINUIT routine for the maximum-likelihood-method, we "minimize" $-2\ln[\mathcal{L}(\vec{\beta})]$.

4.2 Probability Density Function

In this section, we construct the probability density function of the Δt distribution. The probability density function for the proper-time difference is expressed as

$$\mathcal{P}(\Delta t;\ \sin 2\phi_1) = f_{\rm sig}\cdot\mathcal{P}_{\rm sig}(\Delta t;\ \sin 2\phi_1) + (1 - f_{\rm sig})\cdot\mathcal{P}_{\rm bkg}(\Delta t), \tag{4.5}$$

where $f_{\rm sig}$ is a signal fraction of each event. $\mathcal{P}_{\rm sig}$ is a signal probability density function convoluted with resolution function. Using equation (3.7), the $\mathcal{P}_{\rm sig}$ is represented as

$$\mathcal{P}_{\rm sig}(\Delta t;\ \sin 2\phi_1) = \int_{-\infty}^{\infty} {\rm d}(\Delta t')\ \mathcal{R}(\Delta t - \Delta t')\cdot f_{\ell^\pm}(\Delta t';\ \sin 2\phi_1), \tag{4.6}$$

where $\mathcal{R}(x)$ is a resolution function. We use event-by-event wrong tag fraction for $f_{\ell^\pm}(\Delta t; \sin 2\phi_1)$. The background probability density function is represented by $\mathcal{P}_{\rm bkg}$ in equation (4.5). $\mathcal{P}(\Delta t; \sin 2\phi_1)$ consists of many parameters related to the real data:

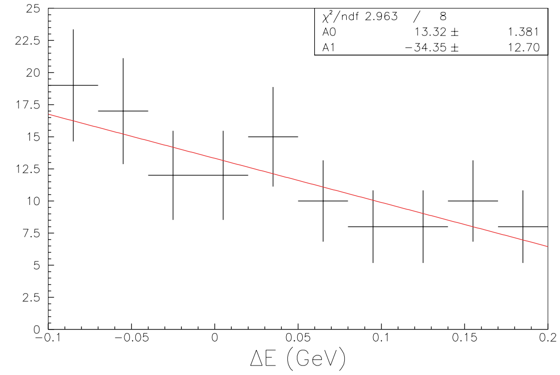


Figure 4.1: Event distribution in ΔE in background dominated region obtained from reconstructed $B^0 \to J/\psi K_S$ candidates in real data. Beam constraint mass selection of $5.200 < M_{\rm bc} < 5.265$ GeV/c^2 is applied. A fit to first order polynomial is also superimposed.

$\sin 2\phi_1$, τ_{B^0}, ΔM, $f_{\rm sig}$, $\mathcal{R}$, and $\mathcal{P}_{\rm bkg}$, but due to the small statistics of $B^0 \to J/\psi K_S$ we can determine only $\sin 2\phi_1$ in the maximum-likelihood-fit. For the fit, τ_{B^0} and ΔM are substituted by world averages. The constructions of function forms and the determinations of their parameters for $f_{\rm sig}$ using real data are described in Subsection 4.2.1. We construct the Δz resolution function form using the Monte Carlo simulation, which is described in Subsection 4.2.2. The resolution function form is also utilized to parameterize the Δt distribution for background events. The determination of background distribution parameters is described in Subsection 4.2.3. After we determine the $f_{\rm sig}$ and $\mathcal{P}_{\rm bkg}$ distribution, we determine the resolution function parameters using real data in Subsection 4.2.4.

4.2.1 Signal Probability

A signal probability of the reconstructed event is not constant over all events. For example, we should apply rather higher signal probability for a event with reconstructed B meson mass that has less deviation from the average value. We represent an event-by-event signal probability, $f_{\rm sig}$, with ΔE and $M_{\rm bc}$, which strongly correlates to it and which are independent parameters of each other. The signal distribution in ΔE-$M_{\rm bc}$ plane is represented by two-dimensional single Gaussian. As for the background, beam constraint mass of the the background events shows ARGUS background function form [58]. We adopt this function for $M_{\rm bc}$ distribution. The distribution for ΔE in the background dominated region ($-0.1 < \Delta E < 0.2$ GeV $\cap$ $5.200 < M_{\rm bc} < 5.265$ GeV/c^2) shows a linear dependency as shown in Figure 4.1. We represent background distribution in ΔE by a first order polynomial. In summary, following functions are used to determine the

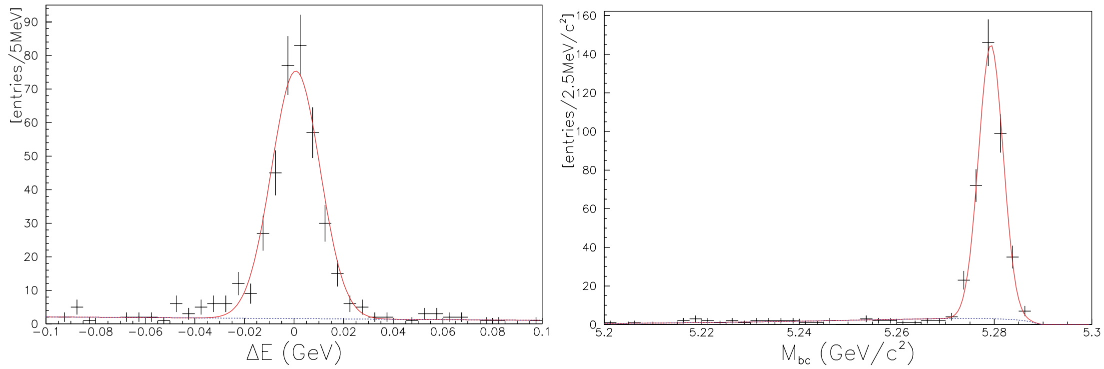


Figure 4.2: Results from fitting ΔE (left) and $M_{\rm bc}$ (right) distributions. The fits are performed on the events that passed all selection criteria. The selection $5.2694 < M_{\rm bc} < 5.2864$ GeV/c^2 for ΔE distribution, and that of $|\Delta E| < 40$ MeV for $M_{\rm bc}$ distribution are applied.

event-by-event signal probability:

$$f_{\rm sig} = \frac{A_{\rm sig}(\Delta E, M_{\rm bc})}{A_{\rm sig}(\Delta E, M_{\rm bc}) + A_{\rm bkg}(\Delta E, M_{\rm bc})}, \tag{4.7}$$

$$A_{\rm sig} = a \cdot G(\Delta E;\ \mu_{\Delta E}, \sigma_{\Delta E}) \cdot G(M_{\rm bc};\ \mu_{M_{\rm bc}}, \sigma_{M_{\rm bc}}), \tag{4.8}$$

$$A_{\rm bkg} = b \cdot M_{\rm bc} \cdot \sqrt{u(M_{\rm bc})} \cdot {\rm e}^{c \cdot u(M_{\rm bc})} \times (1 + d \cdot \Delta E), \tag{4.9}$$

where

$$u(M_{\rm bc}) \equiv 1 - \left(\frac{M_{\rm bc}}{E_{\rm beam}}\right)^2. \tag{4.10}$$

The signal fraction parameters are determined by a fit with the unbinned-maximum-likelihood method in the limited ΔE-$M_{\rm bc}$ region ($-0.1 < \Delta E < 0.2$ GeV $\cap$ $5.200 < M_{\rm bc} < 5.265$ GeV/c^2) or ($-0.05 < \Delta E < 0.05$ GeV $\cap$ $5.265 \leq M_{\rm bc} < 5.290$ GeV/c^2). We limit the ΔE-$M_{\rm bc}$ range to avoid the contamination from the $B^0 \to J/\psi K^{*0}$ events. Figures 4.2 show the distributions for ΔE and $M_{\rm bc}$ with superimposed solid lines representing the distributions of a sum of signal and background, and with the broken lines representing the background distributions, where the lines are determined by the fit. Table 4.1 lists the parameters for the $A_{\rm sig}$ and $A_{\rm bkg}$ obtained by the fit. The background fraction in the signal region is estimated to be 0.038 ± 0.010.

4.2.2 Resolution Function

In this subsection, we discuss the resolution function for the signal events [59]. The resolution function is considered to consist of three components: (i) detector resolution mainly due to tracking uncertainty for reconstructed B_{CP} and $B_{\rm tag}$ vertices, (ii) smearing

Table 4.1: Parameters for $A_{\rm sig}$ and $A_{\rm bkg}$ obtained by the fit.

Fit parameters	Fitted values
$\mu_{M_{\rm bc}}$ (GeV/c^2)	5279.3 ± 0.1
$\sigma_{M_{\rm bc}}$ (GeV/c^2)	2.5 ± 0.1
$\mu_{\Delta E}$ (GeV)	0.71 ± 0.57
$\sigma_{\Delta E}$ (GeV)	9.78 ± 0.54
c: ARGUS param.	-91 ± 12
d: ΔE slope (GeV^{-1})	$-3.3\ ^{+0.6}_{-0.5}$

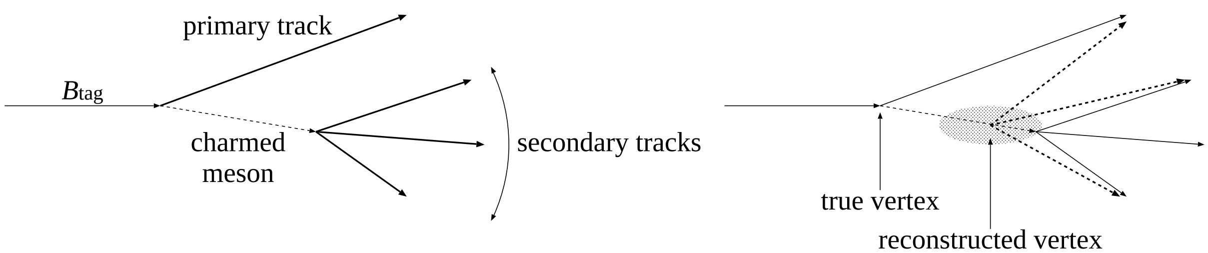


Figure 4.3: Schematic drawing of the smearing effect due to non-primary tracks.

on $B_{\rm tag}$ vertices due to non-primary tracks as shown in Figure 4.3, and (iii) smearing due to the kinematic approximation shown in equation (3.18). The convolution of three components yields the resolution function, $\mathcal{R}$. We use $B^0 \to J/\psi K_S$ signal Monte Carlo to determine the function form of $\mathcal{R}$.

In the following paragraphs, first we study the Δz resolution without the effect of the charmed meson lifetime. The effects of the charm lifetimes are included after understanding the resolution function due to tracking. Finally, the effect of the approximation in the Δt calculation is taken into account.

Detector Resolution

Figure 4.4 shows the distribution for $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ obtained from the Monte Carlo, where $\Delta z_{\rm rec}$ and $\Delta z_{\rm gen}$ denotes the reconstructed and generated Δz, respectively. The Monte Carlo events are generated with setting $\tau_D = 0$ artificially to suppress the charmed meson lifetime, which makes positive shifts on $B_{\rm tag}$ vertex, and thus negative shifts on $\Delta z_{\rm rec}$, as we described. A fit to a sum of double Gaussians, main Gaussian and tail Gaussian, is superimposed in the figure. The fit is not a good representation of the distribution since the vertex resolution depends on track momenta, angles and other factors, and therefore the distribution consists of many Gaussians. We can calculate the track error event by event using a Kalman filtering technique that takes into account

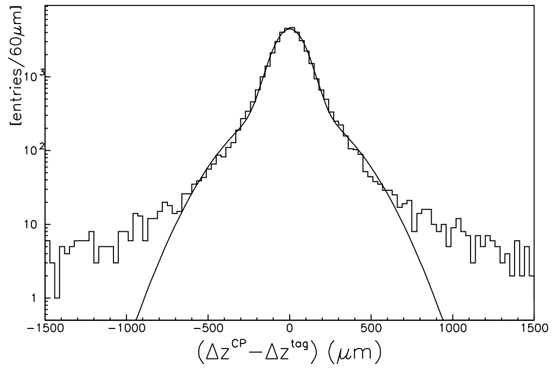


Figure 4.4: Distribution for $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ obtained from signal Monte Carlo. A fit to a sum of double Gaussians is also shown.

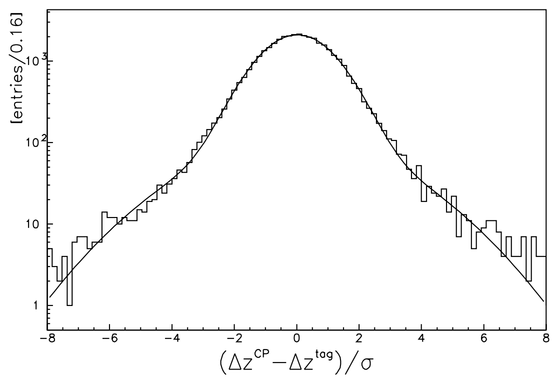


Figure 4.5: Distribution for $(\Delta z_{\rm rec} - \Delta z_{\rm gen})/\sigma_{\Delta z}$ obtained from signal Monte Carlo. A fit to a sum of two Gaussians is superimposed.

the resolution of each hit, multiple scattering and energy loss. Figure 4.5 shows the distribution for $(\Delta z_{\rm rec} - \Delta z_{\rm gen})/\sigma_{\Delta z}$ from the Monte Carlo simulation, where $\sigma_{\Delta z}$ is the calculated Δz error from the vertex fit for each event. The distribution is well represented by a fit to a sum of two Gaussians, shown in the figure by the overlaid solid curve. The fit is performed with binned maximum likelihood method. The standard deviation of the main Gaussian would be 1.0 if the error estimation were perfect, whereas the fit result is 1.08 ± 0.01. The difference is considered to be due to underestimation of the tracking error. The remaining events, represented by the broader tail Gaussian, are considered to be badly measured events, due to miss-association of the SVD hits, incorrect SVD-hit clustering, hard scattering of tracks, poorly reconstructed vertices by single track and so on. Therefore, we can construct a model that the detector resolution is represented by event-by-event vertex reconstruction error scaled with a certain correction referred as “scaling factor”. Figure 4.6 shows the $\sigma_{\Delta z}$ distributions obtained from Monte Carlo simulation (solid line) and from real data (dotted line). As the figure implies,

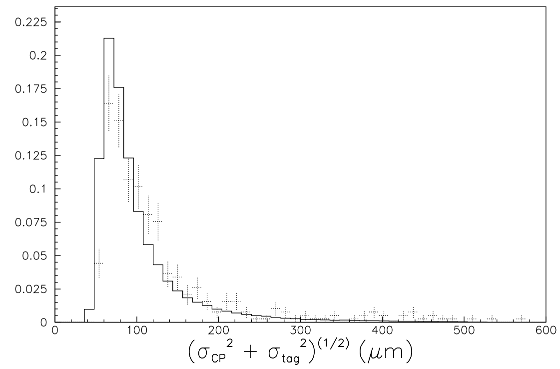


Figure 4.6: Normalized distributions for $\sigma_{\Delta z} \equiv \sqrt{\Delta z_{CP}^2 + \Delta z_{\rm tag}^2}$ obtained from signal Monte Carlo (solid line) and from real data (dotted line).

the simulated detector resolution may differ from the true detector resolution, which is quite natural. However, since we use $\sigma_{\Delta z}$ from real data in the determination of $\sin 2\phi_1$, such discrepancy can be absorbed. Based on the success of this fit, we represent the Δz resolution due to detector performance, $R_{\rm det}$, as a function of the vertex reconstruction error:

$$R_{\rm det}(\Delta z) = (1 - f_{\rm tail})\cdot G(\Delta z;\ \mu^{\Delta z}_{\rm main}, \sigma^{\Delta z}_{\rm main}) + f_{\rm tail}\cdot G(\Delta z;\ \mu^{\Delta z}_{\rm tail}, \sigma^{\Delta z}_{\rm tail}), \tag{4.11}$$

$$\sigma^{\Delta z}_{\rm main} = s_{\rm main}\cdot\sigma_{\Delta z}, \tag{4.12}$$

$$\sigma^{\Delta z}_{\rm tail} = s_{\rm tail}\cdot\sigma_{\Delta z}, \tag{4.13}$$

$$\sigma_{\Delta z} = \sqrt{\sigma_{CP}^2 + \sigma_{\rm tag}^2}, \tag{4.14}$$

where $s_{\rm main}$ and $s_{\rm tail}$ are global scaling factors for the estimated error $\sigma_{\Delta z}$ for the main and tail Gaussian distribution, σ_{CP} and $\sigma_{\rm tag}$ are z vertex resolution analytically calculated from error matrices of the track reconstruction for B_{CP} and $B_{\rm tag}$, respectively, and $f_{\rm tail}$ is the fraction of the tail-part Gaussian. Here the Gaussian distribution is abbreviated by

$$G(x;\ \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{4.15}$$

Figure 4.7 shows the distribution for $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$. A superimposed curve represents the fit to $R_{\rm det}$ distribution, where the fit is performed with unbinned-maximum-likelihood method. The fit gives $s_{\rm main} = 1.140 \pm 0.004$, $s_{\rm tail} = 3.61 \pm 0.06$, and $f_{\rm tail} = 0.048 \pm 0.002$. $\mu^{\Delta z}_{\rm main}$ and $\mu^{\Delta z}_{\rm tail}$ are fixed to zero because the detectors are not expected to generate shifts in vertex reconstruction without canceling.

Figure 4.8 (a) shows the distribution for χ^2/n of the vertex reconstruction of the $B_{\rm tag}$ for the signal Monte Carlo, where χ^2/n represents the reduced χ^2. A small fraction of the events has a large χ^2/n. These outliers are considered to be due to badly measured

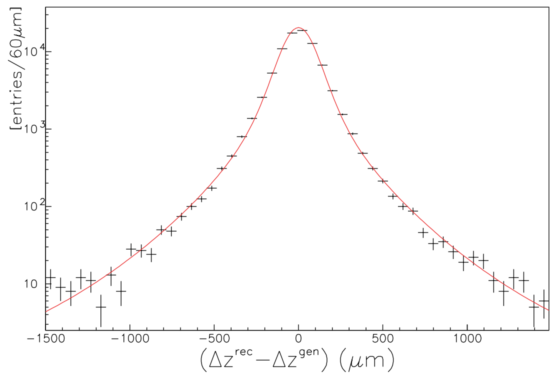


Figure 4.7: Distribution for $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ obtained from the signal Monte Carlo. A fit to $R_{\rm det}$ function is superimposed. $\mu^{\Delta z}_{\rm main}$ and $\mu^{\Delta z}_{\rm tail}$ are fixed to zero. The standard deviation of main Gaussian is fitted to 1.140 ± 0.004, and the fitted standard deviation for the tail Gaussian and its fraction is 3.61 ± 0.06 and 0.048 ± 0.002, respectively.

Table 4.2: Fit parameters form fits to $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distribution with the resolution function.

χ^2/n region	$s_{\rm main}$	$s_{\rm tail}$	$f_{\rm tail}$
$\chi^2/n < 3.0$	1.116 ± 0.004	3.49 ± 0.07	0.034 ± 0.002
$3.0 \le \chi^2/n < 4.5$	1.45 ± 0.03	3.29 ± 0.22	0.062 ± 0.012
$4.5 \le \chi^2/n$	2.10 ± 0.06	3.11 ± 0.16	0.129 ± 0.019
All	1.140 ± 0.004	3.61 ± 0.06	0.048 ± 0.002

tracks, contamination of K_S daughters and so on. We evaluate the effect of those outliers on the resolution by dividing the signal events into three samples according to χ^2/n. The first region is $\chi^2/n < 3.0$, the second is $3.0 \le \chi^2/n < 4.5$, and the third is $4.5 \le \chi^2/n$. We fit the $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distributions with $R_{\rm det}$ functions for these three samples as shown in Figures 4.8 (b)−(d). Table 4.2 lists the fit results for each region. It is apparent that the resolution degrades with larger χ^2/n. To take into account this effect, we redefine $\sigma_{\Delta z}$ for the event with $\chi^2/n \ge 3.0$ as

$$\tilde{\sigma}_{\Delta z} = \sqrt{\tilde{\sigma}^2_{CP} + \tilde{\sigma}^2_{\rm tag}} \tag{4.16}$$

$$\tilde{\sigma}_{CP} = \sqrt{1 + \alpha_{CP}\left[(\chi^2/n)_{CP} - 3\right]} \cdot \sigma_{CP}, \qquad (\chi^2/n)_{CP} \ge 3, \tag{4.17}$$

$$\tilde{\sigma}_{\rm tag} = \sqrt{1 + \alpha_{\rm tag}\left[(\chi^2/n)_{\rm tag} - 3\right]} \cdot \sigma_{\rm tag}, \qquad (\chi^2/n)_{\rm tag} \ge 3, \tag{4.18}$$

where $(\chi^2/n)_{CP}$ and $(\chi^2/n)_{\rm tag}$ are the reduced χ^2/n of B_{CP} and $B_{\rm tag}$ vertex reconstructions, respectively. The redefinition is equivalent to the assignments of different $s_{\rm main}$ for B_{CP} and $B_{\rm tag}$, and the introduction of the χ^2/n dependence to $s_{\rm main}$. Figure 4.9 (a) shows the $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distribution with overlaid $R_{\rm det}$ function using redefined $\sigma_{\Delta z}$

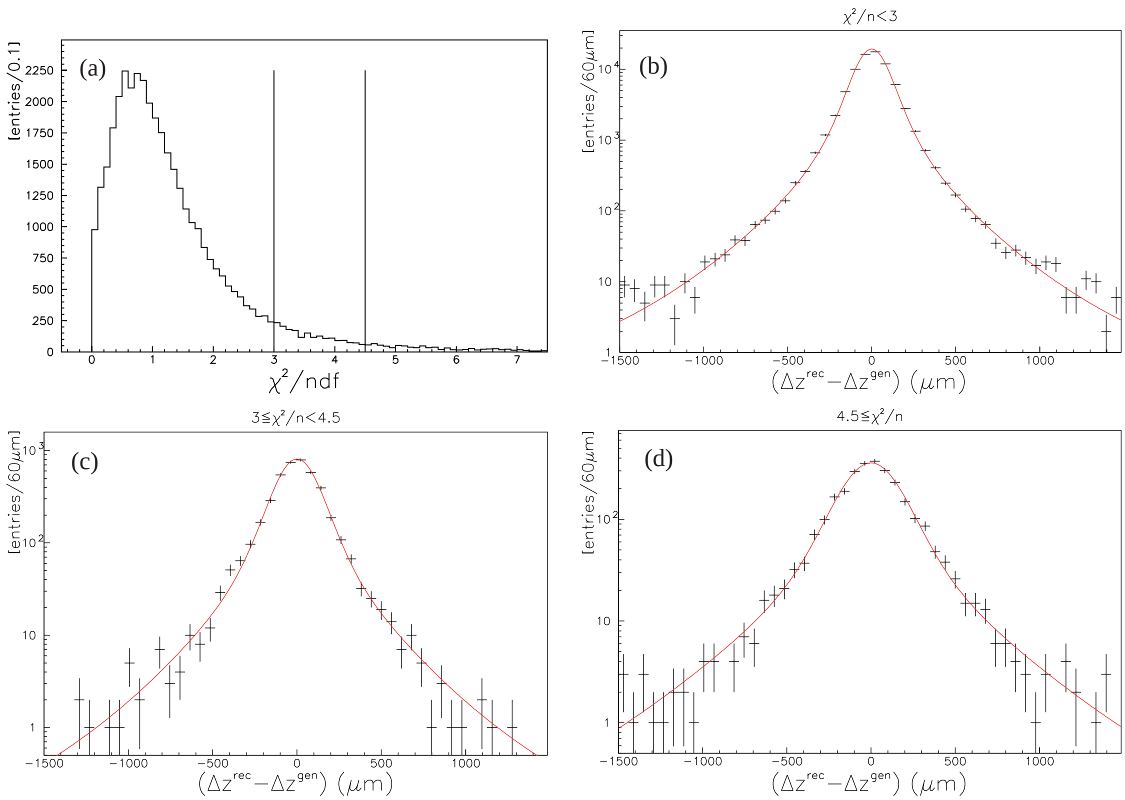


Figure 4.8: (a) Distribution of χ^2/n of the vertex constraint fit for $B_{\rm tag}$, and its of the $R_{\rm det}$ function to $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distribution with (b) $\chi^2/n < 3.0$, (c) $3.0 \leq \chi^2/n < 4.5$, and (d) $4.5 \leq \chi^2/n$, for Monte Carlo signal events.

Table 4.3: Parameters for $R_{\rm det}$ obtained from the fit to $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distribution without charmed meson lifetimes.

Fit parameters	Fitted values
$s_{\rm main}$	1.035 ± 0.003
$s_{\rm tail}$	3.79 ± 0.05
$f_{\rm tail}$	0.033 ± 0.001
α_{CP}	1.02 ± 0.03
$\alpha_{\rm tag}$	1.64 ± 0.05

for all χ^2/n. The obtained parameters from the fits are listed in Table 4.3. Figures 4.9 (b) through (d) show the $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distributions for the different χ^2/n region with $R_{\rm det}$ function superimposed. The fit is performed once for all χ^2/n regions. The fit is well representing each distribution.

Smearing Due to Non-Primary Tracks

We now take into account the effect of charm vertex in the vertex reconstruction of the $B_{\rm tag}$ decay. Since this effect is due to a physics process, we can rely on the Monte Carlo simulation and uses the parameters for $\sin 2\phi_1$ determination obtained by the simulation study. Redefinitions of $\sigma^{\Delta z}_{\rm main}$ and $\sigma^{\Delta z}_{\rm tail}$ are given as

$$\sigma^{\Delta z}_{\rm main} = \sqrt{s_{\rm main}\cdot\sigma^2_{CP} + [s^2_{\rm main} + (s^{\rm NP}_{\rm main})^2]\cdot\sigma^2_{\rm tag}}, \tag{4.19}$$

$$\sigma^{\Delta z}_{\rm tail} = \sqrt{s_{\rm tail}\cdot\sigma^2_{CP} + [s^2_{\rm tail} + (s^{\rm NP}_{\rm tail})^2]\cdot\sigma^2_{\rm tag}}, \tag{4.20}$$

where $s^{\rm NP}_{\rm main}$ and $s^{\rm NP}_{\rm tail}$ corresponds to the smearing from the displacement of charm vertex. Figure 4.10 shows the $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distribution in the Monte Carlo signal events with proper charm lifetimes. A fit to $R_{\rm det}$ function using the new scaling factors is superimposed. $f_{\rm tail}$ is substituted by the value obtained with zero charm lifetime. The mean shifts of the Gaussians in $R_{\rm det}$, $\mu^{\Delta z}_{\rm main}$ and $\mu^{\Delta z}_{\rm tail}$, are allowed to vary in the fit. Table 4.4 lists fit parameters obtained from the fit. Figure 4.11 (a) shows the distribution for $\sigma_{\rm tag}$ obtained from the signal Monte Carlo. Small fraction of events have a large $\sigma_{\rm tag}$. We evaluate the effect of these large $\sigma_{\rm tag}$ on the resolution function by dividing the Monte Carlo signal events into three samples according to $\sigma_{\rm tag}$: $\sigma_{\rm tag} < 50\ \mu$m, $50 \leq \sigma_{\rm tag} < 80\ \mu$m, and $80\ \mu$m $\leq \sigma_{\rm tag}$. We fit the $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distributions with the $R_{\rm det}$ function for those three samples as shown in Figures 4.11 (b)$-$(d). Fit results that are listed in Table 4.5 are also superimposed. Fit parameters are consistent with each other with statistical error except for $\mu^{\Delta z}_{\rm main}$ and $\mu^{\Delta z}_{\rm tail}$. We observe slight dependence of the mean shifts on $\sigma_{\rm tag}$

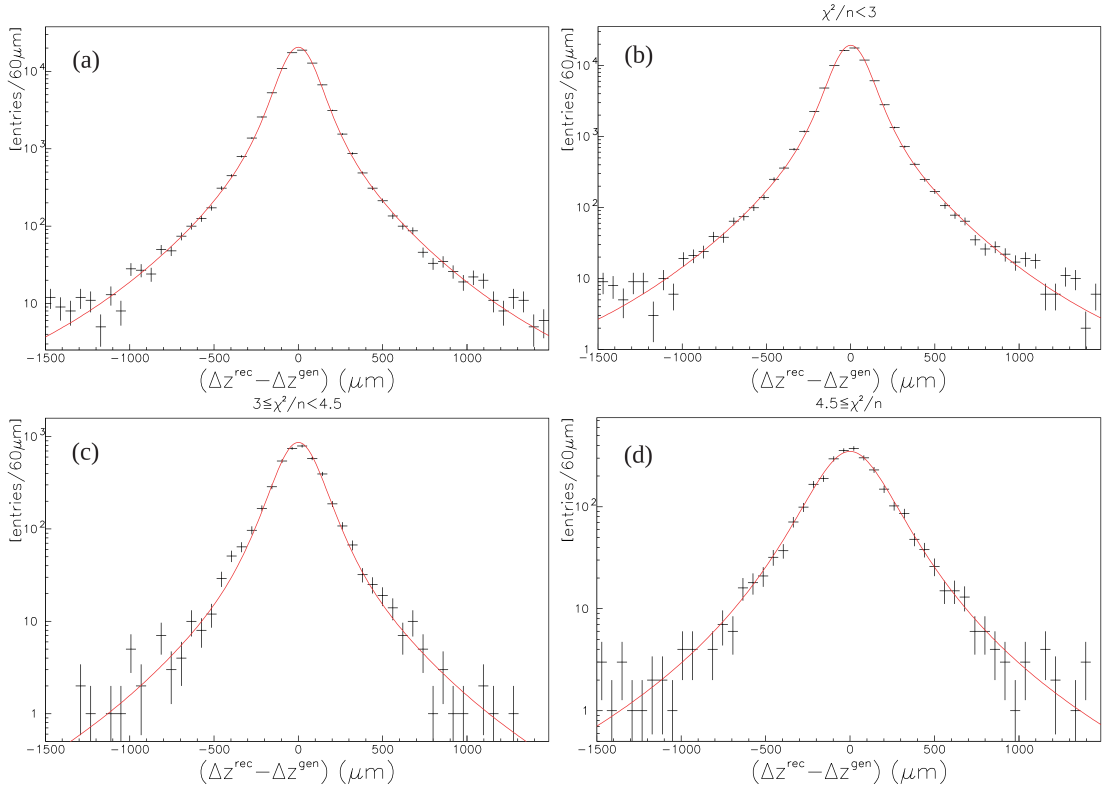


Figure 4.9: (a) $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distribution with a fit of the $R_{\rm det}$ function using the redefined $\sigma_{\Delta z}$. (b)$-$(d) $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distributions for different χ^2/n regions with $R_{\rm det}$ function superimposed. The fit is performed once globally for all χ^2/n region.

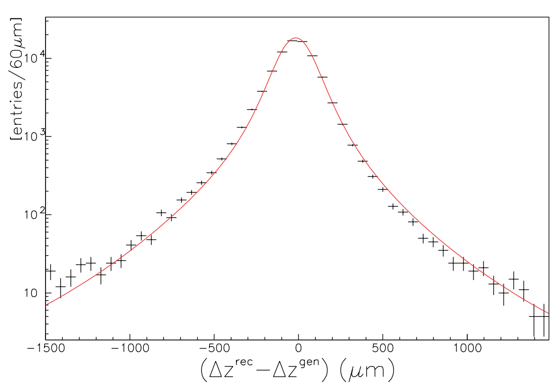


Figure 4.10: $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distribution in the Monte Carlo signal events with proper charm lifetimes. A fit $R_{\rm det}$ convoluted with $R_{\rm np}$ is also superimposed.

Table 4.4: Parameters obtained from a fit to $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distribution with $R_{\rm det}$ function with proper charm lifetimes.

Fit parameters	Fitted values
$s^{\rm NP}_{\rm main}$	0.61 ± 0.01
$s^{\rm NP}_{\rm tail}$	2.40 ± 0.15
$\mu^{\Delta z}_{\rm main}$ (μm)	-16.6 ± 0.4
$\mu^{\Delta z}_{\rm tail}$ (μm)	-94.0 ± 7.3

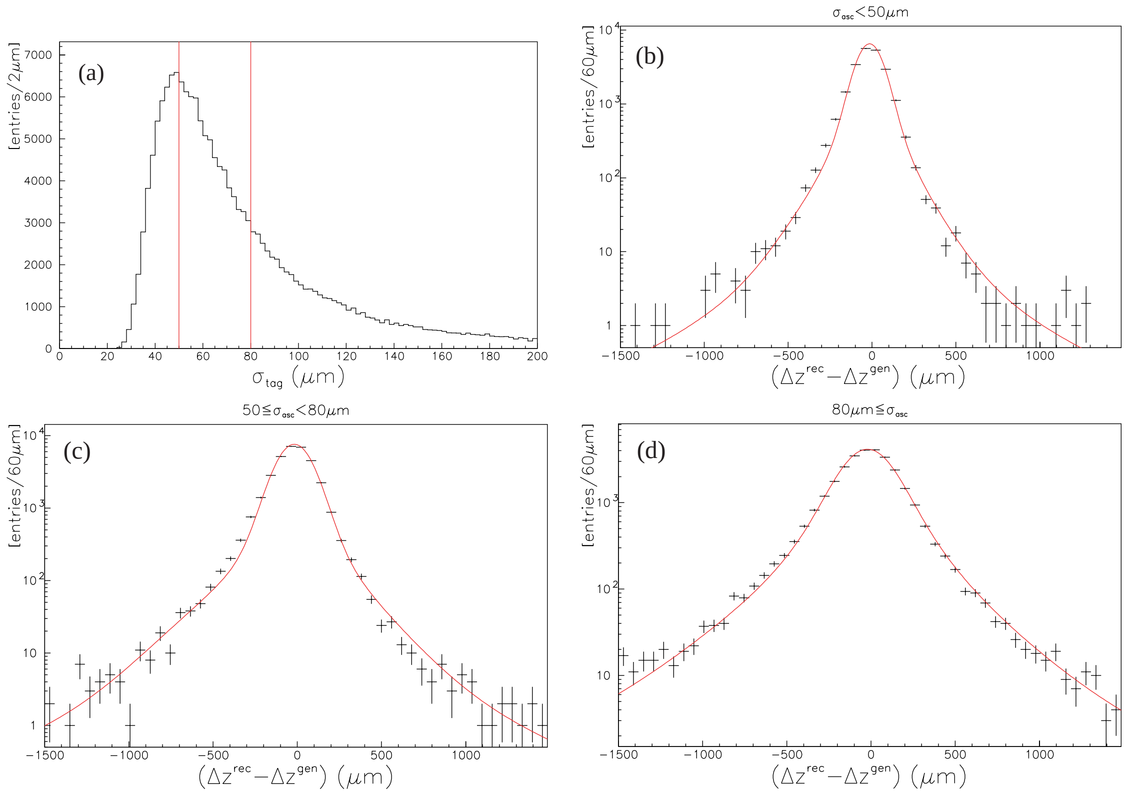


Figure 4.11: (a) Distribution for $\sigma_{\rm tag}$ obtained from signal Monte Carlo events. Small fraction of events has large $\sigma_{\rm tag}$. Fits with convoluted $R_{\rm det}$ function by $R_{\rm np}$ to $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ are shown for (b) lower, (c) higher, and (d) tail regions of $\sigma_{\rm tag}$.

Table 4.5: Parameters obtained from a fit to $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distribution for each $\sigma_{\rm tag}$ region of lower, higher and tail part.

Fit parameters	$\sigma_{\rm tag} < 50\ \mu$m	$50 \leq \sigma_{\rm tag} < 80\ \mu$m	$80\ \mu$m $\leq \sigma_{\rm tag}$
$s^{\rm NP}_{\rm main}$	0.53 ± 0.03	0.68 ± 0.01	0.57 ± 0.01
$s^{\rm NP}_{\rm tail}$	0.00 ± 7.85	2.99 ± 0.23	2.24 ± 0.15
$\mu^{\Delta z}_{\rm main}$ (μm)	-14.5 ± 0.5	-17.6 ± 0.6	-20.7 ± 1.0
$\mu^{\Delta z}_{\rm tail}$ (μm)	-44.6 ± 10.2	-111.88 ± 11.8	-171.9 ± 18.0

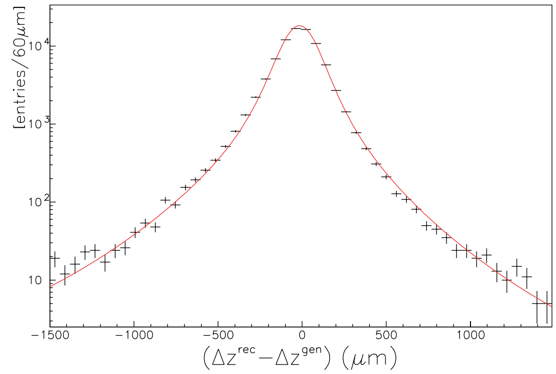


Figure 4.12: Fit to $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ with the convolution of $R_{\rm det}$ and $R_{\rm np}$. The dependence on $\sigma_{\rm tag}$ is introduced for the parameterization of $R_{\rm np}$.

Table 4.6: Parameters obtained from a fit to $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distribution for each $\sigma_{\rm tag}$ region of lower, higher and tail part.

Fit parameters	Fitted Values
$s^{\rm NP}_{\rm main}$	0.59 ± 0.01
$s^{\rm NP}_{\rm tail}$	2.16 ± 0.10
$\mu^0_{\rm main}$ (μm)	-10.9 ± 0.6
$\mu^1_{\rm main}$	-0.097 ± 0.009
$\mu^0_{\rm tail}$ (μm)	-6 ± 12
$\mu^1_{\rm tail}$	-1.42 ± 0.17

and introduce $\sigma_{\rm tag}$ dependent $\mu^{\Delta z}_{\rm main}$ and $\mu^{\Delta z}_{\rm tail}$ as

$$\mu^{\Delta z}_{\rm main} = \mu^0_{\rm main} + \mu^1_{\rm main} \cdot \sigma_{\rm tag}, \qquad (4.21)$$

$$\mu^{\Delta z}_{\rm tail} = \mu^0_{\rm tail} + \mu^1_{\rm tail} \cdot \sigma_{\rm tag}. \qquad (4.22)$$

A fit to the $(\Delta z_{\rm rec} - \Delta z_{\rm gen})$ distributions is made to obtain these parameters. Figure 4.12 shows the fit. Table 4.6 lists the parameters obtained by the fit. During the fit parameters listed in Table 4.3 are fixed to the value in the table.

Smearing Due to Kinematic Approximation

Finally, we combine the error in Δz measurement, $\sigma_{\Delta z}$, and the error due to kinematic approximation in the proper-time calculation, $\sigma_{\rm k}$. As well as the charm meson effect on $B_{\rm tag}$ decay vertex, this smearing is also due to the physics process of two-body decay of $\Upsilon(4S)$. We also utilize the result from the Monte Carlo simulation. We represent the

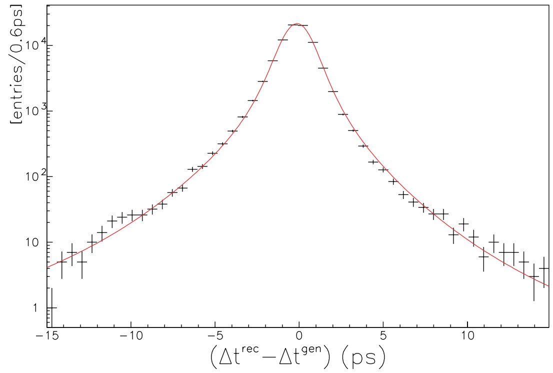


Figure 4.13: Distribution for $(\Delta t_{\rm rec} - \Delta t_{\rm gen})$. The fit to $R_{\rm k}$ function convoluted with the function superimposed in Figure 4.12 is also shown.

proper-time resolution using following functions:

$$\mathcal{R}(x) = (1 - f_{\rm tail})\cdot G(x;\ \mu^{\Delta t}_{\rm main}, \sigma^{\Delta t}_{\rm main}) + f_{\rm tail}\cdot G(x;\ \mu^{\Delta t}_{\rm tail}, \sigma^{\Delta t}_{\rm tail}), \tag{4.23}$$

$$\sigma^{\Delta t}_{\rm main} = \sqrt{\left[\frac{\sigma^{\Delta z}_{\rm main}}{c(\beta\gamma)_\Upsilon}\right]^2 + (\sigma^{\rm main}_{\rm k})^2}, \tag{4.24}$$

$$\sigma^{\Delta t}_{\rm tail} = \sqrt{\left[\frac{\sigma^{\Delta z}_{\rm tail}}{c(\beta\gamma)_\Upsilon}\right]^2 + (\sigma^{\rm tail}_{\rm k})^2}, \tag{4.25}$$

$$\mu^{\Delta t}_{\rm main} = \frac{\mu^{\Delta z}_{\rm main}}{c(\beta\gamma)_\Upsilon}, \tag{4.26}$$

$$\mu^{\Delta t}_{\rm tail} = \frac{\mu^{\Delta z}_{\rm tail}}{c(\beta\gamma)_\Upsilon}. \tag{4.27}$$

We obtain $\sigma^{\rm main}_{\rm k} = 0.287\pm0.004$ ps and $\sigma^{\rm tail}_{\rm k} = 0.32\pm0.19$ ps by a fit to the $(\Delta t_{\rm rec} - \Delta t_{\rm gen})$ distribution with this resolution function as shown in Figure 4.13. All other parameters are fixed in the fit as they are determined independently by the Δz distribution studies.

Summary

Three components of the resolution function are considered: the detector resolution, the smearing due to non-primary tracks, and the smearing due to the kinematic approximation. Second and third components are almost totally constructed by the Monte Carlo simulation, since they are originated by physics processes. Smearing effect on $B_{\rm tag}$ vertex due to non-primary tracks is represented by a single Gaussian, while shifting the mean of the detector resolution apart from zero. The mean shifts depend on the reconstruction error of $B_{\rm tag}$ vertex. The kinematic approximation is also represented by a single Gaussian. On the other hand, the function form of the detector resolution, which is most dominating part among three components, is constructed by the Monte Carlo simulation. We utilize the fact that the detector resolution including the tail part can be represented by a superimposition of event-by-event vertex uncertainty (σ_{CP} and $\sigma_{\rm tag}$) scaled with a correction factors. Thus, the detector resolution is intrinsically derived from the real data. Moreover, we also determine the correction factors from real data, which is described in Subsection 4.2.4. We found small fraction of events belongs to large χ^2 tail of the vertex reconstruction and they have poorer resolution than σ_{CP} or $\sigma_{\rm tag}$ estimation. We introduce a correction to broaden the uncertainty of those events. These correction is estimated by the Monte Carlo simulation, while chi^2 quantities are also obtained by the event-by-event vertex reconstruction.

4.2.3 Background Shape

As for the "Δt" distribution for a fraction of the backgrounds, we consider they consist of two components. An event with B_{CP} reconstructed with the contaminating decay products from $B_{\rm tag}$ tends to be narrower Δt, because B_{CP} vertex is shifted toward the $B_{\rm tag}$ vertex by the contaminating tracks. We call such background "prompt" background. Other background components, whose vertices of B_{CP} and $B_{\rm tag}$ are reconstructed correctly, tends to show lifetime distribution, and we name this "lifetime" background. Such backgrounds can be considered to come from a certain B decay (mainly from $B \to J/\psi X$) one of which decay component is missed and incorrectly identified as $B^0 \to J/\psi K_S$ candidates. We convolute those two components with a "resolution function" [1] assuming it as a sum of two Gaussians determined by event-by-event error similarly to signal resolution function. The probability density function for the background fraction is defined as

$$
\begin{aligned}
\mathcal{P}_{\rm bkg} &= f_\delta \cdot \mathcal{R}_{\rm bkg}^{\rm prompt}(\Delta t) + \frac{(1-f_\delta)}{2\lambda} \cdot \int_{-\infty}^{\infty} {\rm d}\Delta t' \, \mathcal{R}_{\rm bkg}^{\rm life}(\Delta t' - \Delta t) \cdot {\rm e}^{-\frac{\Delta t'}{\lambda}}, \quad (4.28)\\
\mathcal{R}_{\rm bkg}(x) &= (1-f_{\rm tail}^{\rm bkg}) \cdot G(x;\, \mu_{\rm main}^{\Delta t,{\rm bkg}}, \sigma_{\rm main}^{\Delta t,{\rm bkg}}) + f_{\rm tail}^{\rm bkg} \cdot G(x;\, \mu_{\rm tail}^{\Delta t,{\rm bkg}}, \sigma_{\rm tail}^{\Delta t,{\rm bkg}}), \quad (4.29)
\end{aligned}
$$

where $\mathcal{R}_{\rm bkg}$ represents $\mathcal{R}_{\rm bkg}^{\rm prompt}$ and $\mathcal{R}_{\rm bkg}^{\rm life}$. The difference of $\mathcal{R}_{\rm bkg}^{\rm prompt}$ and $\mathcal{R}_{\rm bkg}^{\rm life}$ is the parameters for $\mu_{\rm main}^{\Delta t,{\rm bkg}}$ and $\mu_{\rm tail}^{\Delta t,{\rm bkg}}$. The standard deviations of $\mathcal{R}_{\rm bkg}^{\rm prompt}$ and $\mathcal{R}_{\rm bkg}^{\rm life}$ are defined commonly as

$$
\begin{aligned}
\sigma_{\rm main}^{\Delta t,{\rm bkg}} &= s_{\rm main}^{\rm bkg} \cdot \tilde{\sigma}_{\Delta z}, \quad (4.30)\\
\sigma_{\rm tail}^{\Delta t,{\rm bkg}} &= s_{\rm tail}^{\rm bkg} \cdot \tilde{\sigma}_{\Delta z}, \quad (4.31)
\end{aligned}
$$

where $\tilde{\sigma}_{\Delta z}$ is defined as equation (4.16). The background shape is determined in the sideband region of ΔE-$M_{\rm bc}$ plane with unbinned-maximum-likelihood method. The region is selected far from the signal region to eliminate the signal contamination (Figure 4.15

[1]It is not a real resolution function in an exact sense, because we cannot define "true Δt" for background events.

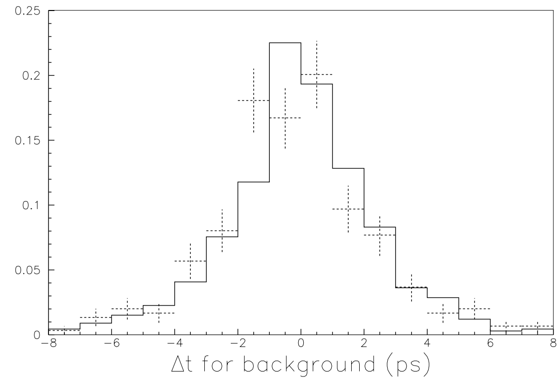


Figure 4.14: Distribution of background shapes in the sideband region (histogram) and in the signal region (crosses). Both distributions are obtained from $b\bar{b}$ Monte Carlo data.

(a)):

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

We checked consistency between the background shape in the sideband region and that in the signal region using $b\bar{b}$ Monte Carlo samples. Figure 4.14 shows the distribution of background shapes in the sideband region (histogram) and in the signal region (crosses). Figure 4.15 shows the distribution of background shape with superimposed fitted curve. The fitted parameters are listed in Table 4.7. Because of low statistics of $B^0 \to J/\psi K_S$ events, we perform fit combining $B^0 \to J/\psi K_S$ and $B^+ \to J/\psi K^+$ modes. In this fit, the fraction of long lifetime component is measured to be consistent to zero, and we set $f_\delta = 1$.

4.2.4 Determination of Resolution Function Parameters with Real Data

We have modeled the resolution function form in Subsection 4.2.2. In this subsection, with the real data, we determine the parameters for $\mathcal{R}$ that is modeled by the Monte Carlo simulation. First, we determine the $(s_{\rm main})_{\rm MC}$ by a scaling factor $s_{\rm data}$:

$$(s_{\rm main})_{\rm data} = s_{\rm data} \cdot (s_{\rm main})_{\rm MC}. \tag{4.32}$$

We replace all $(s_{\rm main})_{\rm MC}$ in the resolution function parameterization by $(s_{\rm main})_{\rm data}$. We determine the $s_{\rm data}$ by the lifetime measurement with $D^0 \to K^+\pi^-$ using both Monte Carlo sample and the real data. In the reconstruction of the D^0 lifetime, we can determine

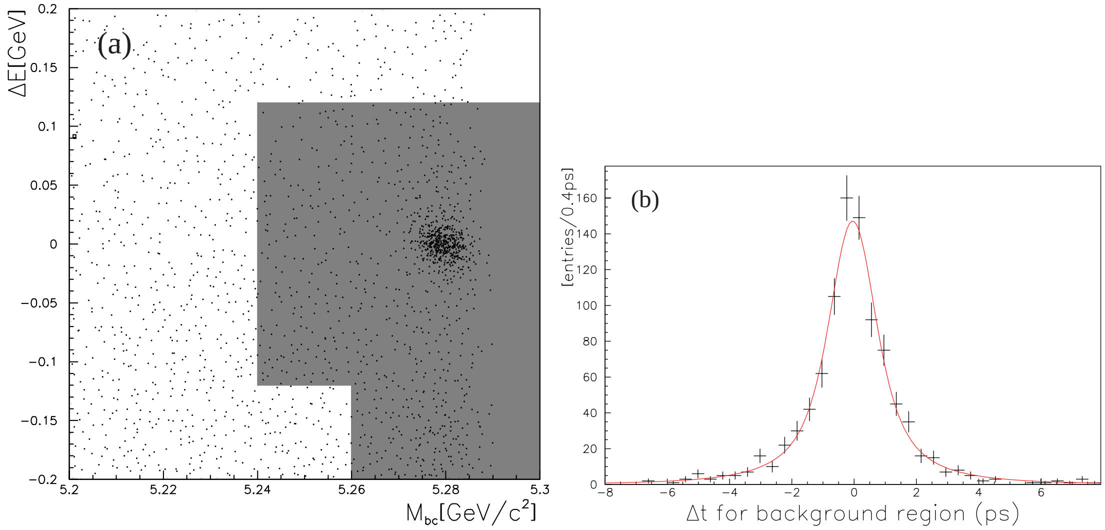


Figure 4.15: (a) ΔE-$M_{\rm bc}$ scattered plot for the combination of $B^0 \to J/\psi K_S$ and $B^+ \to J/\psi K^+$ events. The white area is used to determine the background shape distribution. (b) The distribution for Δt in the white area shown in (a). The fitted curve is superimposed.

Table 4.7: Parameters for background shape obtained from the fit as shown in Figure 4.15.

Fit parameters	Fitted values
$s_{\rm main}^{\rm bkg}$	1.08 ± 0.06
$s_{\rm tail}^{\rm bkg}$	3.3 ± 0.3
$f_{\rm tail}^{\rm bkg}$	0.14 ± 0.04
f_δ	1.00 (fixed)
λ (ps)	—
$\mu_{\rm main}^{\Delta t,{\rm bkg}}$ for prompt (μm)	-5.9 ± 4.5
$\mu_{\rm tail}^{\Delta t,{\rm bkg}}$ for prompt (μm)	-16 ± 34
$\mu_{\rm main}^{\Delta t,{\rm bkg}}$ for lifetime (μm)	—
$\mu_{\rm tail}^{\Delta t,{\rm bkg}}$ for lifetime (μm)	—

Table 4.8: Parameters obtained from the fit of real data.

Fit parameters	Fitted values
$s_{\rm tail}$	3.50 ± 0.89
$f_{\rm tail}$	0.027 ± 0.018
$\mu^0_{\rm main}$ (μm)	-21.4 ± 3.7
$\mu^0_{\rm tail}$ (μm)	151 ± 128

the resolution function parameters such as $s_{\rm main}$ [60]. Taking the ratio of $s_{\rm main}$ for the real data to the Monte Carlo simulation, we determined

$$s_{\rm data} = \frac{(s_{\rm main})_{D^0\ \rm data}}{(s_{\rm main})_{D^0\ \rm MC}} = 0.834 \pm 0.014. \tag{4.33}$$

After the determination of $s_{\rm data}$, we perform fit on the real data of hadronic event samples, $J/\psi K_S$, $J/\psi K^{*0}(K^+\pi^-)$, $D^+\pi^-$, $D^{*+}\pi^-$, $D^{*+}\rho^-$, $J/\psi K^+$, and $D^0\pi^+$, floating B meson lifetimes, $s_{\rm tail}$, and $f_{\rm tail}$. Together with those parameters, we treat $\mu^0_{\rm main}$ and $\mu^0_{\rm tail}$ as free parameters in the fit to estimate the alignment discrepancy between the real data and the Monte Carlo samples. In the fit, we obtain the lifetimes as $\tau_{B^0} = 1.58 \pm 0.04$ ps and $\tau_{B^+} = 1.72 \pm 0.03$ ps. The determined resolution parameters as listed in 4.8. We use above values for the $\sin 2\phi_1$ determination instead of values obtained in Section 4.2.2. To summarize, parameters for the resolution function are:

from the read data:	σ_{CP}, $\sigma_{\rm tag}$, $(\chi^2/n)_{CP}$, $(\chi^2/n)_{\rm tag}$, $s_{\rm data}$, $s_{\rm tail}$, $f_{\rm tail}$, $\mu^0_{\rm main}$, and $\mu^0_{\rm tail}$.
from the Monte Carlo simulation:	$(s_{\rm main})_{\rm MC}$, α_{CP}, $\alpha_{\rm tag}$, $s^{\rm NP}_{\rm main}$, $s^{\rm NP}_{\rm tail}$, $\mu^1_{\rm main}$, $\mu^1_{\rm main}$, $\sigma^{\rm main}_{\rm k}$, and $\sigma^{\rm tail}_{\rm k}$

Figure 4.16 shows average shape of the resolution function for the $J/\psi K_S$ sample in real data. The solid curve represents a sum of the main resolution and the tail resolution, and broken curve represents the tail resolution solely. The typical resolution (r.m.s.) and mean shift are measured to be 1.46 ps and -0.19 ps, respectively. The main part of the Gaussian is fitted to be 0.77 ps, which corresponds to the Δz resolution of 98 μm.

4.3 Fit Results

The fit to the reconstructed Δt distribution is performed with maximum likelihood fit method using the probability density function described above. The following B^0 lifetime is assumed:

$$\tau_{B^0} = 1.548 \pm 0.032 \text{ ps},$$

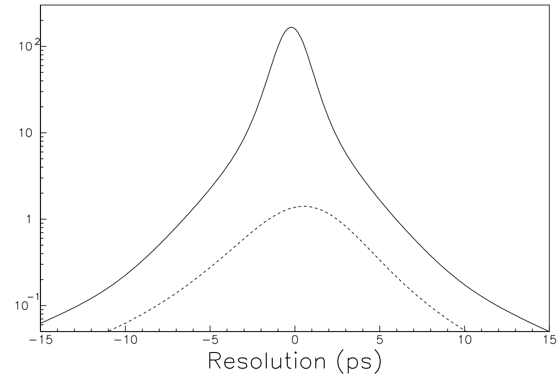


Figure 4.16: Resolution function shape obtained by the $J/\psi K_S$ data sample. Solid line represents a superimposition of a sum two Gaussians and the dotted line represents that of tail Gaussian. The typical resolution (r.m.s.) and mean shift are measured to be 1.46 ps and -0.19 ps, respectively.

which is quoted from the world average [49]. The parameter for B^0-$\overline{B}^0$ mixing is also quoted from world average;

$$\Delta M = 0.472 \pm 0.017 \text{ ps}^{-1}.$$

The central value of each parameter is used. The $\sin 2\phi_1$ value determined from produced 31.3×10^6 $B\overline{B}$ event is

$$\sin 2\phi_1 = 0.81 \pm 0.20(\text{stat}).$$

Figure 4.17 shows the Δt distributions together with the fit curves. The distributions for $q = +1$ and $q = -1$ are separately shown in the figure with white and black points, respectively. The combined distribution of $q = \pm 1$ for Δt and the background distribution are shown in Figure 4.18. The background contamination for the Δt distribution is shown to be quite small. Figure 4.19 shows the distribution for $-2\ln(\mathcal{L}/\mathcal{L}_{\max})$ as a function of $\sin 2\phi_1$. The distribution shows the parabolic shape, and thus the Gaussian assumption for the estimation of the statistical error stands.

4.4 Systematic Uncertainties

We estimate the systematic uncertainties by varying various parameters for the B^0 reconstruction or the fit procedure. The systematic uncertainties from the sources examined in this section are summarized in Table 4.9. A total systematic uncertainty is calculated by a sum of them in quadrature. The total uncertainty is estimated to be ± 0.04, which is one fifth of the statistical uncertainty.

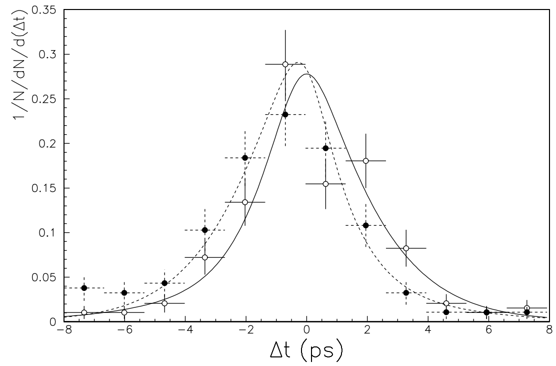


Figure 4.17: Δt distributions for the events with $q = +1$ (white points) and $q = -1$ (black points). The results of global fit (with $\sin 2\phi_1 = 0.81$) are shown as solid and dotted curves, respectively.

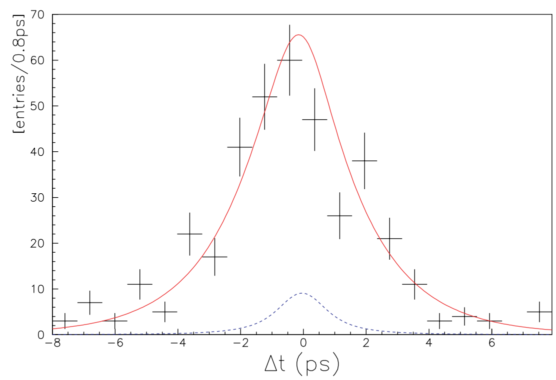


Figure 4.18: Sum of the Δt distributions for $q = \pm 1$ with fit curve. The broken line represent the distribution for background event.

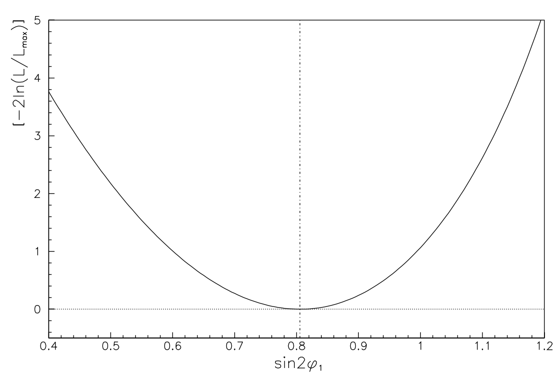


Figure 4.19: Values of $-2\ln(\mathcal{L}/\mathcal{L}_{\max})$ versus $\sin 2\phi_1$.

- B decay position dependence
  The interaction point distribution obtained from $B\overline{B}$ event samples differs by 10 μm from Bhabha events. We take into account the uncertainty of $\sin 2\phi_1$ due to this discrepancy, which is measured to be quite small. The interaction point profile includes the effect of the B meson flight length in the r-ϕ plane. We approximate this effect by convolution of Gaussian but it does not well reproduce the distribution since it is actually an exponential function. We also estimate the uncertainty due to the approximation by varying the standard deviation of the B meson flight length by ± 10 μm from the nominal value of 21 μm.

- $|\Delta t|$ tail
  We changed the track selection criteria to estimate the systematic trend of $\sin 2\phi_1$ on the evaluation of relative poorly reconstructed Δt. Each track selection criteria, $|\delta z| < 1.8$ mm, $|\sigma z| < 500$ μm, and $|\delta r| < 500$ μm, has been varied by 10% to estimate the systematic error associated with the modeling of the badly measured events.

  The criteria for the reconstructed vertices with the χ^2/n for B_{CP}, is varied from 3 to 27, and repeated the fit to estimate the validity of representation for the Δt tail part.

  Another check of the Δt-tail effect is performed by varying the $|\Delta t|$ region used for the $\sin 2\phi_1$ determination. We use the region of $|\Delta t| < 15$ ps in the fit. The fit region is varied by -10 ps, *i.e.*, $|\Delta t| < 5$ ps.

- Fit bias and Monte Carlo statistics
  A possible bias is studied using Monte Carlo samples that are generated by fully simulating the detector response. The difference between the input value of $\sin 2\phi_1$ for the signal Monte Carlo and the $\sin 2\phi_1$ value obtained from the fit may indicate the bias on $\sin 2\phi_1$ measurement. The resolution function parameters and wrong tag fractions are substituted by the appropriate values for the Monte Carlo sample. We obtain $\sin 2\phi_1 = 0.495 \pm 0.012$, while the input value is $\sin 2\phi_1 = 0.500$. The fit result is well consistent with the input value. We quote the statistical error of the fit as systematic error for possible bias in the fit procedure.

- Dependence on physics parameters
  The dependence on the B^0 lifetime, and the B^0-$\overline{B}^0$ mixing parameter are evaluated by varying the parameters by $\pm 1\sigma$ of world average [49]. The dependence on the B^0 lifetime is estimated to be quite small, and the uncertainty of the mixing parameter, ΔM, dominates the uncertainty due to the physics parameters, although it is still small compared with other systematic sources.

- Resolution function
  The $\sin 2\phi_1$ dependence on the resolution function is tested. The parameters obtained from Monte Carlo simulation are varied by $\pm 2\sigma$, and those from real data are varied by $\pm 1\sigma$. The dependences on these parameters are negligible except for $f_{\rm tail}$, which dominates the systematic uncertainties due to resolution function. The $f_{\rm tail}$ is determined from the fit to the real data, as described in Section 4.2.4, and the statistical error of $f_{\rm tail}$ is large, which causes large systematic uncertainty due to $f_{\rm tail}$. The effect of mean shift of resolution function directly changes the number of events in $\Delta t < 0$ region and $\Delta t > 0$ region. The parameters of $\mu^0_{\rm main}$ and $\mu^0_{\rm tail}$ are determined from the real data. However, the systematic dependences on $\mu^0_{\rm main}$ and $\mu^0_{\rm tail}$ are relatively small, $\lesssim 7 \times 10^{-3}$, and the dependences on $\mu^1_{\rm main}$ and $\mu^1_{\rm tail}$ are much smaller than them, $\lesssim 5 \times 10^{-4}$.

- Wrong tag fraction
  Each wrong tag fraction in the six r regions is varied by $\pm 1\sigma$ and the fits are repeated. The total uncertainty is estimated to be ± 0.024, which is the largest components among all the sources. This is due to large statistical errors in the determination of the wrong tag fractions.

- Signal probability
  The event-by-event signal fraction is determined from the fit to the ΔE-$M_{\rm bc}$ distribution in the data. The systematic error associated to the fit is estimated by varying the parameters determining signal fractions by $\pm 1\sigma$. The difference is measured to be quite small.

- Background fraction
  The systematic dependence of the background fraction is tested by varying ΔE-$M_{\rm bc}$ signal region from 20 MeV to 60 MeV for ΔE (nominal value is 40 MeV) and from 7 MeV/c^2 to 13 MeV/c^2 for $M_{\rm bc}$ (nominal value is 10 MeV/c^2) and fits are repeated. Small systematic trend on the ΔE or $M_{\rm bc}$ range is observed.

- Background shape
  The systematic error associated to the determination of the background shape is estimated by varying the parameters by $\pm 1\sigma$. Because B meson reconstruction via $B^0 \to J/\psi K_S$ mode is quite cleanly performed with few background fraction, the systematic difference is almost negligible.

4.5 Validity Examinations

In the following subsections, we discuss the validity of the fit procedure of $\sin 2\phi_1$ and its error. First, we test the fit procedure using a parameterized toy Monte Carlo simulation

Table 4.9: Summary of systematic uncertainties for $\sin 2\phi_1$ measurement with $B^0 \to J/\psi K_S$ decay. The errors are combined in quadrature.

	Uncertainties	
B decay position	-0.001	$+0.010$
$\|\Delta t\|$ tail	-0.021	$+0.012$
Monte Carlo statistics	-0.012	$+0.012$
Physics parameters	-0.008	$+0.003$
Wrong flavor tagging	-0.024	$+0.024$
Resolution function	-0.014	$+0.016$
Signal probability	-0.004	$+0.006$
Background fraction	-0.005	$+0.006$
Background shape	-0.001	$+0.001$
Total	-0.038	$+0.036$

in Sections 4.5.1 and 4.5.2. Then, we compare the observed uncertainty of $\sin 2\phi_1$ and the expected statistical uncertainty from analytical computation in Section 4.5.3. The data samples from CP mixed final states are also analyzed to check possible bias in the determination of $\sin 2\phi_1$ value in Section 4.5.4. In Section 4.5.5, we evaluate the correlation of the $\sin 2\phi_1$ value with the flavor tag. We also check correlations of the $\sin 2\phi_1$ value and ΔM or $|\lambda_{J/\psi K_S}|$ in the end of this section.

4.5.1 Ensemble Test

To check the fit procedure, we use parameterized Monte Carlo samples, which is generated according to the probability density function, $\mathcal{P}(\Delta t;\ \sin 2\phi_1)$, as we established in Subsection 4.2. The input variables for the probability density function such as σ_{CP}, $(\chi^2/n)_{CP}$, $\sigma_{\rm tag}$, and $(\chi^2/n)_{\rm tag}$ are also generated according to the corresponding distributions observed in the real data. We generate 1000 sets of Monte Carlo samples, each of which contains 387 $B^0 \to J/\psi K_S$ events. The input value of $\sin 2\phi_1$ is 0.81. Figure 4.20 (a) shows the $\sin 2\phi_1$ distributions determined by the fits. The mean value of the distribution is 0.81, which is consistent to the input $\sin 2\phi_1$. The standard deviation of the distribution is 0.19, which is also consistent with the uncertainty obtained from the data. Figure 4.20 (b) shows the distribution for $(\sin 2\phi_1(\text{fit}) - \sin 2\phi_1(\text{input}))/(\text{error of fit})$. The standard deviation of the distribution is 1.00. The result indicates that our fit correctly estimates the errors of $\sin 2\phi_1$. Figures 4.20 (c) and (d) show the distribution of negative and positive errors estimated by the fits, respectively. The vertical line indicates the error estimated by the fit using the real data. The estimated error also resides within the Monte Carlo prediction.

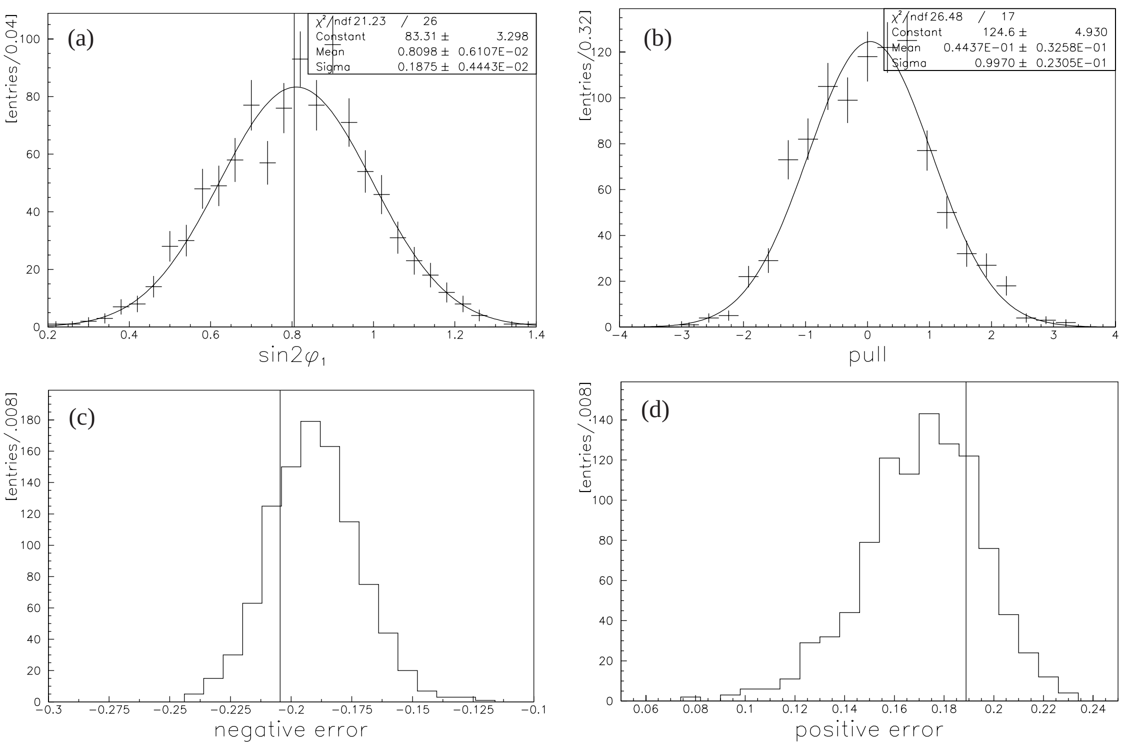


Figure 4.20: Results of toy Monte Carlo simulation. Each plot shows the distribution for (a) $\sin 2\phi_1$ values obtained from the fits, (b) $(\sin 2\phi_1(\text{fit}) - \sin 2\phi_1(\text{input}))/(\text{error of fit})$, (c) the negative error, and (d) positive error for $\sin 2\phi_1$.

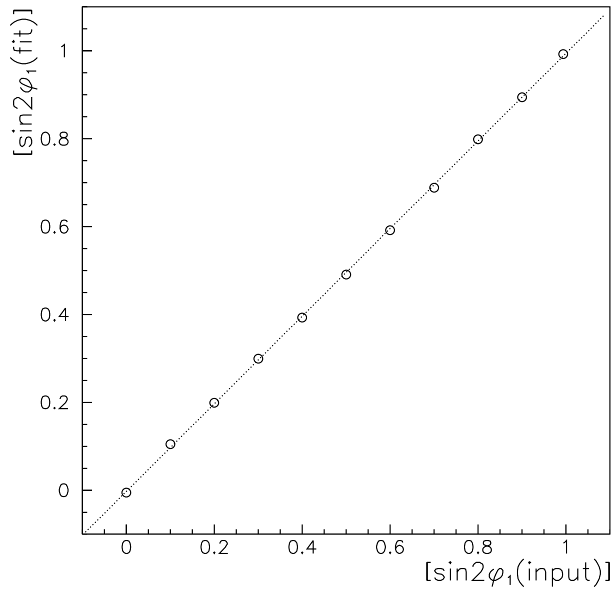


Figure 4.21: Result of linearity test. Horizontal axis represents the input of $\sin 2\phi_1$ value and vertical axis represents the $\sin 2\phi_1$ value obtained from the fits. The broken line shows a fit to a linear function.

4.5.2 Linearity Test

Another test using a parameterized Monte Carlo simulation is made to check a linearity of $\sin 2\phi_1$ determined by the fit with respect to the true $\sin 2\phi_1$. We generate 1000 experiments containing 387 events for every input value of $\sin 2\phi_1$. Figure 4.21 shows the obtained $\sin 2\phi_1$ value as a function of the input $\sin 2\phi_1$ value. The points represent the mean of the obtained $\sin 2\phi_1$ values. The dotted line in the figure represents a fit to a function:

$$\sin 2\phi_1(\text{fit}) = a + b \times \sin 2\phi_1(\text{input}).$$

The fit yields $a = -0.002 \pm 0.004$ and $b = 0.997 \pm 0.006$, which indicates a good linearity.

4.5.3 Expected Statistical Uncertainty

The statistical error of $\sin 2\phi_1$ ($\sigma_{\sin 2\phi_1}$) can be analytically computed using a number of events (N), a "dilution" factor[2] (d), and the wrong tag fraction (w) as

$$\sigma^2_{\sin 2\phi_1} \simeq \frac{1}{d^2(1-2w)^2 N}, \qquad (N \gg 1). \tag{4.34}$$

[2]The definition of the "dilution" factor is complicated. When d is large, the significance of $\sin 2\phi_1$ is enhanced.

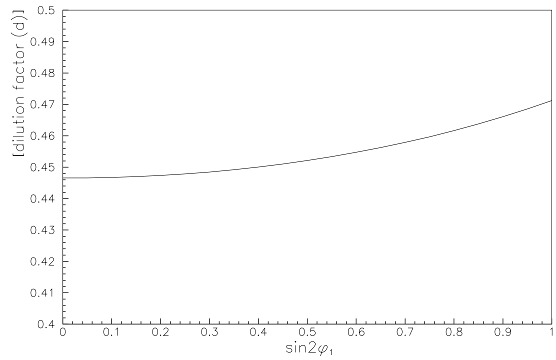


Figure 4.22: Dilution factor as a function of $\sin 2\phi_1$ for the nominal resolution function shown in Figure 4.16. The wrong tag fraction of 0.240 is used.

We can ignore the background contribution to the dilution factor because it is negligibly small. Approximating the event-by-event resolution function as a global resolution function shown in Figure 4.16, it is obtained from following equation [61]:

$$d^2 = \int_{-\infty}^{\infty} \mathrm{d}(\Delta t)\, \frac{1}{\mathcal{P}_{\mathrm{sig}}(\Delta t)} \left[\frac{\partial \mathcal{P}_{\mathrm{sig}}(\Delta t)}{\partial\left((1-2w)\cdot\sin 2\phi_1\right)} \right]^2 . \tag{4.35}$$

Figure 4.22 shows the dilution factor as a function of $\sin 2\phi_1$. For this plot we use the average wrong tag fraction of $w = 0.240$, which is obtained from $\varepsilon_{\mathrm{eff}} = 27.0\%$. The error is calculated to be 0.21 by equation (4.34), which is consistent with the statistical error obtained from the fit to the real data.

4.5.4 Tests Using Non-CP Samples

We extract "$\sin 2\phi_1$" value from the flavor-specific decay final states in the real data to verify the procedures for the flavor tag algorithm the proper-time difference reconstruction, the resolution function, and the fit procedures. Since the flavor specific decay modes are not expected to give CP violation in this analysis, the "$\sin 2\phi_1$" values should be zero. $B^0 \to J/\psi K^{*0}(K^+\pi^-)$, $B^0 \to D^+\pi^-$, $B^0 \to D^{*+}\pi^-$, $B^0 \to D^{*+}\rho^-$, and $B^0 \to D^{*-}\ell^+\nu_\ell$ samples are used for this analysis. The selection of "control" samples are described in Appendix A. The fit results are listed in Table 4.10 with statistical errors only. Combining all the modes, we obtain "$\sin 2\phi_1$" $= 0.05 \pm 0.04$. The measurements are consistent with zero within the statistical errors. The result indicates no significant bias in the procedure to determine the $\sin 2\phi_1$ value.

We can make similar test using the $J/\psi K_S$ sample by assuming all B_{tag} mesons as B^0 (or $\overline{B}^0$) mesons. The "$\sin 2\phi_1$" is obtained to be $-0.14^{+0.21}_{-0.20}$, which also show no significant bias.

Table 4.10: "$\sin 2\phi_1$" values from non-CP samples.

	$J/\psi K^{*0}(K^+\pi^-)$	DX	$D^{*-}\ell^+\nu_\ell$
# of events	816	5560	10232
"$\sin 2\phi_1$"	0.01 ± 0.14	0.12 ± 0.06	0.01 ± 0.05

Table 4.11: r dependence of the $\sin 2\phi_1$ values. Only statistical errors are shown.

	$0 < r \leq 0.5$	$0.5 < r \leq 0.75$	$0.75 < r \leq 0.875$	$0.875 < r \leq 1$
# of events	213	75	42	57
$\sin 2\phi_1$	$0.41^{+0.88}_{-0.91}$	$0.30^{+0.40}_{-0.41}$	$1.43^{+0.43}_{-0.50}$	$0.91^{+0.22}_{-0.27}$

4.5.5 Tests Associated with Flavor Identification

The flavor dependence of $\sin 2\phi_1$ determination is also tested. When the Δt asymmetry of the resolution function is incorrectly modeled, the effect will show up in the discrepancy of the $\sin 2\phi_1$ values between $q = +1$ events and $q = -1$ events. The results are $\sin 2\phi_1 = 0.70^{+0.29}_{-0.31}$ for $B_{\rm tag} = B^0$ events, and $\sin 2\phi_1 = 0.96^{+0.29}_{-0.43}$ for $B_{\rm tag} = \overline{B}^0$ events. They are consistent with each other indicating no significant discrepancy of the asymmetry in the Δt reconstruction. We also repeat $\sin 2\phi_1$ determination for different r regions. The results are listed in Table 4.11. The $\sin 2\phi_1$ value for smaller r region is determined with larger statistical errors due to small information about the flavor determination. We find no systematic variation in the fit results.

4.5.6 Fit with ΔM

We performed a fit while treating $\sin 2\phi_1$ and ΔM as free parameters. The result is consistent with that obtained fixing ΔM within statistical errors:

$$\begin{aligned} \sin 2\phi_1 &= 0.81\ ^{+0.18}_{-0.20}\ (\text{stat}), \\ \Delta M &= (0.43\ ^{+0.09}_{-0.10}\ (\text{stat}))\ \text{ps}^{-1}. \end{aligned}$$

Figure 4.23 shows the contour plot for ΔM and $\sin 2\phi_1$.

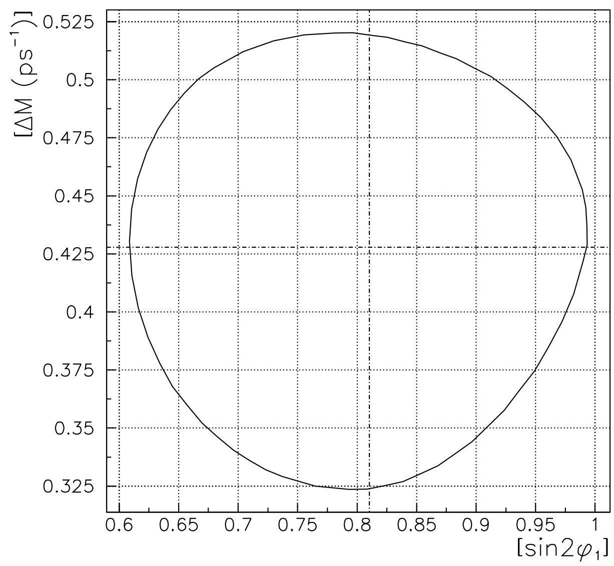


Figure 4.23: Contour plot for ΔM versus $\sin 2\phi_1$. The curve represents the border of 1σ from the best fitted position.

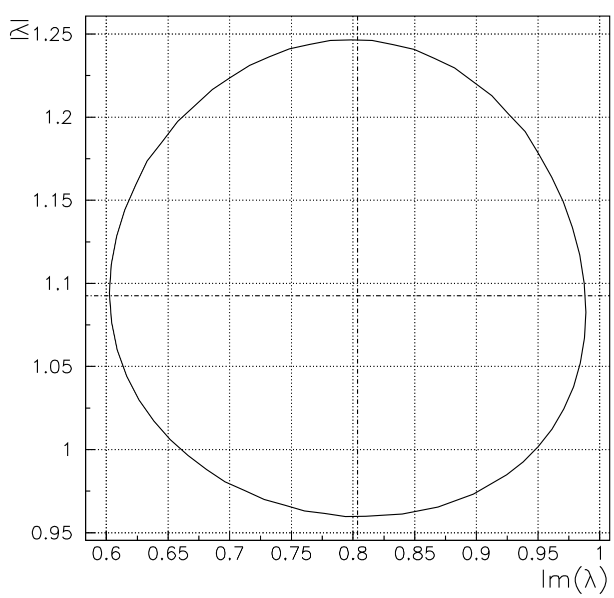


Figure 4.24: Contour plot for $|\lambda_{J/\psi K_S}|$ versus $\mathcal{I}m(\lambda_{J/\psi K_S})$. The curve represents the border of 1σ from the best fitted position.

4.5.7 Fit without $|\lambda_{J/\psi K_S}| = 1$ Assumption

The the probability density function without the assumption of $|\lambda_{J/\psi K_S}| = 1$ is derived from equations (1.69) and (1.70) as

$$\begin{aligned} f_{\ell^\pm}(\Delta t;\ \sin 2\phi_1) \quad &= \quad \frac{1}{1+|\lambda_{J/\psi K_S}|^2}\frac{\mathrm{e}^{-\frac{|\Delta t|}{\tau_{B^0}}}}{2\tau_{B^0}}\left[\frac{1+|\lambda_{J/\psi K_S}|^2}{2}\right. \\ &\quad \left. + \, q\cdot\mathcal{I}m(\lambda_{J/\psi K_S})\cdot\sin(\Delta M\Delta t) - q\cdot\frac{1-|\lambda_{J/\psi K_S}|^2}{2}\cdot\cos(\Delta M\Delta t)\right]. \end{aligned} \tag{4.36}$$

The normalization factor is calculated so that a sum of probability density functions for B^0 and $\overline{B}^0$ be unity. We perform a fit with $|\lambda_{J/\psi K_S}|$ and $\mathcal{I}m(\lambda_{J/\psi K_S})$ treated as free parameters on the Δt distribution of $B^0 \to J/\psi K_S$ events. The fit yields

$$\begin{aligned} |\lambda_{J/\psi K_S}| \quad &= \quad 1.09 \pm 0.14(\mathrm{stat}), \\ \mathcal{I}m(\lambda_{J/\psi K_S}) \quad &= \quad 0.80 \pm 0.19(\mathrm{stat}). \end{aligned}$$

Figure 4.24 shows the contour plot for $|\lambda_{J/\psi K_S}|$ and $\mathcal{I}m(\lambda_{J/\psi K_S})$.

Chapter 5

Discussions

So far we have described the observation of the CP violation and the measurement of $\sin 2\phi_1$. In this section, we discuss the significance of our measurement.

The obtained $\sin 2\phi_1$ is

$$\sin 2\phi_1 = 0.81 \pm 0.20 \text{ (stat)} \pm 0.04 \text{ (syst)}.$$

The CP conservation, *i.e.* $\sin 2\phi_1 = 0$, is ruled out at 4σ level. To obtain the confidence intervals of $\sin 2\phi_1$ we use the Feldman-Cousins method [62][63] since the obtained value is close to the physical boundary. Feldman-Cousins method correctly incorporates the effect of the physical boundary as described in Appendix D. We determine confidence intervals as $0.60 \leq \sin 2\phi_1 \leq 1.00$ and $0.40 \leq \sin 2\phi_1 \leq 1.00$ at the confidence levels of 68% and 95%, respectively. Figure 5.1 shows the Feldman-Cousins confidence intervals for ϕ_1 in the ρ-η plane (ρ is horizontal). Broken two lines indicate the obtained value of $\sin 2\phi_1$ *i.e.* 0.81. The lightly and heavily hatched area represent the 68% and 95% confidence intervals, respectively. A recent theoretical constraint, $\sin 2\phi_1 = 0.70 \pm 0.07$, is also superimposed. The experimental result and the theoretical constraint cover each other. The measured value of $\sin 2\phi_1$ is consistent with the Standard Model constraint. However, the statistical error is not small enough to be conclusive. For the verification of the Standard Model through KM-mechanism, more data are necessary.

We perform the measurement of the CP violation by combining following fully reconstructed B decays with $B^0 \to J/\psi K_S$:

- $J/\psi K_S, \quad K_S \to \pi^0\pi^0$,
- $\psi(2S)K_S, \quad \psi(2S) \to \ell^+\ell^-$,
- $\psi(2S)K_S, \quad \psi(2S) \to J/\psi\pi^+\pi^-$,
- $\chi_{c1}K_S, \quad \chi_{c1} \to J/\psi\gamma$,

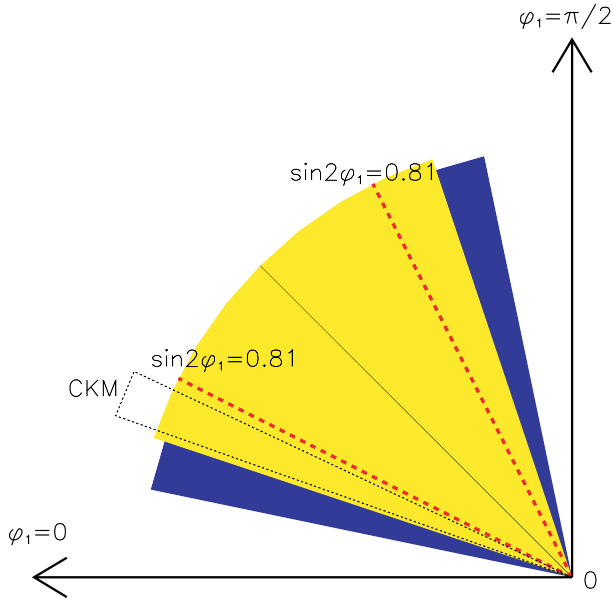


Figure 5.1: Obtained $\sin 2\phi_1$ (broken two lines) and Feldman-Cousins confidence intervals (light hatch for 68% and heavy hatch for 95% confidence levels, respectively). Theoretical constraint of $\sin 2\phi_1 = 0.70 \pm 0.07$ is also superimposed. The diagonal dotted line represents the boundary of $\sin 2\phi_1 = 1$.

- $\eta_c K_S, \quad \eta_c \to K_S K^\pm \pi^\mp$,
- $\eta_c K_S, \quad \eta_c \to K^+ K^- \pi^0$,
- $J/\psi K^{*0}, \quad K^{*0} \to K_S \pi^0$,
- $J/\psi K_L$.

All CP eigenvalues of above decay modes are -1 except for $J/\psi K_L$ and the $J/\psi K^{*0}$. The CP eigenvalue of the $B^0 \to J/\psi K_L$ is $+1$. The final state of $J/\psi K^{*0}$ is a mixture of the CP even and the CP odd states, where the fraction of $CP = -1$ is $0.19 \pm 0.04\,(\text{stat}) \pm 0.04\,(\text{syst})$ [64]. We reconstruct 1137 B^0 candidates from those decay modes including $J/\psi K_S$ in 29.1 fb^{-1} data. We perform unbinned-maximum-likelihood fit on the Δt distribution and determine $\sin 2\phi_1 = 0.99 \pm 0.14(\text{stat}) \pm 0.06(\text{syst})$. The probability to observe $\sin 2\phi_1 > 0.99$ if CP is conserved actually, is negligibly small. Using Feldman-Cousins method, we determine the confidence intervals as $0.85 \leq \sin 2\phi_1 \leq 1.00$ and $0.70 \leq \sin 2\phi_1 \leq 1.00$, at the 68% and 95% confidence levels, respectively.

To test the internal consistency, we perform $\sin 2\phi_1$ extraction from sub data samples. We obtained $\sin 2\phi_1 = 1.00 \pm 0.40(\text{stat}) \pm 0.08(\text{syst})$ with $CP = -1$ modes except for $B^0 \to J/\psi K_S$, $\sin 2\phi_1 = 1.31 \pm^{+0.19}_{-0.13}\,(\text{stat}) \pm 0.12(\text{syst})$ with $B^0 \to J/\psi K_L$, and $\sin 2\phi_1 = 0.97 \pm^{+1.38}_{-1.40}\,(\text{stat}) \pm 0.19(\text{syst})$ with $B^0 \to J/\psi K^{*0}$. The χ^2 of the fit is 2.3 for two degrees

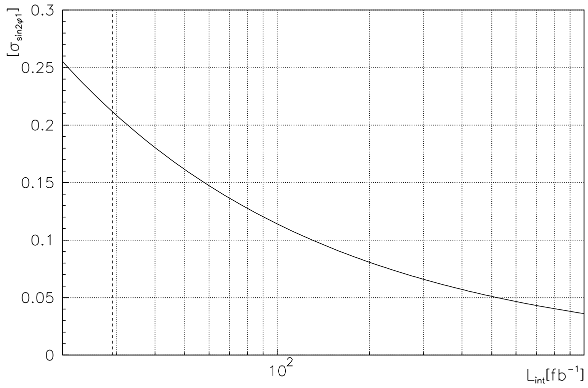


Figure 5.2: Expected statistical error of $\sin 2\phi_1$ as a function of integrated luminosity for $\sin 2\phi_1 = 0.80$. The broken vertical line indicates the integrated luminosity of 29.1 fb^{-1}.

of freedom, which indicates a good internal consistency.

To evaluate the validity of the KM-model, we still need more data, since the statistical error is the dominant contribution to the $\sin 2\phi_1$ uncertainty. By the end of 2002, we will accumulate ~ 100 fb^{-1} data sample. The statistical error is expected to be reduced from 0.2 to 0.1. Figure 5.2 shows the expected statistical error of $\sin 2\phi_1$ as a function of integrated luminosity for $\sin 2\phi_1 = 0.8$. The largest contribution to the systematic uncertainty is the uncertainties of the wrong tag fractions. The uncertainties in the fractions is mainly due to the statistical limit. The increase of the data will reduce the systematic error associated to this source. Another significant source of the systematic uncertainty is the understanding of the tail part of the resolution function. Although they are still small with respect to the statistical error at the integrated luminosity of 29.1 fb^{-1}, the improvement of systematic uncertainties toward the enormous integrated luminosity must be inevitable. The major reason for this source is that the resolutions for badly reconstructed vertices are not described very properly. Examination of the contributing events for the tail components of the resolution function will provide elaborated result. Better description of the resolution function is in progress in the B meson lifetime analysis, where much larger statistics is available.

Chapter 6

Conclusions

We have collected 31.3 million $B\overline{B}$ pairs at the KEK B-factory. We have reconstructed 387 $B^0 \to J/\psi K_S$ candidates associated with the flavor information and the proper-time difference of the two B decays. We perform unbinned-maximum-likelihood fit on the proper-time difference distributions and obtain the CP violation parameter in the Standard Model as

$$\sin 2\phi_1 = 0.81 \pm 0.20(\text{stat}) \pm 0.04(\text{syst}).$$

CP invariance is ruled out at 4σ level. We also study the CP violation with combining other charmonium+K^0 decay modes with $B^0 \to J/\psi K_S$. We determine

$$\sin 2\phi_1 = 0.99 \pm 0.14(\text{stat}) \pm 0.06(\text{syst}).$$

The probability to obtain $\sin 2\phi_1 \geq 0.99$ when the true value of $\sin 2\phi_1$ is zero is negligibly small. These results are the first definitive observation of the CP violation in the B meson system. Although the results are consistent to the theoretical prediction, we need more statistics to conclude whether the CP violation resides within or beyond the Standard Model. The measurement of the CP violation stimulate the further efforts for the precise $\sin 2\phi_1$ determination of the unitarity triangles.

Appendix A

Reconstruction of Other Modes than $B^0 \to J/\psi K_S$ Mode

In this chapter, we give brief description of the reconstruction of following decay modes:

- $B^0 \to J/\psi K^{*0}(K^+\pi^-)$,
- $B^0 \to D^+\pi^-$,
- $B^0 \to D^{*+}\pi^-$,
- $B^0 \to D^{*+}\rho^-$,
- $B^0 \to D^{*-}\ell^+\nu_\ell$,

for neutral B decays, and

- $B^+ \to J/\psi K^+$,
- $B^+ \to D^0\pi^+$,

for charged B decays. These decay modes are used to reconstruct B^0 or B^+ for the estimation of wrong flavor identification probability, for the tests on control sample. We discuss the lifetime analysis in Appendix E in terms of the improvement of the resolution function using hadronic modes among the decay modes listed above.

We name reconstruction side B meson as B_{rec} instead of B_{CP} and the associated side B meson as B_{asc} instead of B_{tag}, because the decay final state of B_{rec} is not CP eigenstate.

A.1 Light Meson Reconstruction

Decay modes of several light mesons used to reconstruct the listed B decays are summarized in Table A.1. Invariant mass ranges used to select them are also shown. A

Table A.1: Decay modes and invariant mass ranges used to select the light mesons.

Decay mode	Invariant mass range (MeV/c^2)
$\pi^0 \to \gamma\gamma$	$124 < M_{\gamma\gamma} < 146$
$\rho^- \to \pi^-\pi^0$	$\lvert M_{\pi^-\pi^0} - M_{\rho^-}\rvert < 150$
$K^{*0} \to K^-\pi^+$	$\lvert M_{K^-\pi^+} - M_{K^{*0}}\rvert < 150$

mass constraint fit is performed to improve the π^0 momentum resolution. A minimum momentum of 200 MeV is required for π^0 candidates. K^{*0} reconstruction is performed with oppositely charged two tracks, where track should be positively tagged as kaon with as $P(K) > 0.4$, and one track should be negated to be kaon.

A.2 Charmed Meson Reconstruction

Three D^0 decay modes, $D^0 \to K^-\pi^+$, $D^0 \to K^-\pi^+\pi^0$, $D^0 \to K^-\pi^+\pi^+\pi^-$, are used to reconstruct the D^0 meson, and $D^+ \to K^-\pi^+\pi^+$ decay is used to reconstruct the D^+ meson. At least two tracks are required to satisfy the SVD hit association criterion for $\overline{B}^0 \to D^+\pi^-$, $\overline{B}^0 \to D^{*+}\pi^-$, $\overline{B}^0 \to D^{*+}\rho^-$, and $B^- \to D^0\pi^-$ ($\overline{B} \to D\pi$) analysis.

D^{*+} candidates are formed by combining D^0 candidates with soft π^+. The kinematic selection of D^{*+} is made on the mass difference: $\Delta M_{D^{*+}} \equiv M_{D^0\pi^+} - M_{D^0}$. The $\Delta M_{D^{*+}}$ resolution is greatly improved by re-fitting the soft π^+ track with a constraint that it originates from the D^0 production point (or B^0 decay point). Figures A.1 show ΔM distribution for D^{*+} candidates.
The points with error bar indicate data points while the solid curve indicates the fit result. Signal is represented by a sum of two Gaussians while background is represented by a phase-space function; $a(\Delta M_{D^*} - x_0)^p \cdot \mathrm{e}^{-b(\Delta M_{D^*} - x0)}$.
The background function is indicated by the dotted curve in the figure. The typical standard deviation for main Gaussians are approximately 0.35 MeV/c^2 for $\Delta M_{D^{*+}}$.
We apply $\Delta M_{D^{*+}}$ and invariant mass range selections for the charmed meson reconstruction. The ranges depend on the decay modes, where each parameter is listed in Table A.2.

A.3 B Meson Reconstruction

In this section, we describe the reconstructions of B mesons with hadronic decay final states and semileptonic decay final state.

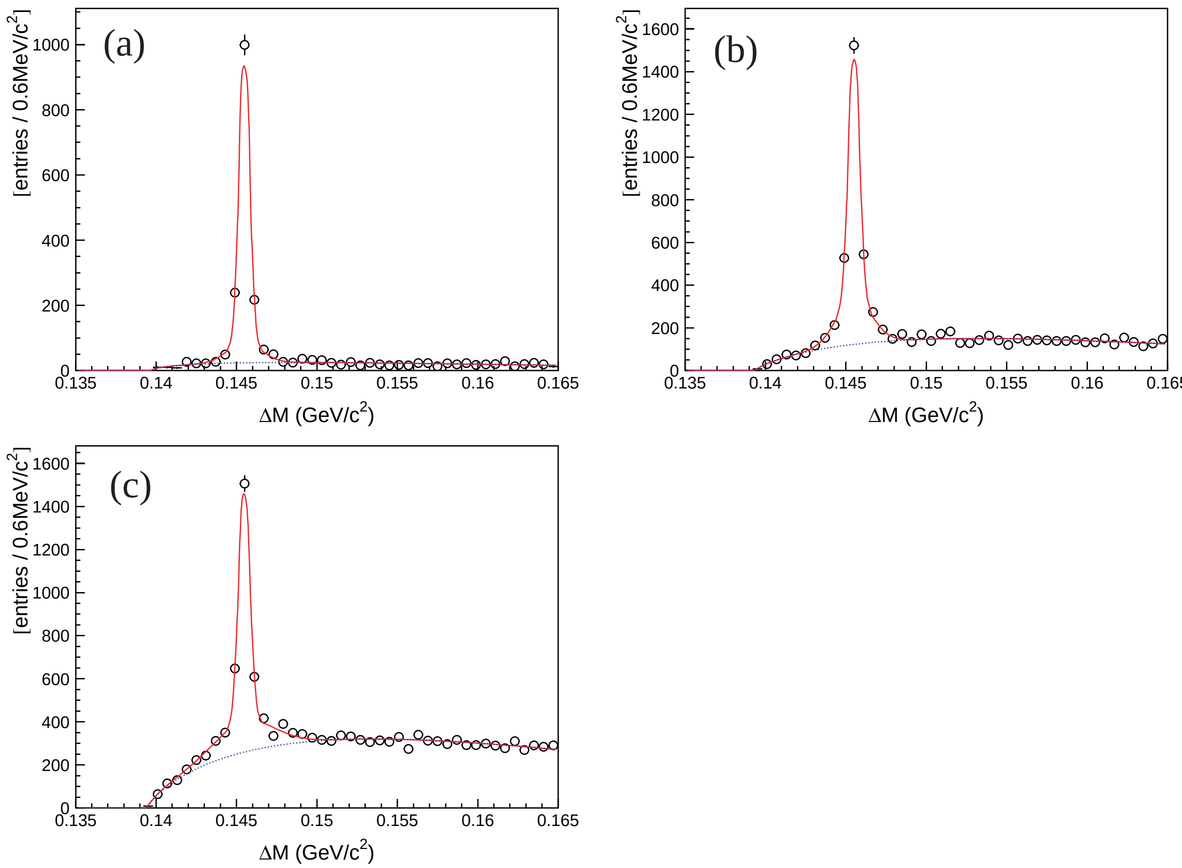


Figure A.1: $\Delta M_{D^{*+}}$ distributions for D^{*+} candidates. The D^0 is reconstructed using the modes: (a) $D^0 \to K^-\pi^+$, (b) $D^0 \to K^-\pi^+\pi^0$, and (c) $D^0 \to K^-\pi^+\pi^+\pi^-$.

Table A.2: Selection criteria for $B \to DX$ modes. "—" means no applicant of the selection for the decay mode.

B decay mode	D decay mode	$\|M_{K+n\pi} - M_D\|$	ΔM_{D^*}	R_2	$\cos\theta_{\rm th}$
$B^0 \to D^-\pi^+$	$D^+ \to K^-\pi^+\pi^+$	$< 2.5\sigma$	—	< 0.5	< 0.995
$B^0 \to D^{*-}\pi^+$	$D^0 \to K^-\pi^+$	$< 10\sigma$	< 5 MeV/c^2	—	—
	$D^0 \to K^-\pi^+\pi^0$	$< 3.5\sigma$	< 3 MeV/c^2	—	< 0.98
	$D^0 \to K^-\pi^+\pi^+\pi^-$	$< 4\sigma$	< 4 MeV/c^2	< 0.6	—
$B^0 \to D^{*-}\rho^+$	$D^0 \to K^-\pi^+$	$< 7\sigma$	< 4 MeV/c^2	< 0.6	< 0.95
	$D^0 \to K^-\pi^+\pi^0$	$< 3.5\sigma$	< 12 MeV/c^2	< 0.7	< 0.98
	$D^0 \to K^-\pi^+\pi^+\pi^-$	$< 3.5\sigma$	< 3 MeV/c^2	—	< 0.92
$B^+ \to \overline{D}^0\pi^+$	$D^0 \to K^-\pi^+$	$< 4\sigma$	—	—	—
	$D^0 \to K^-\pi^+\pi^0$	$< 3\sigma$		< 0.45	
	$D^0 \to K^-\pi^+\pi^+\pi^-$	$< 2\sigma$		< 0.45	

A.3.1 Hadronic B Decay Reconstruction

For the elimination of $e^+e^- \to q\overline{q}$ background, the selection criteria on the R_2 and cosine of thrust angle ($\cos\theta_{\rm thr}$) are applied for $B \to DX$ events. The selection criteria are listed in Table A.2. As for J/ψ inclusive modes, most reconstruction procedure are same as $B^0 \to J/\psi K_S$. In $B^0 \to J/\psi K^{*0}$ events, we reconstruct J/ψ by requiring two positively identified leptons, and the invariant mass ranges within $2.50 < M_{e^+e^-} < 3.15$ GeV/c^2 and within $3.05 < M_{\mu^+\mu^-} < 3.15$ GeV/c^2. The remaining difference is the species of the strange mesons for the reconstruction, K^+ or K^{*0} instead of K_S.
The B mesons reconstructed with all hadronic event in the list are finally selected by requiring ΔE and $M_{\rm bc}$ ranges. For $B \to D\pi$ modes and $B^+ \to J/\psi K^+$, the ranges are $5.27 < M_{\rm bc} < 5.29$ GeV/c^2 and $|\Delta E| < 40$ MeV, and for $B^0 \to J/\psi K^{*0}$, the ranges are $5.27 < M_{\rm bc} < 5.29$ GeV/c^2 and $|\Delta E| < 30$ MeV. Figures A.2 show the $M_{\rm bc}$ distributions for $B \to DX$ decays. The background contributions are estimated by fitting $M_{\rm bc}$ and ΔE distributions in the same way as $B^0 \to J/\psi K_S$ mode.

The decay vertices of $B_{\rm rec}$ is substituted by that of J/ψ for J/ψ inclusive modes as we do in $B^0 \to J/\psi K_S$ decay. In the $B \to DX$ modes, D^0 or D^+ decay vertex is reconstructed first, and then B decay vertex is determined using reconstructed "pseudo D track" and the primary track originated from B meson. The nominal B decay point constraint is applied during the vertex reconstruction for both modes. The $B_{\rm asc}$ decay vertices are reconstructed in exactly same way as $B_{\rm tag}$ vertex reconstruction.

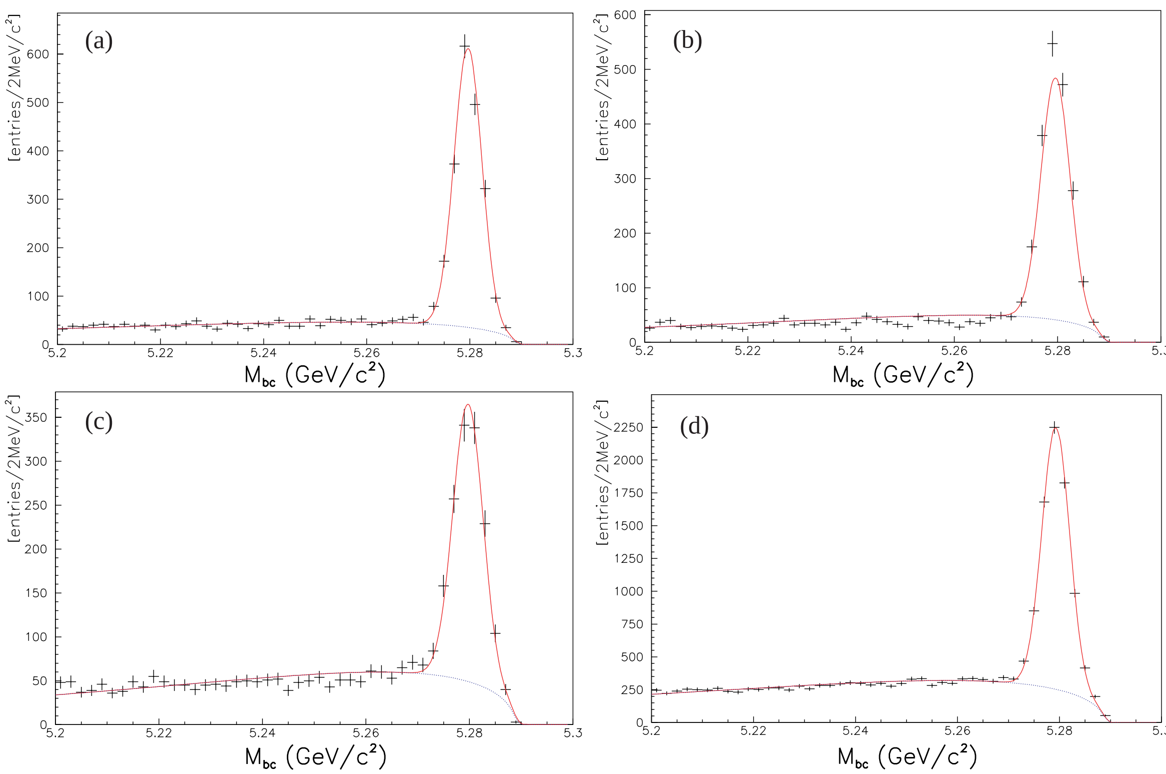


Figure A.2: $M_{\rm bc}$ distributions reconstructed via $B \to DX$ modes: (a) $B^0 \to D^- \pi^+$, (b) $B^0 \to D^{*-}\pi^+$, (c) $B^0 \to D^{*-}\rho^+$, and (d) $B^+ \to \overline{D}^0\pi^+$. The background function is indicated by the dotted curve in the figure.

A.3.2 Semileptonic B Decay Reconstruction

The $D^{*\mp}$ candidates reconstructed in Section A.2 are combined with $e^\pm$ or $\mu^\pm$ candidates that has opposite charge to $D^{*\mp}$. Lepton candidate should have the cms momentum between 1.4 GeV/c and 2.4 GeV/c. The semileptonic decay is not fully reconstructed due to neutrinos. We make two-dimensional scatter distribution for a missing mass-squared and a product of B and $D^* + \ell$ momenta, where the missing mass-squared is defined as $M^2_{\rm miss} = (E_B - E_{D^*\ell})^2 - |\vec{p_B}|^2 - |\vec{p_{D^*\ell}}|^2$ in the cms, and the momenta product is defined as $C \equiv 2\vec{p_B} \cdot \vec{p_{D^*\ell}}$. Drawing a triangle in the scatter plot, we select the B^0 candidate decaying into semileptonic mode from the inner area of the triangle. The triangle is enclosed by three lines; $C = (1.5/1.65)M^2_{\rm miss}$, $C = (1.2/1.3)M^2_{\rm miss}$, and $C = (0.3/2.95)(M^2_{\rm miss}+1.65)+1.5$, where the units are in GeV^2 and (GeV/c)^2. The signal fraction among them are estimated to be 79.4%, and the fraction of each background component is estimated to be 10.4% for fake D^*, 4.1% for wrong combination of $D^{*\mp}$ and $\ell^\pm$, 2.1% for continuum events, and 4.0% for $B^\pm$ decays. The decay vertex reconstruction of $B_{\rm rec}$ is done by the combination of "pseudo D^0 vertex" from K and π and the primary lepton from B^0 decay. The slow pion from D^{*-} decay is also used.

Appendix B

Estimation of Wrong Tagging Probability

In this chapter, we describe the estimation procedure of the wrong tagging probability.

Since we know the flavors of both B mesons for these decays, we can observe the time evolution of the neutral B-meson pair with the opposite flavor (OF) and the same flavor (SF) that are given by

$$\mathcal{P}_{\rm OF}(\Delta t) \;\propto\; 1 + (1-2w)\cos(\Delta M \Delta t), \tag{B.1}$$

$$\mathcal{P}_{\rm SF}(\Delta t) \;\propto\; 1 - (1-2w)\cos(\Delta M \Delta t). \tag{B.2}$$

The time dependent OF-SF asymmetry is

$$A_{\rm mix} \;\equiv\; \frac{\rm OF - SF}{\rm OF + SF} = (1-2w)\cos(\Delta M \Delta t). \tag{B.3}$$

We can obtain the w by measuring the amplitude of the OF-SF asymmetry. We perform an unbinned-maximum-likelihood fit of Δt distributions for both OF and SF onto the time evolution function, $\mathcal{P}_{\rm OF}$ and $\mathcal{P}_{\rm SF}$. The description of the unbinned-maximum-likelihood method is given in Appendix 4.1. In this study, the B^0-$\overline{B}^0$ mixing parameter, ΔM is fixed to the world average of $\Delta M = (0.472 \pm 0.017)$ ps^{-1} [49]. The reconstruction procedure of Δt is described in the next section. Figures B.1 and B.2 show the time dependent asymmetry for each of r region. Superimposed curves represent the asymmetry functions B.3. The estimation of wrong flavor tagging are performed individually with semileptonic B decay and hadronic B decays. The results are listed in Table B.1. In the list, we categorized the systematic uncertainty due to common systematic source for both semileptonic and hadronic modes, such as B meson lifetime and ΔM, into "correlated" uncertainty, and the uncertainty due to remaining systematic sources are categorized to "uncorrelated" uncertainty. Then we calculate correlated and uncorrelated systematic

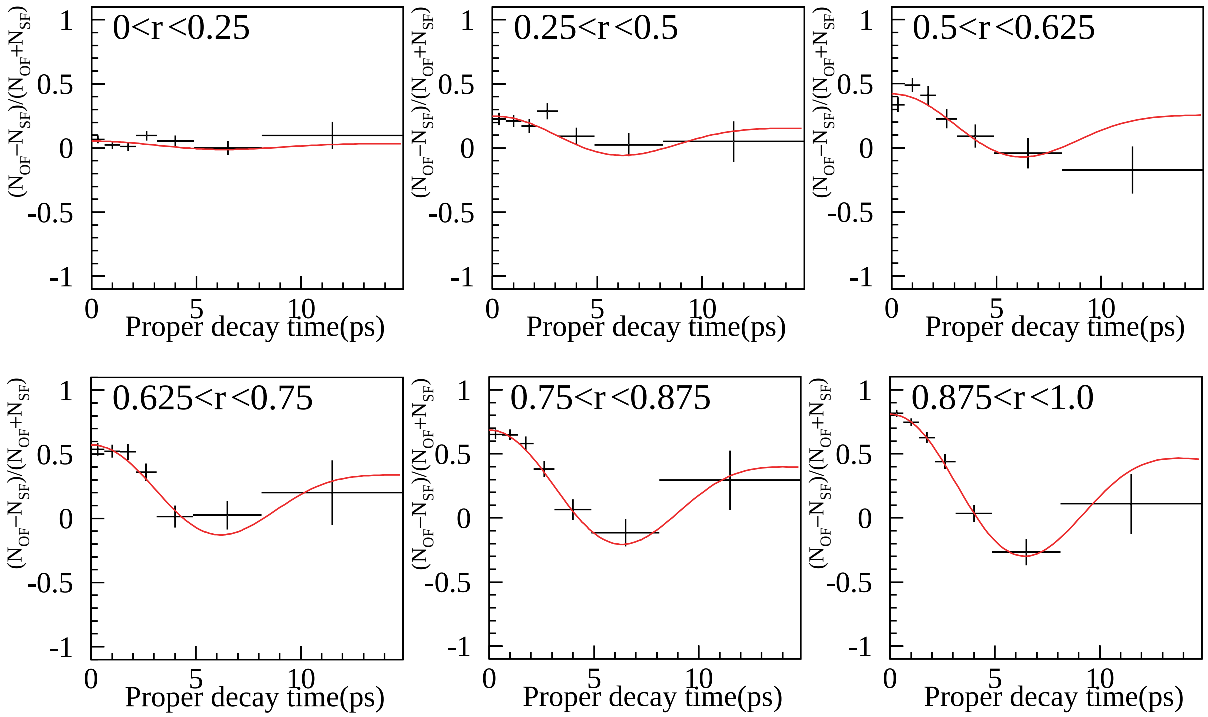


Figure B.1: Time dependent asymmetry as a function of proper-time difference for six regions of r for $B^0 \to D^{*-}\ell^+\nu$ decay. Fitted curves to the asymmetry function are superimposed.

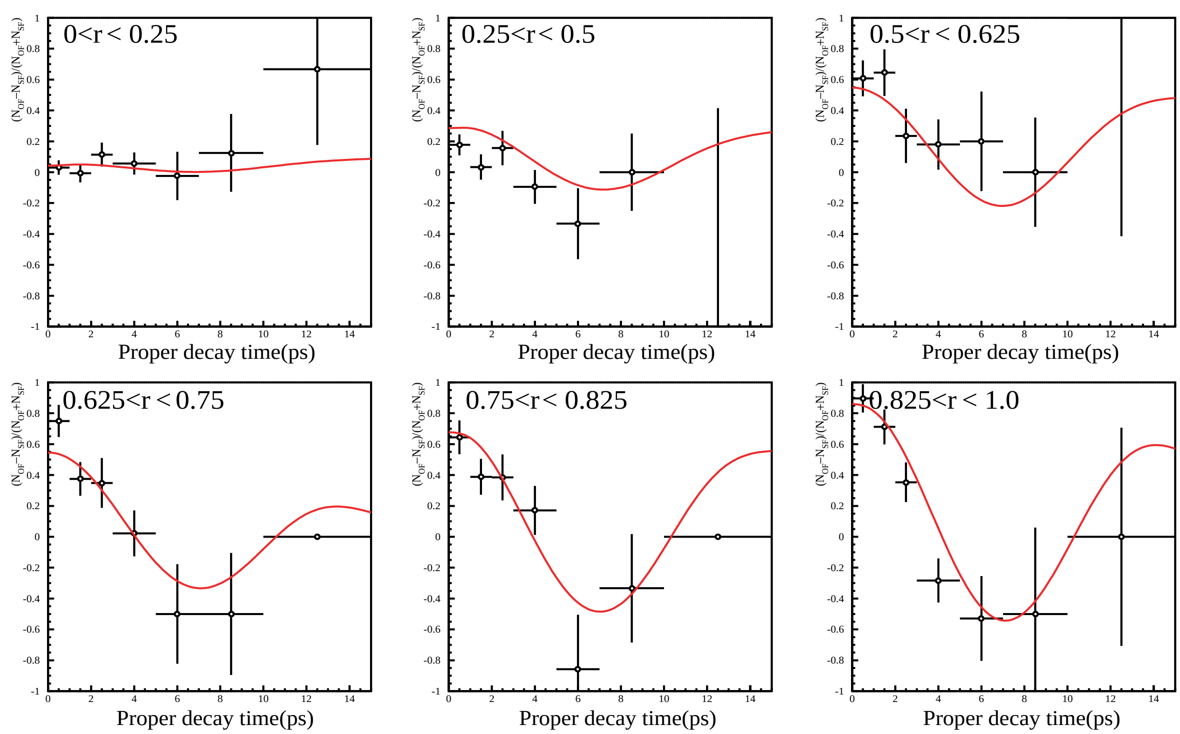


Figure B.2: Time dependent asymmetry as a function of proper-time difference for six regions of r for all combined hadronic decay modes. Fitted curves to the asymmetry function are superimposed.

Table B.1: Wrong tagging probabilities, w_ℓ, for each r-category. Values for semileptonic B decay and hadronic B decays are listed individually. “stat.” means statistical error, “uncor. syst.” means uncorrelated systematic uncertainty, and “corr. syst.” means correlated systematic uncertainty of w_ℓ.

w_ℓ obtained from $D^*\ell\nu$					
ℓ	r_ℓ	w_ℓ	stat.	uncorr. syst.	corr. syst.
1	$0.000 - 0.250$	0.463	± 0.010	± 0.004	± 0.000
2	$0.250 - 0.500$	0.351	± 0.016	${}^{-0.006}_{+0.010}$	± 0.001
3	$0.500 - 0.625$	0.254	± 0.019	${}^{-0.007}_{+0.008}$	± 0.003
4	$0.625 - 0.750$	0.169	± 0.017	± 0.007	± 0.003
5	$0.750 - 0.875$	0.107	${}^{-0.013}_{+0.014}$	± 0.006	${}^{-0.003}_{+0.004}$
6	$0.875 - 1.000$	0.041	${}^{-0.009}_{+0.010}$	± 0.005	± 0.004
w_ℓ obtained from hadronic modes					
ℓ	r_ℓ	w_ℓ	stat.	uncorr. syst.	corr. syst.
1	$0.000 - 0.250$	0.469	± 0.014	${}^{-0.007}_{+0.006}$	± 0.000
2	$0.250 - 0.500$	0.352	± 0.024	${}^{-0.012}_{+0.011}$	± 0.001
3	$0.500 - 0.625$	0.219	${}^{-0.027}_{+0.028}$	± 0.014	± 0.001
4	$0.625 - 0.750$	0.192	${}^{-0.024}_{+0.025}$	${}^{-0.014}_{+0.013}$	± 0.001
5	$0.750 - 0.875$	0.127	${}^{-0.024}_{+0.025}$	± 0.019	± 0.002
6	$0.875 - 1.000$	0.041	${}^{-0.017}_{+0.018}$	${}^{-0.016}_{+0.015}$	± 0.001

uncertainties by adding in quadrature, respectively. This procedure is necessary to combine individual result properly later. The systematic uncertainties for semileptonic mode are dominated by the uncertainties of background fractions and comparable with the statistical errors. For hadronic modes, the main contribution to the systematic uncertainties is uncertainty of fit bias obtained from Monte Carlo simulation, but the statistical errors dominate over the systematic uncertainty. To combine those results, we calculate the weighted average and its error, taking into account the correlated systematic uncertainties properly. The discrepancy of the estimated w_ℓ between semileptonic and hadronic B decays are also taken into account for the systematic uncertainties. The combined results are summarized in Table 3.2.

Appendix C

Computational Recipes of Probability Density Functions

In this chapter, we describe techniques for computation of the probability density functions.

C.1 Convolution

The probability density function for the extraction of $\sin 2\phi_1$ is basically represented by a linear combination of following convolution,

$$\frac{A}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \mathrm{d}t'\ G(t'-t;\ \mu,\sigma) \times \mathrm{e}^{-|t'|} \cdot (1+\sin 2\phi_1 \sin \gamma t'), \tag{C.1}$$

where t, μ, and σ are normalized by B meson lifetime and A denotes a constant. Since the signal probability density function consists of two Gaussians, it is represented by a sum of two convolutions above. As for the background, the background fractions are considered not to have asymmetry, the convolution is calculated with turning off the $\sin 2\phi_1$ term. We do not compute above convolution with naive trapezoidal integration but compute with analytic calculation technique for rapid computation. In this section, we show the recipes to compute above convolution [65].

First, we define following quantities

$$a \equiv \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \mathrm{d}t'\ G(t'-t;\ \mu,\sigma) \times \mathrm{e}^{-|t'|} \tag{C.2}$$

$$b \equiv \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \mathrm{d}t'\ G(t'-t;\ \mu,\sigma) \times \mathrm{e}^{-|t'|} \sin \gamma t'. \tag{C.3}$$

Convolution (C.1) is expressed using a and b as

$$A \cdot (a + \sin 2\phi_1 b). \tag{C.4}$$

C.1.1 Computation of $\boldsymbol{a}$

The computation of a is expressed as

$$a = \frac{1}{2}\mathrm{e}^{\frac{\sigma^2}{2}}\left[\ \mathrm{e}^{+\delta}\cdot\mathrm{erfc}\left(\frac{\sigma^2+\delta}{\sqrt{2}\sigma}\right) + \mathrm{e}^{-\delta}\cdot\mathrm{erfc}\left(\frac{\sigma^2-\delta}{\sqrt{2}\sigma}\right)\ \right], \tag{C.5}$$

where δ is $t-\mu$. $\mathrm{erfc}(x)$ is complementary error function, defined as

$$\mathrm{erfc}(x) \equiv \frac{2}{\sqrt{\pi}}\int_x^{\infty} \mathrm{d}t\ \mathrm{e}^{-t^2}. \tag{C.6}$$

C.1.2 Computation of $\boldsymbol{b}$

The computation of b is expressed as

$$b = \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{\delta^2}{2\sigma^2}}\left[\ I\left(\frac{1}{2\sigma^2},\ 1-\frac{\delta}{\sigma^2},\ \gamma\right) - I\left(\frac{1}{2\sigma^2},\ 1+\frac{\delta}{\sigma^2},\ \gamma\right)\ \right], \tag{C.7}$$

where I is defined as

$$I(k,\ell,m) = \begin{cases} \dfrac{\sqrt{\pi}}{2\sqrt{k}}\mathcal{I}m\left[w\left(\dfrac{m+\mathrm{i}\ell}{2\sqrt{k}}\right)\right] & \text{for } \ell \geq 0 \\ \dfrac{\sqrt{\pi}}{2\sqrt{k}}\mathcal{I}m\left[w\left(\dfrac{m+\mathrm{i}|\ell|}{2\sqrt{k}}\right) + 2\exp\left[\left(\dfrac{|\ell|+im}{2\sqrt{k}}\right)^2\right]\right] & \text{for } \ell < 0 \end{cases}. \tag{C.8}$$

w, called as Wofz function, in equation (C.8) is defined in complex plane as

$$w(z)\ \equiv\ \mathrm{e}^{-z^2}\mathrm{erfc}(-\mathrm{i}z). \tag{C.9}$$

The computation of Wofz function can be performed approximately with an accuracy better than $\mathcal{O}(10^{-10})$. The detail of the computation is described in [66].

C.2 Approximation

The calculation of

$$X \equiv \exp(\kappa)\ \mathrm{erfc}(\lambda) \tag{C.10}$$

is often seen in the convolution of probability density function and resolution function. The orders of κ and λ are $\kappa \sim \mathcal{O}(\delta t^2)$ and $\lambda \sim \mathcal{O}(\delta t)$, where δt is proper-time difference normalized by measurement error. $\exp(\delta t^2)$ diverges to $+\infty$ at $\delta t \to \infty$, while $\mathrm{erfc}(\delta t)$ converges to zero. The product also converges to zero. However, the computation of X by computers sometimes fails when $\delta t \gtrsim 26$ because the floating calculation of multiplying zero to $+\infty$ is performed. We can escape from this failure by an approximation [67]:

$$\begin{aligned} f(x) &\equiv \exp(x^2)\,\mathrm{erfc}(x) \\ &\simeq \frac{1}{\sqrt{\pi}}x^{-1} - \frac{1}{2\sqrt{\pi}}x^{-3}. \end{aligned} \tag{C.11}$$

Thus, X is approximated as

$$X \simeq \exp(\kappa - \lambda^2)\left(\frac{1}{\sqrt{\pi}}\lambda^{-1} - \frac{1}{2\sqrt{\pi}}\lambda^{-3}\right). \tag{C.12}$$

Appendix D

Feldman-Cousins Method

We describe the technique of the Feldman-Cousins confidence interval in this chapter. The Feldman-Cousins confidence interval is effective to discuss the significance of the observed physics result when the observed value is close to the theoretical boundary.

D.1 Confidence Belt

We consider a parameter α whose true value is fixed but unknown. The properties of our experimental apparatus are expressed in the function $\mathcal{F}(x;\alpha)$ that gives the probability of observing data x if the true value of the parameter is α. This function must be known in order to interpret the results of an experiment. For a large complex experiment, $\mathcal{F}$ is usually determined numerically using Monte Carlo simulation. $\mathcal{F}$ is naively a likelihood function of α.

For given $\mathcal{F}(x;\alpha)$, we can find for every value of α, two values $x_1(\alpha,\varepsilon)$ and $x_2(\alpha,\varepsilon)$ such that

$$P(x_1 < x < x_2 \mid \alpha) = 1 - \varepsilon = \int_{x_1}^{x_2} \mathcal{F}(x;\alpha)\,\mathrm{d}x. \tag{D.1}$$

This is shown graphically in Figure D.1 [49]: a horizontal line segment $[x_1(\alpha,\varepsilon), x_2(\alpha,\varepsilon)]$ is drawn for representative values of α. the union of all intervals $[x_1(\alpha,\varepsilon), x_2(\alpha,\varepsilon)]$, designated in the figure as the domain $D(\varepsilon)$, is known as "confidence belt". Typically the curves $x_1(\alpha,\varepsilon)$ and $x_2(\alpha,\varepsilon)$ are monotonic functions of α, which we assume for this discussion.

Upon performing an experiment to measure x and obtaining the value x_0, one draws a vertical line through x_0 on the horizontal axis. The confidence interval for α is the union 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. The confidence interval is an interval $[\alpha_1(x_0), \alpha_2(x_0)]$, where $\alpha_1(x_0)$ and $\alpha_2(x_0)$ are on the boundary of $D(\varepsilon)$. Thus, the boundaries of $D(\varepsilon)$ can be considered to be functions $x(\alpha)$ when constructing D, and then to be functions

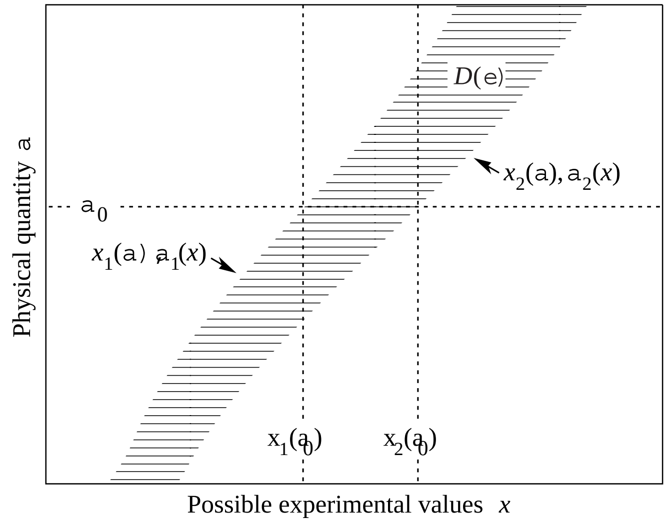


Figure D.1: Confidence intervals for a single unknown parameter, α.

$\alpha(x)$ when reading off confidence intervals. Such confidence intervals are said to have "confidence level" equal to $1-\varepsilon$.

Now suppose that some unknown particular value of α, say α_0 (indicated in the figure), is the true value of α. 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 $x_2(\alpha_0)$. Thus, we can write:

$$P[x_1(\alpha_0) < x < x_2(\alpha_0)] = 1-\varepsilon = P[\alpha_2(x) < \alpha_0 < \alpha_1(x)]. \tag{D.2}$$

And since, by construction, this is true for any value α_0, we can drop the subscript 0 and obtain:

$$P[\alpha_2(x) < \alpha < \alpha_1(x)] = 1-\varepsilon. \tag{D.3}$$

Suppose an experiment that determines a parameter α bounded from -1 to $+1$. We define a likelihood function as $\mathcal{L}(x;\alpha)$. According to the maximum-likelihood-method, $x=\alpha^*$ that gives the maximum $\mathcal{L}$ is considered to be "true" α. Based on the assumption that $\mathcal{L}(x;\alpha)$ is Gaussian, the standard deviation of $\mathcal{L}(x;\alpha)$ is considered the statistical error of the reconstructed α^*. For $1-\varepsilon = 0.68$, the "horizontal interval" can be defined as $[x_1(\alpha), x_2(\alpha)]$ where $x_1(\alpha) = \alpha - \sigma_\alpha$ and $x_2(\alpha) = \alpha + \sigma_\alpha$. σ_α is the standard deviation of $\mathcal{L}(x;\alpha)$. Figures D.2 show the distributions for $\mathcal{L}(x;\alpha=0.00)$ and $\mathcal{L}(x;\alpha=0.99)$. The naive domain $D(\varepsilon)$ can be constructed as Figure D.3. The discussion with the naive domain is not appropriate because we cannot determine the confidence interval, for example, when $x=+3$. In the Feldman-Cousins confidence method the construction of x_1 and x_2 is elaborated and for any observed x, the confidence interval can be defined.

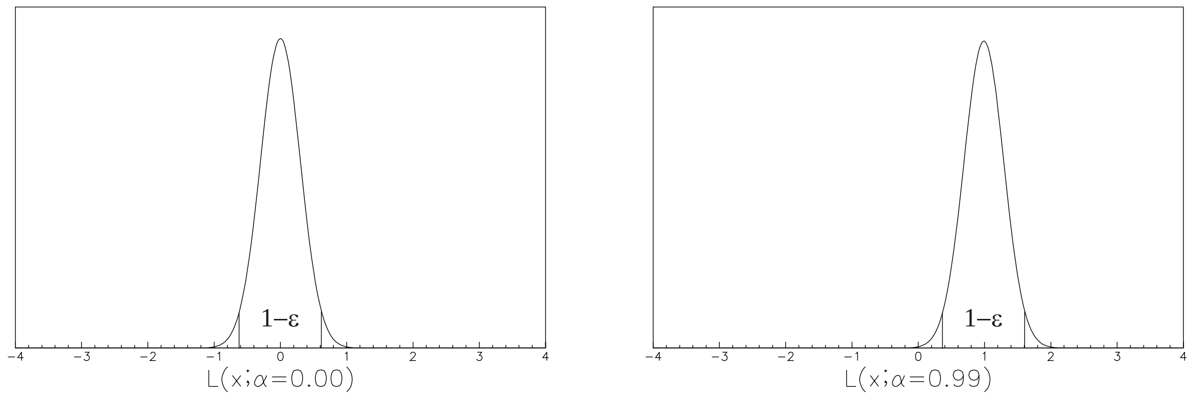


Figure D.2: Likelihood distributions for $\mathcal{L}(x;\alpha=0.00)$ (left) and $\mathcal{L}(x;\alpha=0.99)$ (right). Vertical lines indicate the "horizontal intervals" as a function of α. Vertical ticks are omitted because they are meaningless.

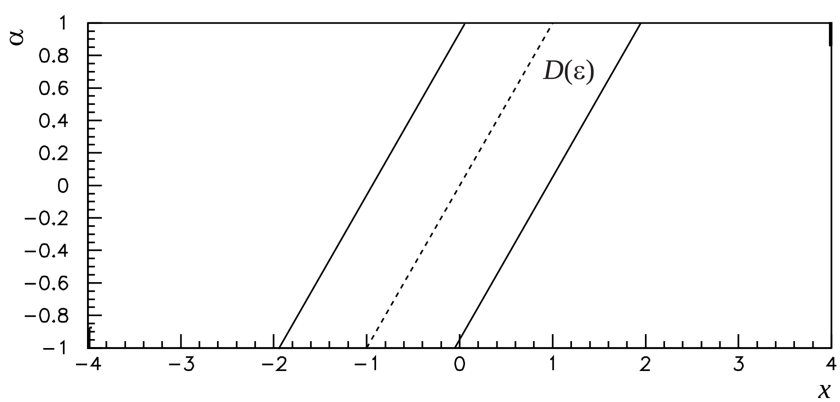


Figure D.3: Confidence belt constructed naively. The Gaussian distribution is assumed for the likelihood function.

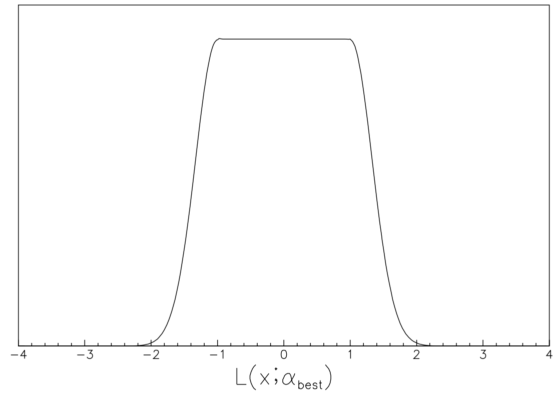


Figure D.4: Distribution for $\mathcal{L}(x;\alpha_{\mathrm{best}})$.

D.2 Feldman-Cousins Confidence Interval

The advantage of the Feldman-Cousins method is the capability to determine the confidence interval for any observed x. The construction of $\mathcal{F}(x;\alpha)$ by the Feldman-Cousins method is described in this section. Consider that how can be maximize $\mathcal{L}(x;\alpha_{\mathrm{best}})$ for a particular x. α_{best} is allowed to vary from -1 to $+1$. When x lies between -1 and $+1$, $\mathcal{L}$ is maximized by $\alpha_{\mathrm{best}} = x$. When $x > 1$, the maximum $\mathcal{L}$ is given by $\alpha_{\mathrm{best}} = 1$, and when $x < -1$, the maximum is given by $\alpha_{\mathrm{best}} = -1$. $\mathcal{L}(x;\alpha_{\mathrm{best}})$ corresponds to relative possibility of measuring x. The relative possibility to measure $-1 \le x \le +1$ is same, however, the possibility to measure $x > +1$ or $x < -1$ is smaller than that to obtain $-1 \le x \le +1$ because of the theoretical boundary of α. The degradation from the boundary edge is expected to be Gaussian shape with the standard deviation of measurement error of x, namely a quadratic sum of statistical and systematic uncertainties. Figure D.4 shows the distribution of $\mathcal{L}(x;\alpha_{\mathrm{best}})$.

According to the probability ordering principle, we take into account the relative possibility, $\mathcal{L}(x;\alpha_{\mathrm{best}})$, to the likelihood calculation:

$$\mathcal{R}(x;\alpha) = \frac{\mathcal{L}(x;\alpha)}{\mathcal{L}(x;\alpha_{\mathrm{best}})}. \tag{D.4}$$

Figures D.5 show the modified likelihood distributions for $\mathcal{R}(x;\alpha = 0.00)$ and $\mathcal{R}(x;\alpha = 0.99)$. x_1 and x_2 for the confidence level of $(1-\varepsilon)$ are defined as

$$1 - \varepsilon = \frac{\displaystyle\int_{x_1}^{x_2} \mathrm{d}x\ \mathcal{R}(x;\alpha)}{\displaystyle\int_{-\infty}^{\infty} \mathrm{d}x\ \mathcal{R}(x;\alpha)}, \qquad \mathcal{R}(x_1;\alpha) = \mathcal{R}(x_2;\alpha). \tag{D.5}$$

Figure D.6 shows the confidence belt obtained by the Feldman-Cousins method. For any measured x, we can define the confidence interval.

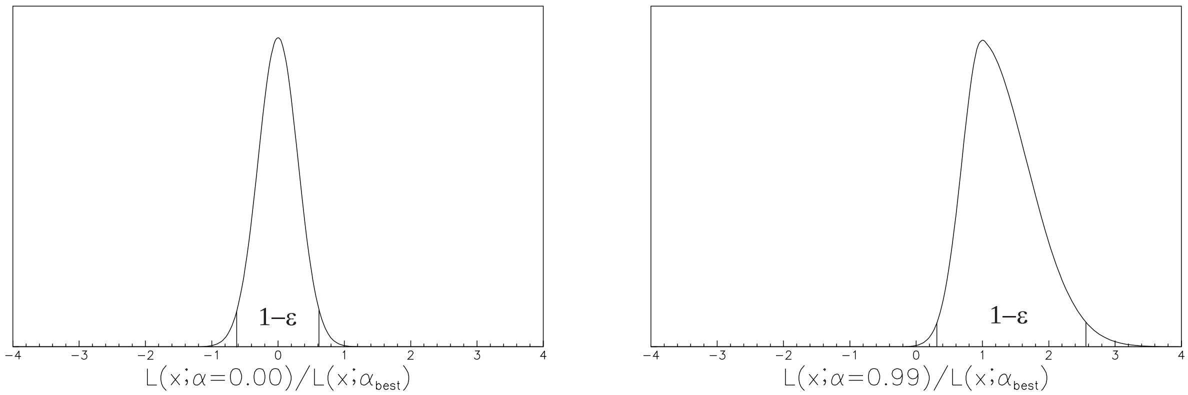


Figure D.5: Distributions for $\mathcal{R}(x;\alpha=0.00)$ (left) and $\mathcal{R}(x;\alpha=0.99)$ (right).

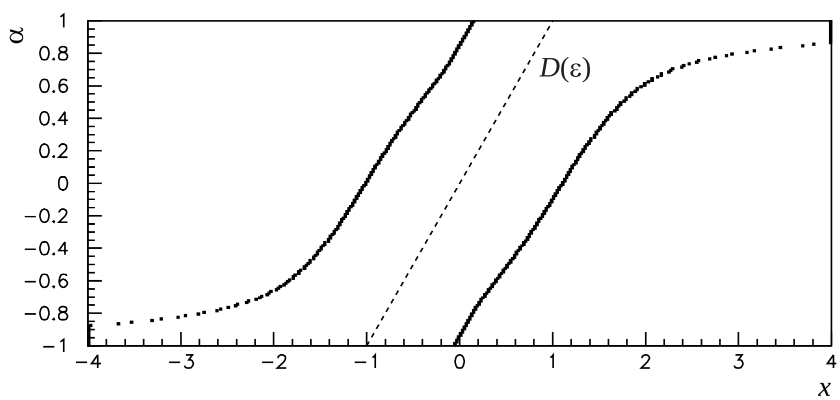


Figure D.6: Confidence belt constructed with the Feldman-Cousins method.

Appendix E

Lifetime Measurement of B Mesons with Hadronic Decay Modes

In this chapter, we present lifetime measurement of B^0 and B^+ mesons using 29.1 fb^{-1} data sample. The lifetime measurement provides key input to determine the Kobayashi-Maskawa matrix element, $|V_{cb}|$ [2]. Moreover, the ratio of the lifetimes is sensitive to effects beyond the spectator model, such as Pauli interference and W-exchange. In the framework of the heavy flavor expansion, the lifetime ratio, $r_\tau \equiv \tau_{B^+}/\tau_{B^0}$ is predicted to be equal one, up to small corrections proportional to $1/m_b^3$ [68]. A recent theoretical study predicts $r_\tau = 1.07 \pm 0.03$ [69], whereas this ratio is measured to be $r_\tau = 1.073 \pm 0.027$ on average [49][70]. Accurate determinations of B^0 and B^+ lifetimes also provide essential input to analyses of CP violation and neutral B mixing.

We measure B meson lifetimes from the distance of two decay vertices of the B mesons, as we measured Δz in the study of CP violation. In the Belle experiment, the lifetime measurement is quite sensitive to the representation of resolution function unlike $\sin 2\phi_1$ measurement, because the order of the detector resolution is similar to the lifetimes of the B mesons. Therefore, the main efforts to obtain the lifetimes is to improve the representation of the resolution function. After the lifetime measurement, we consequently have the elaborate resolution function form. This will be fed quite usefully to other analysis measuring Δz distribution.

The focused decay modes to extract B meson lifetimes are hadronic B decays: $B^0 \to J/\psi K_S$, $B^0 \to J/\psi K^{*0}(K^+\pi^-)$, $B^0 \to D^-\pi^+$, $B^0 \to D^{*-}\pi^+$, $B^0 \to D^{*-}\rho^+$, $B^+ \to J/\psi K^+$, and $B^+ \to \overline{D}^0\pi^+$. The reconstruction of $B^0 \to J/\psi K_S$ is described in Section 3, and the reconstruction of the other modes are given in Section A, where first five modes are used for B^0 lifetime measurement and the remaining two modes are used for B^+ lifetime measurement. The reconstruction of B mesons via those decay modes are described in Section 3.3 and in Appendix A. The lifetime is reconstructed with using proper-time difference of two B mesons as decay products of $\Upsilon(4S)$, in the similar method as described in Section 3.5 and in Appendix A.

In the following sections, we discuss the proper-time difference reconstruction, the probability density function especially concerning the improvement of the resolution function, the results of lifetime fitting, and the systematic uncertainties, in this order.

E.1 Proper-Time Difference Reconstruction

As we described already, the proper-time difference is reconstructed as

$$\Delta t \simeq \frac{\Delta z}{(\beta\gamma)_\Upsilon}, \qquad \Delta z \equiv z_{\rm rec} - z_{\rm asc}, \tag{E.1}$$

where $z_{\rm rec}$ and $z_{\rm asc}$ are decay vertices of $B_{\rm rec}$ and $B_{\rm asc}$, respectively. The procedure of the proper-time difference reconstruction is almost same as the method described in Section 3.5 and in Appendix A. The difference is as follows. In the lifetime reconstruction, we turn off the impact parameter selection of $|\delta z| < 1800$ μm, because this selection makes apparent shoulder in the Δz distribution at $|\Delta z| = 1800$ μm and it is difficult to take into account this shoulder to the probability density function of Δz properly. In the $\sin 2\phi_1$ analysis, due to the small statistics, such effect is expected to be quite small, and it is rather effective in the suppression of the Δz tail from the miss-reconstruction of the vertices.

We introduce new definition of a goodness of the vertex reconstruction. Since we apply the B decay point constraint in the r-ϕ plane in the vertex reconstruction, nominal χ^2 of the vertex fit strongly correlates with the distance from the interaction point to the reconstructed vertex. Figure E.1 (a) shows the nominal χ^2 as a function of a distance from the interaction point to the reconstructed vertex. We define the goodness of the fit as a reduced χ^2 projected onto beam direction:

$$\xi \equiv \frac{1}{2n} \sum_i^n \left(\frac{z^i_{\rm bfr} - z^i_{\rm aft}}{\varepsilon^i_{\rm trk}} \right)^2, \tag{E.2}$$

where n is a number of tracks $z^i_{\rm bfr}$ is a z position of input track, $z^i_{\rm aft}$ is z position of fitted track, and $\varepsilon^i_{\rm trk}$ is an estimated error of track position measurement in the z direction. Figure E.1 (b) shows the distribution of ξ as a function of the distance. The newly defined goodness of fit does not depend on the distance. Figure E.2 is the distribution of ξ of the vertex reconstruction with more than one track. We require the ξ of the reconstructed vertices less than 100 to eliminate badly reconstructed vertices for both $z_{\rm rec}$ and $z_{\rm asc}$.

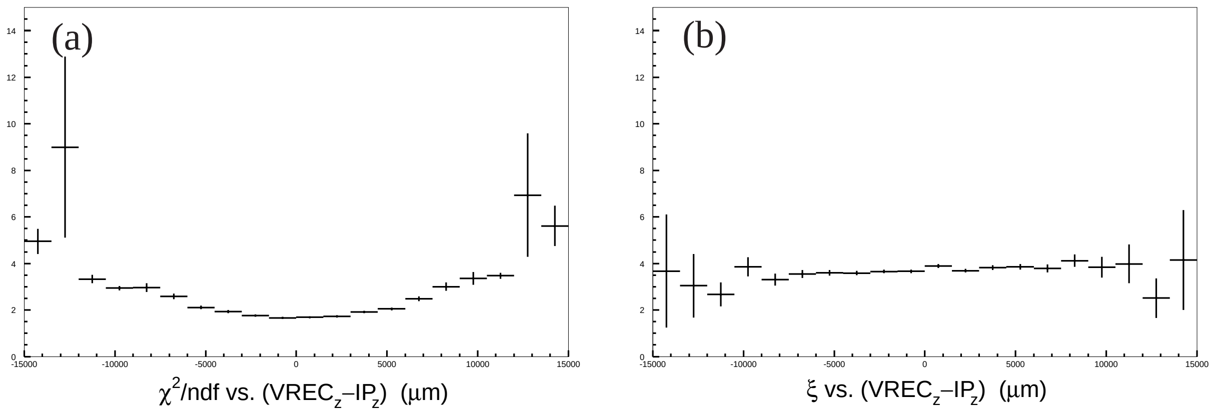


Figure E.1: (a) average of nominal χ^2/n.d.f., and (b) newly defined goodness of fit, ξ, in each bin of the distance form interaction point to the reconstructed vertex. Nominal χ^2/n.d.f. shows a dependence on the distance, while ξ shows no dependence. The histograms are obtained from the reconstruction of $z_{\rm rec}$ with $B^0 \to J/\psi K_S$ Monte Carlo samples.

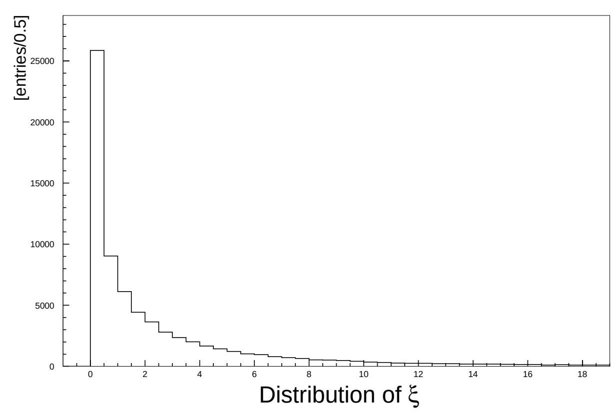


Figure E.2: Distribution of ξ of the vertex reconstruction that is obtained from the reconstruction of $z_{\rm rec}$ with $B^0 \to J/\psi K_S$ Monte Carlo samples, with using multiple tracks.

For vertices reconstructed with only singe track, we cannot define ξ, because the vertex reconstruction with single track is equivalent to obtain a crossing point between the track and the distribution of the B decay point. Thus, we do not discuss the ξ for single track vertices, hereafter.

We also reject the poorly reconstructed vertices by a selection of $|\Delta t| < 70$ ps. The selection is greater than 40 times of the B meson lifetimes.

E.2 Probability Density Function

In this section, we construct the probability density function of the Δt distribution. The probability density function is expressed as

$$\mathcal{P}(\Delta t;\ \tau_B) \quad = \quad f_{\rm sig}\cdot\mathcal{P}_{\rm sig}(\Delta t;\ \tau_B) + (1-f_{\rm sig})\cdot\mathcal{P}_{\rm bkg}(\Delta t) \tag{E.3}$$

where $f_{\rm sig}$ is signal probability of each event, $\mathcal{P}_{\rm sig}$ is a signal probability density function convoluted with resolution function, $\mathcal{P}_{\rm bkg}$ is a background probability density function, and $\mathcal{P}_{\rm ol}$ is a probability density function for outlier component.

The signal probability, $f_{\rm sig}$, is calculated in the same way as we did in the CP violation study. The signal component is represented by two-dimensional single Gaussian in ΔE and $M_{\rm bc}$ plane, background component in ΔE is represented by first order polynomial, and that in $M_{\rm bc}$ is represented by ARGUS background function. The detail discussion for $f_{\rm sig}$ is given in 4.2.1 and we omit the discussion of $f_{\rm sig}$ in this Appendix.

The signal probability density function is

$$f(\Delta t;\ \tau_B) = \frac{1}{2\tau_B}\exp\left(-\frac{|\Delta t|}{\tau_B}\right). \tag{E.4}$$

Using above equation, $\mathcal{P}_{\rm sig}$ is represented as

$$\mathcal{P}_{\rm sig}(\Delta t;\ \tau_B) = \int_{-\infty}^{\infty} {\rm d}(\Delta t')\ \mathcal{R}(\Delta t - \Delta t')\cdot f(\Delta t';\ \tau_B), \tag{E.5}$$

where $\mathcal{R}(x)$ is a resolution function. The detail of the resolution function is described in subsection E.2.1.

We construct elaborate $\mathcal{P}_{\rm bkg}$ function form than the CP violation study. This is described in subsection E.2.2.

To extract the B meson lifetimes we use the unbinned-maximum-likelihood method by maximizing the likelihood $\mathcal{L}$ defined as

$$\mathcal{L}(\tau_{B^0}, \tau_{B^+}) \equiv \mathcal{L}_{\tau_{B^0}} \cdot \mathcal{L}_{\tau_{B^+}}, \tag{E.6}$$

where $\mathcal{L}_{\tau_{B^0}}$ and $\mathcal{L}_{\tau_{B^+}}$ are

$$\begin{aligned} \mathcal{L}_{\tau_{B^0}} &\equiv \mathcal{L}_{J/\psi K_S} \cdot \mathcal{L}_{J/\psi K^{*0}} \cdot \mathcal{L}_{D^-\pi^+} \cdot \mathcal{L}_{D^{*-}\pi^+} \cdot \mathcal{L}_{D^{*-}\rho^+} = \prod_i \mathcal{P}(\Delta t_i;\ \tau_{B^0}), \\ \mathcal{L}_{\tau_{B^+}} &\equiv \mathcal{L}_{J/\psi K^+} \cdot \mathcal{L}_{\overline{D}^0\pi^+} = \prod_i \mathcal{P}(\Delta t_i;\ \tau_{B^+}), \end{aligned} \tag{E.7}$$

respectively. As equation (E.6) implies, we reconstruct τ_{B^0} and τ_{B^+} in simultaneous fit. In addition to B meson lifetimes, we also determine parameters for resolution function and $\mathcal{P}_{\rm ol}$ in the same fit. Detail of fitted parameters is described in Section E.3.

E.2.1 Resolution Function

As we described in Section 4.2.2, the resolution function is represented as a convolution of three contributions: detector resolution, smearing due to non-primary tracks, and kinematic approximation. In addition, we introduce outlier component for the probability density function so that the fitted lifetimes become independent of the selection of $|\Delta t|$ range. In the following paragraphs, we describe the improvement of the each component, and the outlier.

Detector Resolution

In the study of detector resolution, all tracks in the associated B side is set to be "pseudo" primary tracks by artificially shifting the generation vertices of the tracks toward the decay vertex of associated B that is their ancestor with using generator information. This procedure is performed to distinguish the detector resolution from the smearing effect due to non-primary tracks. The smearing effect from non-primary tracks is discussed in the next paragraph.

In the study of CP violation we represented the resolution function by a sum of two Gaussians defined in event-by-event. We examined the vertex resolution in detail and found that: (i) for the multiple track vertices, the resolution can be represented by a single Gaussian defined in event-by-event using the estimated error of the vertex reconstruction and its goodness of the fit, and (ii) for the single track vertices, the resolution function can be represented by a sum of double Gaussians defined in event-by-event.

First, we discuss the vertices obtained by multiple tracks. Figures E.3 show the pull distributions that is defined as the residual divided by the estimated error. We divide the distribution into seven according to the goodness of the fit, ξ. The fitted curves to single Gaussian for each distribution are superimposed. The distributions are obtained by $B^0 \to J/\psi K_S$ signal Monte Carlo. Figures E.4 show the standard deviation of the pull distribution as a function of ξ. We can see clear linear dependence of the standard deviation on the ξ. Based on the success of these fits, we represent the detector resolution of $z_{\rm rec}$ and $z_{\rm asc}$ obtained with multiple tracks by following function:

$$R_{\rm rec}^{\rm multiple}(\delta z_{\rm rec}) = G(\delta z_{\rm rec}; s_{\rm rec}\ \sigma_{\rm rec}), \tag{E.8}$$

$$R_{\rm asc}^{\rm multiple}(\delta z_{\rm asc}) = G(\delta z_{\rm asc}; s_{\rm asc}\ \sigma_{\rm asc}), \tag{E.9}$$

where $s_{\rm rec}$ and $s_{\rm asc}$ is a first order polynomial of ξ to represent the linear dependence:

$$s_{\rm rec} \equiv s_{\rm rec}^0 + s_{\rm rec}^1 \times \xi_{\rm rec}, \tag{E.10}$$

$$s_{\rm asc} \equiv s_{\rm asc}^0 + s_{\rm asc}^1 \times \xi_{\rm asc}, \tag{E.11}$$

and $\sigma_{\rm rec}$ and $\sigma_{\rm asc}$ are vertex reconstruction errors for $B_{\rm rec}$ and $B_{\rm asc}$, respectively. In above equations, $\xi_{\rm rec}$ and $\xi_{\rm asc}$ are the goodness of the vertex fit defined as equation (E.2) for $B_{\rm rec}$ and $B_{\rm asc}$, respectively. Figures E.5 show the residual distribution of (a) $z_{\rm rec}$ and (b) $z_{\rm asc}$. The fitted curves to $R_{\rm rec}^{\rm multiple}(\delta z_{\rm rec})$ and $R_{\rm rec}^{\rm multiple}(\delta z_{\rm asc})$ are also superimposed. The convolution of two resolution function, $R_{\rm rec}^{\rm multiple}(\delta z_{\rm rec})$ and $R_{\rm asc}^{\rm multiple}(\delta z_{\rm asc})$ generates the detector resolution, $R_{\rm det}$ for multiple tracks vertices. The convoluted curve is shown in Figure E.5 (c), with the residual distribution of Δz. The resolution function represented by a sum of double Gaussians in the analysis of $\sin 2\phi_1$ extraction corresponds to the superimposition of infinite number of single Gaussians whose standard deviation varies depending on the goodness of vertex fit linearly.

For vertices reconstructed with single track, we cannot introduce any ξ dependence, because the goodness of the fit cannot be defined for those vertices. Therefore, we use a sum of two Gaussians whose standard deviations varies according to the vertex reconstruction error, as we did in Section 4.2.2. Since vertex reconstruction situation is expected to be exactly same for $z_{\rm rec}$ and $z_{\rm asc}$ in the single track case, we use same shape for both resolutions of $z_{\rm rec}$ and $z_{\rm asc}$. The resolution of single track vertices, $R_{\rm rec}^{\rm single}(\delta z_{\rm rec}) = R_{\rm asc}^{\rm single}(\delta z_{\rm asc})$, is defined as follows:

$$R_{\rm rec}^{\rm single}(\delta z_{\rm rec}) = (1 - f_{\rm tail})\ G(\delta z_{\rm rec}; s_{\rm main}\sigma_{\rm rec}) + f_{\rm tail}\ G(\delta z_{\rm rec}; s_{\rm tail}\sigma_{\rm rec}), \tag{E.12}$$

$$R_{\rm asc}^{\rm single}(\delta z_{\rm asc}) = (1 - f_{\rm tail})\ G(\delta z_{\rm asc}; s_{\rm main}\sigma_{\rm asc}) + f_{\rm tail}\ G(\delta z_{\rm asc}; s_{\rm tail}\sigma_{\rm asc}). \tag{E.13}$$

Figures E.6 show the residual distributions of $z_{\rm rec}$ and $z_{\rm asc}$ reconstructed with single track. Fits to $R_{\rm rec}^{\rm single}(\delta z_{\rm rec})$ and $R_{\rm asc}^{\rm single}(\delta z_{\rm rec})$ are also superimposed.

To summarize the detector resolution, we use an event-by-event single Gaussian with ξ dependence for multiple track vertices, and an event-by-event sum of two Gaussians for single track vertices.

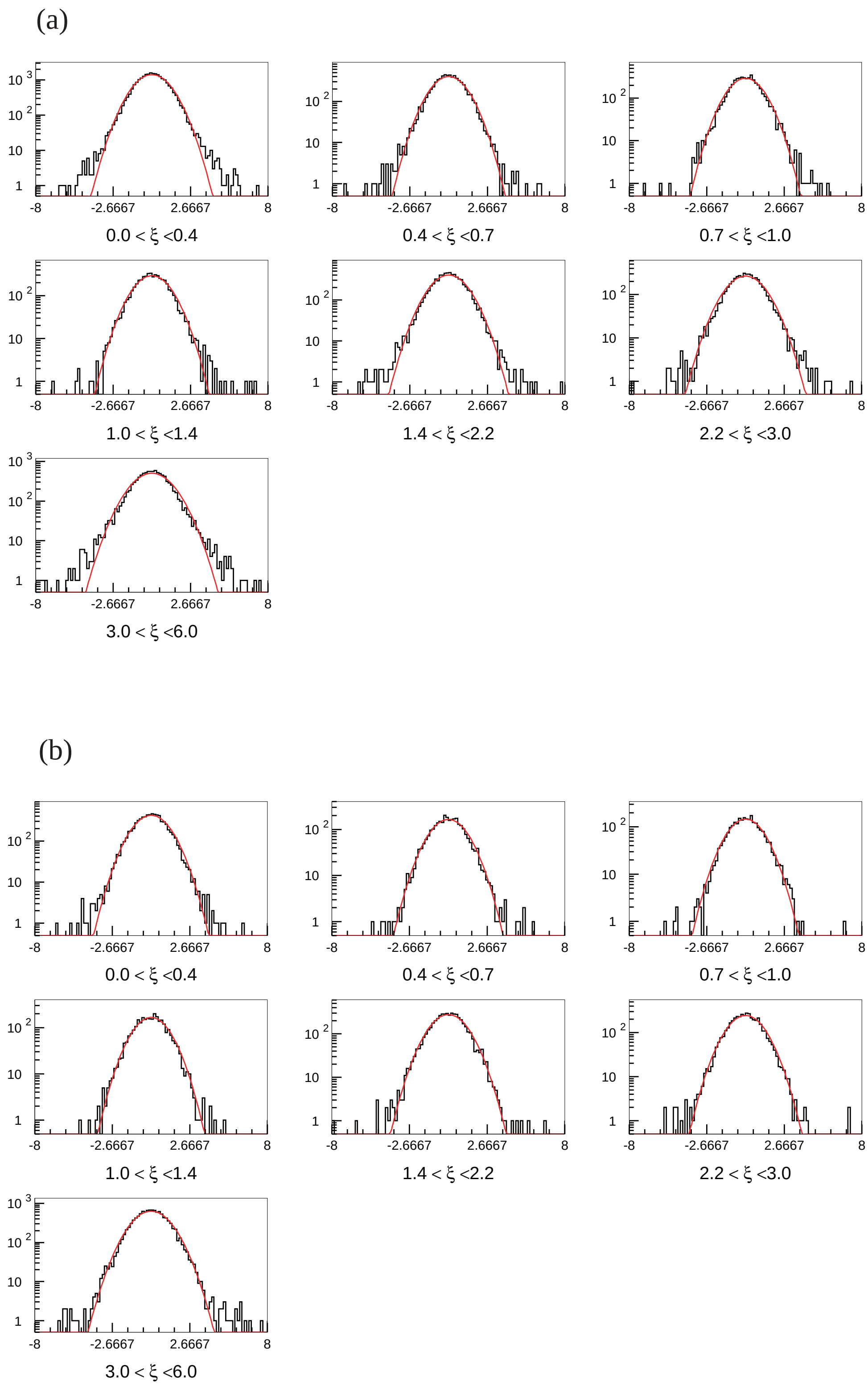


Figure E.3: Pull distributions of (a) reconstructed B side, and (b) associated B side, divided by the associated goodness of vertex fit. The vertex reconstruction is performed with using multiple tracks. The fitted curves to single Gaussian are superimposed. The distributions are obtained from $B^0 \to J/\psi K_S$ Monte Carlo samples.

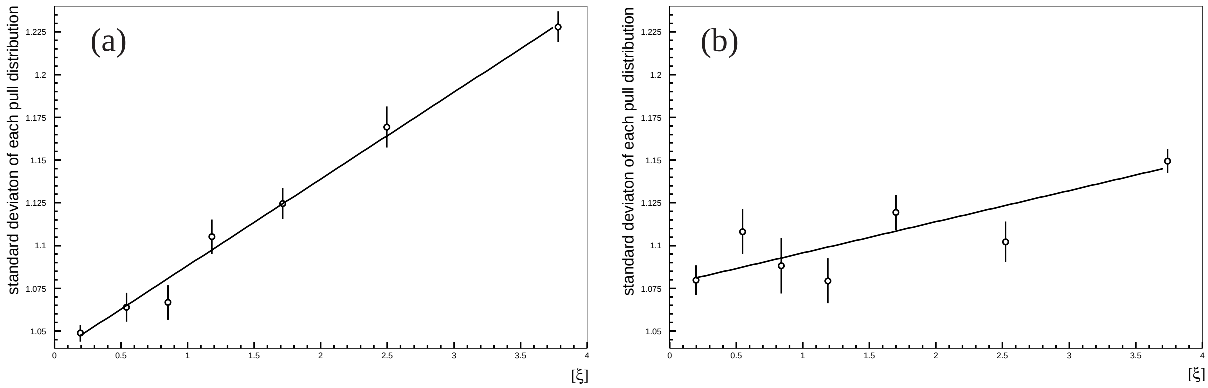


Figure E.4: Standard deviation of the pull distributions fitted to single Gaussian, as a function of ξ. (a) shows the dependence for reconstructed B vertex, and (b) shows that for associated B vertex. The distributions are obtained form $B^0 \to J/\psi K_S$ Monte Carlo samples.

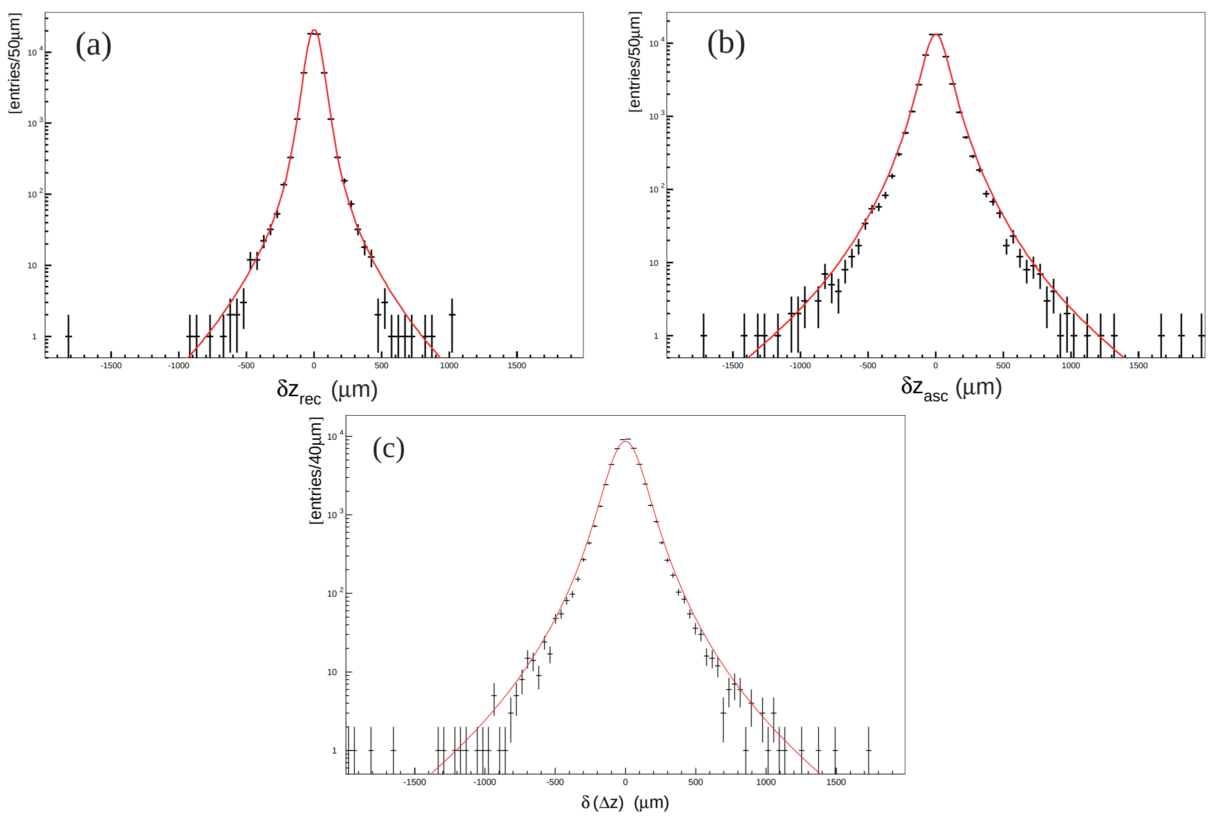


Figure E.5: Residual distribution of (a) $z_{\rm rec}$ and (b) $z_{\rm asc}$ with fitted curve to event by event single Gaussian. Figure (c) is the residual distribution of $\delta\Delta z$, and the superimposed curve in the figure is the convolution of two resolution $R^{\rm multiple}_{\rm rec}(\delta z_{\rm rec})$ and $R^{\rm multiple}_{\rm asc}(\delta z_{\rm asc})$

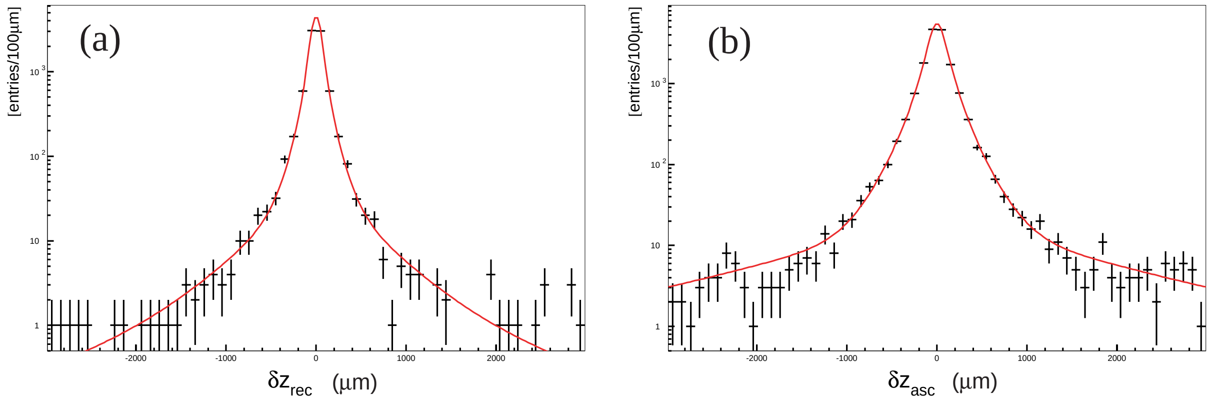


Figure E.6: Residual distribution of (a) $z_{\rm rec}$ and (b) $z_{\rm asc}$ reconstructed with single track. The superimposed curves are fits to a sum of event by event two Gaussian.

Smearing Due to Non-Primary Tracks

The effect of non-primary tracks, mainly due to the decay products of charmed mesons, on the resolution of the associated B meson vertex is considered in this paragraph. In the following discussion, we turn off the artificial shift for tracks to pretend "pseudo" primary tracks. Although the requirement for the impact parameter being $|\delta r| < 500\ \mu$m and the iteration comprised in the vertex reconstruction algorithm of $B_{\rm asc}$ rejects non-primary tracks, remaining non-primary tracks smear the resolution function and its contribution will be non-Gaussian.

We define two decay vertices of $B_{\rm asc}$: $z_{\rm asc}^{\rm nc}$ and $z_{\rm asc}^{\rm np}$. The $z_{\rm asc}^{\rm nc}$ is defined as a vertex that is reconstructed with tracks that are artificially set to be "pseudo" primary tracks. It can be considered that $z_{\rm asc}^{\rm nc}$ is not smeared by non-primary tracks. The $z_{\rm asc}^{\rm np}$ is defined as a vertex that is reconstructed with the tracks as they are. This vertex is smeared by the effect due to non-primary tracks together with the detector resolution. The distribution of $(z_{\rm asc}^{\rm np} - z_{\rm asc}^{\rm nc})$ is shown in Figure E.7. The events with $z_{\rm asc}^{\rm np} = z_{\rm asc}^{\rm nc}$ are removed from the histogram. The histogram could be considered the representation of the smearing effect due to the non-primary tracks, if the long tail distribution could be negligible. We use $R_{\rm np}$ for the smearing of the non-primary tracks hereafter. In reality, since the detector resolution cannot be totally deconvoluted from the smearing, it is slightly broader than true $R_{\rm np}$. Therefore, the convolution of the distribution shown in Figure E.7 with $R_{\rm asc}$ generates overestimated resolution than expected resolution. Instead of this, we can utilize the distribution to understand the structure of $R_{\rm np}$. We can see lifetime structure in the distribution, which is natural because $R_{\rm np}$ is mainly originated by the finite lifetimes

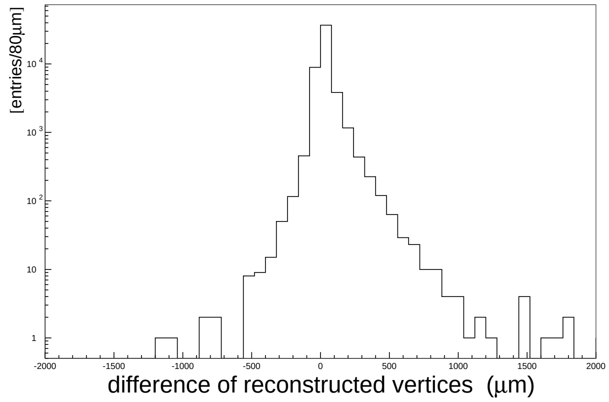


Figure E.7: Distribution of $(z^{\rm np}_{\rm asc} - z^{\rm nc}_{\rm asc})$ is shown. Events that match $z^{\rm np}_{\rm asc} = z^{\rm nc}_{\rm asc}$ are removed, for they will be involved in the Dirac's δ-function in equation (E.14). The histogram is obtained from $B^0 \to J/\psi K_S$ signal Monte Carlo, whose $B_{\rm asc}$ vertex is reconstructed with multiple tracks.

of the charm mesons. We determine $R_{\rm np}$ as follows:

$$\begin{aligned} R_{\rm np}(\delta z_{\rm asc}) \;\equiv\; & f_\delta \times \delta(\delta z_{\rm asc}) \\ & + (1-f_\delta) \times \Big[\, f_p \; E_p(\delta z_{\rm asc}; \tau^p_{\rm np}) \;+\; (1-f_p)\; E_n(\delta z_{\rm asc}; \tau^n_{\rm np}) \,\Big], \end{aligned} \tag{E.14}$$

where $\delta(x)$ is Dirac's δ-function, and E_p and E_n are defined as

$$E_p(x;\tau) \;\equiv\; \frac{1}{\tau}\exp\left(-\frac{x}{\tau}\right) \qquad \text{for } x > 0, \quad \text{otherwise } 0, \tag{E.15}$$

$$E_n(x;\tau) \;\equiv\; \frac{1}{\tau}\exp\left(+\frac{x}{\tau}\right) \qquad \text{for } x \le 0, \quad \text{otherwise } 0. \tag{E.16}$$

In the following paragraphs, we discuss the $\tau^p_{\rm np}$ and $\tau^n_{\rm np}$ in equation (E.14). Since it is impossible to deconvolute $R_{\rm np}$ from the convolution totally, as described already, we always study the $R_{\rm np}$ together with $R_{\rm asc}$ by convolution of them. The convoluted function is expressed by $R_{\rm asc} \otimes R_{\rm np}$.

As well as the detector resolution, we discuss the $R_{\rm np}$ separately for two cases: that the vertices reconstructed multiple tracks, and that the vertices reconstructed with single track. First, we discuss the case of the multiple tracks. Figure E.8 shows the residual distribution of $z_{\rm asc}$ with superimposed curve fitted to $R_{\rm asc} \otimes R_{\rm np}$. The lifetimes, $\tau^p_{\rm np}$ and $\tau^n_{\rm np}$, in equation (E.14) are constants. The detector resolution is fixed to the shape that is determined from the fit in the previous paragraph with using signal Monte Carlo. The obtained resolution does not describe tail region efficiently. Figures E.9 represent the correlation (a) between $\sigma_{\rm asc}$ and $(z^{\rm np}_{\rm rec} - z^{\rm nc}_{\rm rec})$, and (b) between $\xi_{\rm asc}$ and $(z^{\rm np}_{\rm rec} - z^{\rm nc}_{\rm rec})$. The vertices for the distributions are required to be reconstructed with multiple tracks. Events with $z^{\rm np}_{\rm rec} = z^{\rm nc}_{\rm rec}$ are removed from the distributions. In addition to that, the events

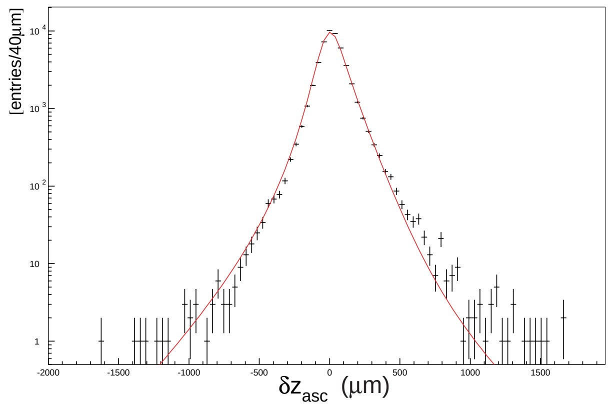


Figure E.8: Residual distribution of $z_{\rm asc}$. Superimposed curve is a fit to $R_{\rm asc} \otimes R_{\rm np}$, where $\tau^p_{\rm np}$ and $\tau^n_{\rm np}$ in $R_{\rm np}$ are constants. The fitted resolution underestimates the residual distribution. The histogram is obtained from $B^0 \to J/\psi K_S$ signal Monte Carlo, whose $B_{\rm asc}$ vertex is reconstructed with multiple tracks.

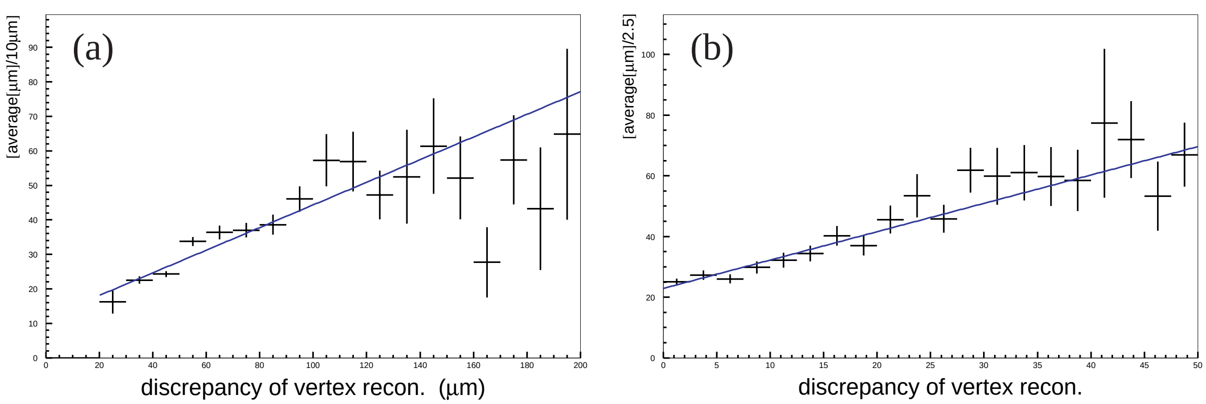


Figure E.9: Averages of $z^{\rm np}_{\rm rec} - z^{\rm nc}_{\rm rec}$ as a function of (a) $\sigma_{\rm asc}$ and (b) $\xi_{\rm np}$. The vertices are reconstructed with at least two tracks. The events with $z^{\rm np}_{\rm rec} < z^{\rm nc}_{\rm rec}$ are excluded.

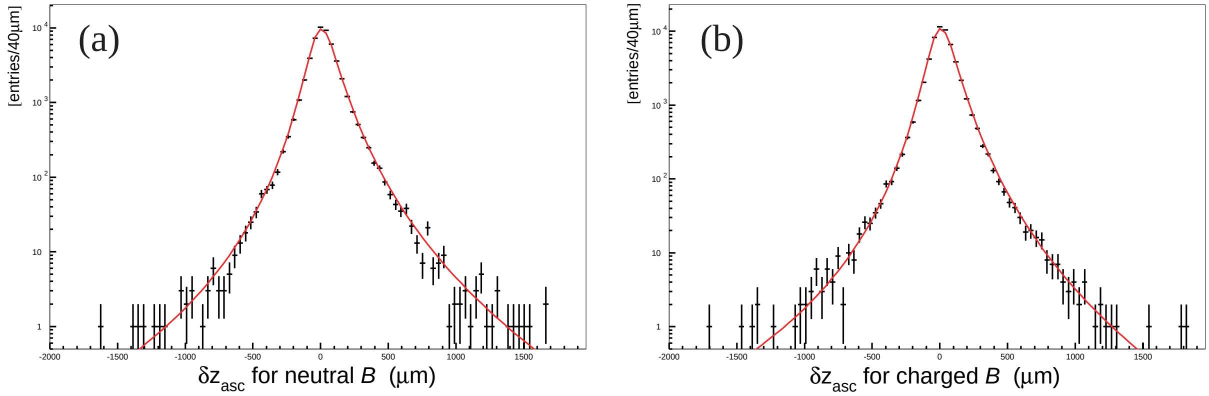


Figure E.10: Residual distributions of $z_{\rm asc}$ for the $B_{\rm asc}$ mesons (a) which are neutral, and (b) which are charged. The decay vertices of B are reconstructed from multiple tracks. (a) is obtained from $B^0 \to J/\psi K_S$ signal Monte Carlo, and (b) is obtained from $B^+ \to J/\psi K^-$ signal Monte Carlo.

are rejected that do not hold long living charm mesons, D^0, D^+, or D_s^+ in the decay chain originated by associated B.

$$\tau^p_{\rm np} = \tau^0_p + \tau^1_p \times (s^0_{\rm asc} + s^1_{\rm asc}\xi_{\rm asc}) \times \sigma_{\rm asc}, \tag{E.17}$$

$$\tau^n_{\rm np} = \tau^0_n + \tau^1_n \times (s^0_{\rm asc} + s^1_{\rm asc}\xi_{\rm asc}) \times \sigma_{\rm asc}. \tag{E.18}$$

As a results of this study, we introduces linear dependency to $\tau^p_{\rm np}$ and $\tau^n_{\rm np}$ on $\sigma_{\rm asc}$ and $\xi_{\rm asc}$. When $\tau^p_{\rm np}$ and $\tau^n_{\rm np}$ depends on both $\sigma_{\rm asc}$ and $\xi_{\rm asc}$, it is natural to consider that $\tau^p_{\rm np}$ and $\tau^n_{\rm np}$ depends on the product, as we discussed in the detector resolution paragraph. Figure E.10 shows the residual distribution of $z_{\rm asc}$ reconstructed by multiple tracks with a fitted curve to the convolution of $R_{\rm asc} \otimes R_{\rm np}$. The linear dependences as equations (E.17) and (E.18) to $\tau^p_{\rm np}$ and $\tau^n_{\rm np}$ in $R_{\rm np}$ are introduced. The distributions are (a) $z_{\rm asc}$ residual obtained from $\Upsilon(4S)$ decayed into neutral B pairs, and (b) that obtained from charged B pairs. We assign different $R_{\rm np}$ shapes for neutral B and charged B, because the yields of charmed mesons, D^0, D^+, and D_s^+ that have different lifetimes are not same between neutral B and charged B. Therefore, the charge dependence of B meson in $R_{\rm np}$ should be taken into account.

Above discussion was made for multiple track vertices. Then we discuss about the smearing due to non-primary track on single track vertices. Figure E.11 shows the distribution of $(z^{\rm np}_{\rm rec} - z^{\rm nc}_{\rm rec})$ as a function of $\sigma_{\rm rec}$, which is obtained similarly as Figure E.9 (a). As we saw in Figure E.9 (a), linear correlation is also observed in the single track case. Since $R_{\rm asc}$ for single track events is defined as a sum of two Gaussians, if we introduce a dependence of vertex reconstruction error on $\tau_{\rm np}$, we should convolute different $R_{\rm np}$ for each of main Gaussian and tail Gaussian, since each Gaussian is considered to have different

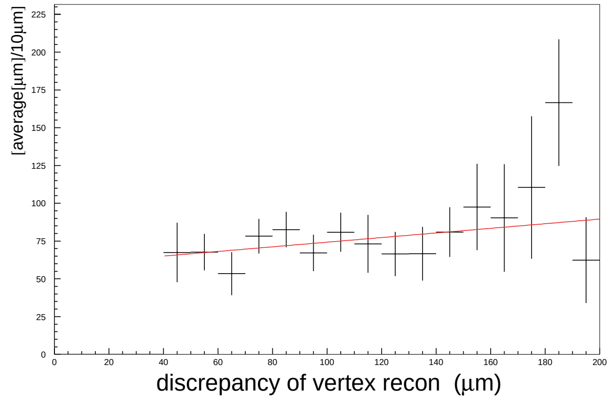


Figure E.11: Averages of $z_{\rm rec}^{\rm np} - z_{\rm rec}^{\rm nc}$ meson flight length as a function of $\sigma_{\rm asc}$. The events with $z_{\rm rec}^{\rm np} < z_{\rm rec}^{\rm nc}$ are excluded.

estimation of the errors of the vertex reconstruction. With definitions of $R_{\rm np}$ lifetimes

$$\begin{cases} (\tau_{\rm np}^{p})_{\rm main} = \tau_p^0 + \tau_p^1 \times s_{\rm main} \times \sigma_{\rm asc} \\ (\tau_{\rm np}^{n})_{\rm main} = \tau_n^0 + \tau_n^1 \times s_{\rm main} \times \sigma_{\rm asc} \end{cases} \tag{E.19}$$

$$\begin{cases} (\tau_{\rm np}^{p})_{\rm tail} = \tau_p^0 + \tau_p^1 \times s_{\rm tail} \times \sigma_{\rm asc} \\ (\tau_{\rm np}^{n})_{\rm tail} = \tau_n^0 + \tau_n^1 \times s_{\rm tail} \times \sigma_{\rm asc} \end{cases} \tag{E.20}$$

we define the convolution of $R_{\rm asc}$ and $R_{\rm np}$ as follows:

$$\int_{-\infty}^{\infty} {\rm d}t \left[R_{\rm main}^{\rm asc}(\delta(\Delta t) - t) \cdot R_{\rm main}^{\rm np}(t) + R_{\rm tail}^{\rm asc}(\delta(\Delta t) - t) \cdot R_{\rm tail}^{\rm np}(t) \right]. \tag{E.21}$$

Figure E.12 shows the residual distribution of $z_{\rm asc}$ reconstructed with single track. Superimposed curves are fit to the resolution defined as equation (E.21).

In summary, the function form of $R_{\rm np}$ and the convolution of $R_{\rm det}$ and $R_{\rm np}$ depends on the number of tracks in associated side B. When $z_{\rm asc}$ is multiple track vertex, the convolution of $R_{\rm det}$ and $R_{\rm np}$ is written as

$$R_{\rm det} \otimes R_{\rm np} = \int_{-\infty}^{\infty} {\rm d}t R_{\rm rec}(\delta(\Delta t) - t) \int_{-\infty}^{\infty} {\rm d}t' R_{\rm asc}^{\rm multiple}(t - t') R_{\rm np}^{\rm multiple}(t'), \tag{E.22}$$

where $R_{\rm np}^{\rm multiple}$ is defined as equation (E.14). In case of single track vertices, the convolution is

$$\begin{aligned} R_{\rm det} \otimes R_{\rm np} &= \int_{-\infty}^{\infty} {\rm d}t R_{\rm rec}(\delta(\Delta t) - t) \\ &\cdot \int_{-\infty}^{\infty} {\rm d}t' \left[(R_{\rm asc}^{\rm single})_{\rm main}(t - t') R_{\rm main}^{\rm np}(t') + (R_{\rm asc}^{\rm single})_{\rm tail}(t - t') R_{\rm tail}^{\rm np}(t') \right]. \end{aligned} \tag{E.23}$$

The parameters for $R_{\rm np}$ determined by the unbinned-maximum-likelihood fit to the Δt distribution obtained with the Monte Carlo simulation.

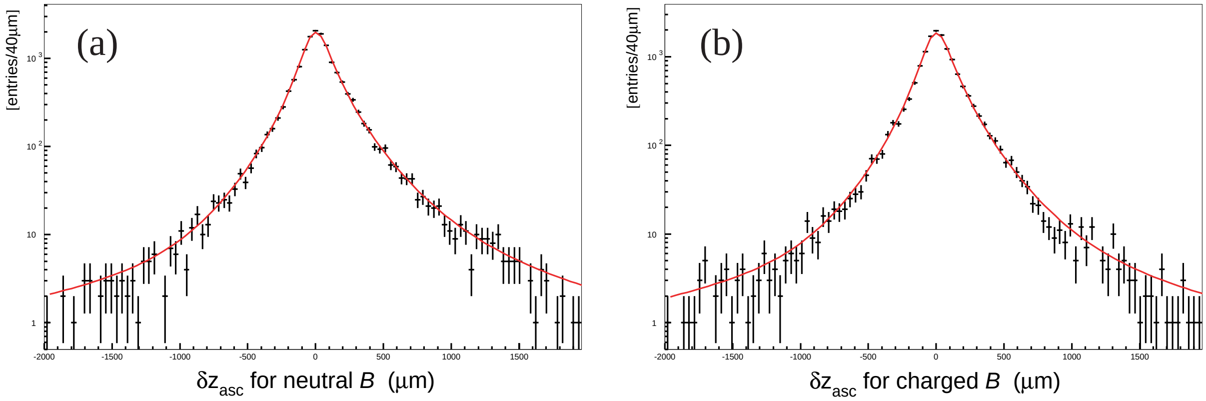


Figure E.12: Residual distributions of z_{asc} for the B_{asc} mesons (a) which are neutral, and (b) which are charged. The decay vertices of B are reconstructed from single track. (a) is obtained from $B^0 \to J/\psi K_S$ signal Monte Carlo, and (b) is obtained from $B^+ \to J/\psi K^-$ signal Monte Carlo.

Table E.1: List of fitted parameters to R_{np} that are determined with the Monte Carlo simulation. The unit of τ_p^0 for both B^0 decays and B^+ decays and for multiple tracks and single track is ps, and the unit of τ_p^1 is ps/ps.

(a) Parameters for B^0 decays.

Fit parameters	Fitted values
$(f_\delta)^{\mathrm{multiple}}$	0.676 ± 0.007
$(f_p)^{\mathrm{multiple}}$	0.955 ± 0.004
$(\tau_p^0)^{\mathrm{multiple}}$	-0.010 ± 0.011
$(\tau_p^1)^{\mathrm{multiple}}$	$0.927^{+0.025}_{-0.024}$
$(\tau_n^0)^{\mathrm{multiple}}$	$-0.194^{+0.078}_{-0.077}$
$(\tau_n^1)^{\mathrm{multiple}}$	$1.990^{+0.182}_{-0.169}$
$(f_\delta)^{\mathrm{single}}$	$0.787^{+0.010}_{-0.011}$
$(f_p)^{\mathrm{single}}$	$0.790^{+0.020}_{-0.021}$
$(\tau_p^0)^{\mathrm{single}}$	$0.108^{+0.068}_{-0.067}$
$(\tau_p^1)^{\mathrm{single}}$	$1.321^{+0.099}_{-0.094}$
$(\tau_n^0)^{\mathrm{single}}$	$-0.281^{+0.130}_{-0.147}$
$(\tau_n^1)^{\mathrm{single}}$	$1.583^{+0.213}_{-0.184}$

(b) Parameters for B^+ decays.

Fit parameters	Fitted values
$(f_\delta)^{\mathrm{multiple}}$	0.650 ± 0.010
$(f_p)^{\mathrm{multiple}}$	0.963 ± 0.004
$(\tau_p^0)^{\mathrm{multiple}}$	0.037 ± 0.012
$(\tau_p^1)^{\mathrm{multiple}}$	0.674 ± 0.025
$(\tau_n^0)^{\mathrm{multiple}}$	-0.269 ± 0.099
$(\tau_n^1)^{\mathrm{multiple}}$	$2.070^{+0.235}_{-0.213}$
$(f_\delta)^{\mathrm{single}}$	$0.763^{+0.017}_{-0.018}$
$(f_p)^{\mathrm{single}}$	$0.757^{+0.025}_{-0.026}$
$(\tau_p^0)^{\mathrm{single}}$	$-0.019^{+0.066}_{-0.065}$
$(\tau_p^1)^{\mathrm{single}}$	$1.113^{+0.099}_{-0.092}$
$(\tau_n^0)^{\mathrm{single}}$	$-0.375^{+0.111}_{-0.122}$
$(\tau_n^1)^{\mathrm{single}}$	$1.548^{+0.207}_{-0.182}$

Smearing Due to Kinematic Approximation

Δt error due to kinematic approximation, $\Delta t \simeq \Delta z/[c(\beta\gamma)_\Upsilon]$ is studied in this paragraph. $(\beta\gamma)$ of reconstructed B and associated B are calculated as

$$\begin{aligned}(\beta\gamma)_{\rm rec} &= \frac{(\beta\gamma)_\Upsilon E_B^{\rm cms} + \gamma_\Upsilon p_B^{\rm cms}\cos\theta_B^{\rm cms}}{M_B} = (\beta\gamma)_\Upsilon\left(\frac{E_B^{\rm cms}}{M_B} + \frac{p_B^{\rm cms}\cos\theta_B^{\rm cms}}{\beta_\Upsilon M_B}\right) \\ &\equiv (\beta\gamma)_\Upsilon(a_k + c_k), \qquad \text{(E.24)} \\ (\beta\gamma)_{\rm asc} &= (\beta\gamma)_\Upsilon\left(\frac{E_B^{\rm cms}}{M_B} - \frac{p_B^{\rm cms}\cos\theta_B^{\rm cms}}{\beta_\Upsilon M_B}\right) = (\beta\gamma)_\Upsilon(a_k - c_k), \qquad \text{(E.25)}\end{aligned}$$

where $E_B^{\rm cms}$, $p_B^{\rm cms}$, and $\theta_B^{\rm cms}$ are the energy, momentum and angle from the beam direction for the $B_{\rm rec}$ in the cms, and M_B is a mass of B meson. Let proper times of $B_{\rm rec}$ and $B_{\rm asc}$ as $t_{\rm rec}^{\rm gen}$ and $t_{\rm asc}^{\rm gen}$. The true proper time difference is $\Delta t_{\rm true} = t_{\rm rec}^{\rm gen} - t_{\rm asc}^{\rm gen}$. The flight length of each B meson can be written as

$$\begin{aligned}z_{\rm rec}^{\rm gen} - z_\Upsilon^{\rm gen} &= c(\beta\gamma)_{\rm rec}\, t_{\rm rec}^{\rm gen} = c(\beta\gamma)_\Upsilon(a_k + c_k)\, t_{\rm rec}^{\rm gen}, \qquad \text{(E.26)} \\ z_{\rm asc}^{\rm gen} - z_\Upsilon^{\rm gen} &= c(\beta\gamma)_{\rm asc}\, t_{\rm asc}^{\rm gen} = c(\beta\gamma)_\Upsilon(a_k - c_k)\, t_{\rm asc}^{\rm gen}. \qquad \text{(E.27)}\end{aligned}$$

Using above equations, nominal proper time difference as defined in equation (E.1) becomes

$$\begin{aligned}\Delta t_{\rm nom} &= \frac{\Delta z_{\rm true}}{c(\beta\gamma)_\Upsilon} \\ &= (a_k + c_k)\, t_{\rm rec}^{\rm gen} - (a_k - c_k)\, t_{\rm asc}^{\rm gen} \qquad \text{(E.28)}\end{aligned}$$

The difference between $\Delta t_{\rm true}$ and $\Delta t_{\rm nom}$ generates the smearing. $R_{\rm k}$ is represented as follows

$$\begin{aligned}R_{\rm k}(\Delta t_{\rm nom} - \Delta t_{\rm true}) =& \\ &\Bigg[\iint {\rm d}t_{\rm rec}^{\rm gen}\, {\rm d}t_{\rm asc}^{\rm gen}\, f_{\rm rec}(t_{\rm rec}^{\rm gen}) \cdot f_{\rm asc}(t_{\rm asc}^{\rm gen}) \\ &\cdot \delta\left(\Delta t_{\rm true} - (t_{\rm rec}^{\rm gen} - t_{\rm rec}^{\rm gen})\right) \cdot \delta\left(\Delta t_{\rm nom} - ((a_k + c_k)\, t_{\rm rec}^{\rm gen} - (a_k - c_k)\, t_{\rm asc}^{\rm gen})\right)\Bigg] \\ &\Bigg/ \left[\iint {\rm d}t_{\rm rec}^{\rm gen}\, {\rm d}t_{\rm asc}^{\rm gen}\, f_{\rm rec}(t_{\rm rec}^{\rm gen})\, f_{\rm asc}(t_{\rm asc}^{\rm gen})\, \delta(\Delta t_{\rm true} - (t_{\rm rec}^{\rm gen} - t_{\rm rec}^{\rm gen}))\right]. \qquad \text{(E.29)}\end{aligned}$$

Here $f_{\rm rec}(t_{\rm asc}^{\rm rec})$ and $f_{\rm asc}(t_{\rm asc}^{\rm gen})$ are the probability density functions for $t_{\rm asc}^{\rm rec}$ and $t_{\rm asc}^{\rm gen}$:

$$\begin{aligned}f_{\rm rec}(t_{\rm rec}^{\rm gen}) &= E_p(t_{\rm rec}^{\rm gen}; \tau_B), \qquad \text{(E.30)} \\ f_{\rm asc}(t_{\rm asc}^{\rm gen}) &= E_p(t_{\rm asc}^{\rm gen}; \tau_B). \qquad \text{(E.31)}\end{aligned}$$

Equation (E.29) is calculated as follows:

$$R_\mathrm{k}(\Delta t_\mathrm{nom} - \Delta t_\mathrm{true}) = \\ \begin{cases} E_p\left((\Delta t_\mathrm{nom} - \Delta t_\mathrm{true}) - \left[(a_k - 1)\Delta t_\mathrm{true} + c_k|\Delta t_\mathrm{true}|\right];\ |c_k|\tau_B\right) & (c_k > 0) \\ \delta((\Delta t_\mathrm{nom} - \Delta t_\mathrm{true}) - (a_k - 1)\Delta t_\mathrm{true}) & (c_k = 0) \\ E_n\left((\Delta t_\mathrm{nom} - \Delta t_\mathrm{true}) - \left[(a_k - 1)\Delta t_\mathrm{true} + c_k|\Delta t_\mathrm{true}|\right];\ |c_k|\tau_B\right) & (c_k < 0) \end{cases} \tag{E.32}$$

The numerator of (E.32) is a probability to obtain $t_\mathrm{rec}^\mathrm{gen}$ and $t_\mathrm{asc}^\mathrm{gen}$ under the constraint that we have to observe Δt and Δz. Since R_k must be defined for any Δt, the constraint for Δt must be canceled. The cancellation is made by the denominator of (E.32). Here is an example of the usage of R_k. If equation (E.32) is convoluted with signal probability density function, equation (E.4), the distribution of Δz_true is obtained:

$$f_{\Delta z}(\Delta z_\mathrm{true}) = \frac{1}{2a_k\tau_B}\exp\left(-\frac{|\Delta z_\mathrm{true}|}{c(\beta\gamma)_\Upsilon\ (a_k \pm c_k)\tau_B}\right), \tag{E.33}$$

where compound sign mark takes same sign that of Δz_true. Figure E.13 (a) shows distribution of $\Delta z/[c(\beta\gamma)_\Upsilon] - \Delta t$. Figure E.13 (b) and (c) show same distribution as (a), where (b) for $0° \leq \cos\theta_B^\mathrm{cms} < 60°$, or $120° < \cos\theta_B^\mathrm{cms} \leq 180°$, and (c) for $60° \leq \cos\theta_B^\mathrm{cms} \leq 120°$. Superimposed lines show R_k. When the B is generated with its momentum parallel to beam axis (case of (b)), the residual distribution tends to be narrower around $\Delta z/[c(\beta\gamma)_\Upsilon] = \Delta t$, since $(\beta\gamma)_B \simeq (\beta\gamma)_\Upsilon$ becomes good approximation. Otherwise, (case of (c)) the distribution has peak other than $\Delta z/[c(\beta\gamma)_\Upsilon] = \Delta t$. In both cases, obtained resolution R_k represents the distribution well. Figure E.14 shows the distribution of generated Δz, where superimposed line shows convoluted signal probability with R_k as shown in equation (E.33).

Outlier

The suppression of impact parameter selection of $|\delta z| < 1800\ \mu$m causes significant outlier component due to faults in the vertex reconstructions. Unless the outlier is not correctly considered, the measured lifetime becomes dependent on $|\Delta t|$ selection range, since undescribed event by the probability density function with $\Delta t \gg 1$ can easily increase the fitted lifetimes. Because the standard deviation of the outlier is expected to be much larger than B meson lifetimes as $\sim 20\tau_B$, we omit the convolution of this component with $f(\Delta t;\ \tau_B)$. We represent the outlier component, $p_\mathrm{ol}(\Delta t)$, by single Gaussian:

$$p_\mathrm{ol}(\Delta t) = G(\Delta t; \sigma_\mathrm{ol}). \tag{E.34}$$

The fraction of outlier, f_ol is different between the events where both z_rec and z_asc are obtained with multiple tracks and the events where at least one of z_rec and z_asc is obtained

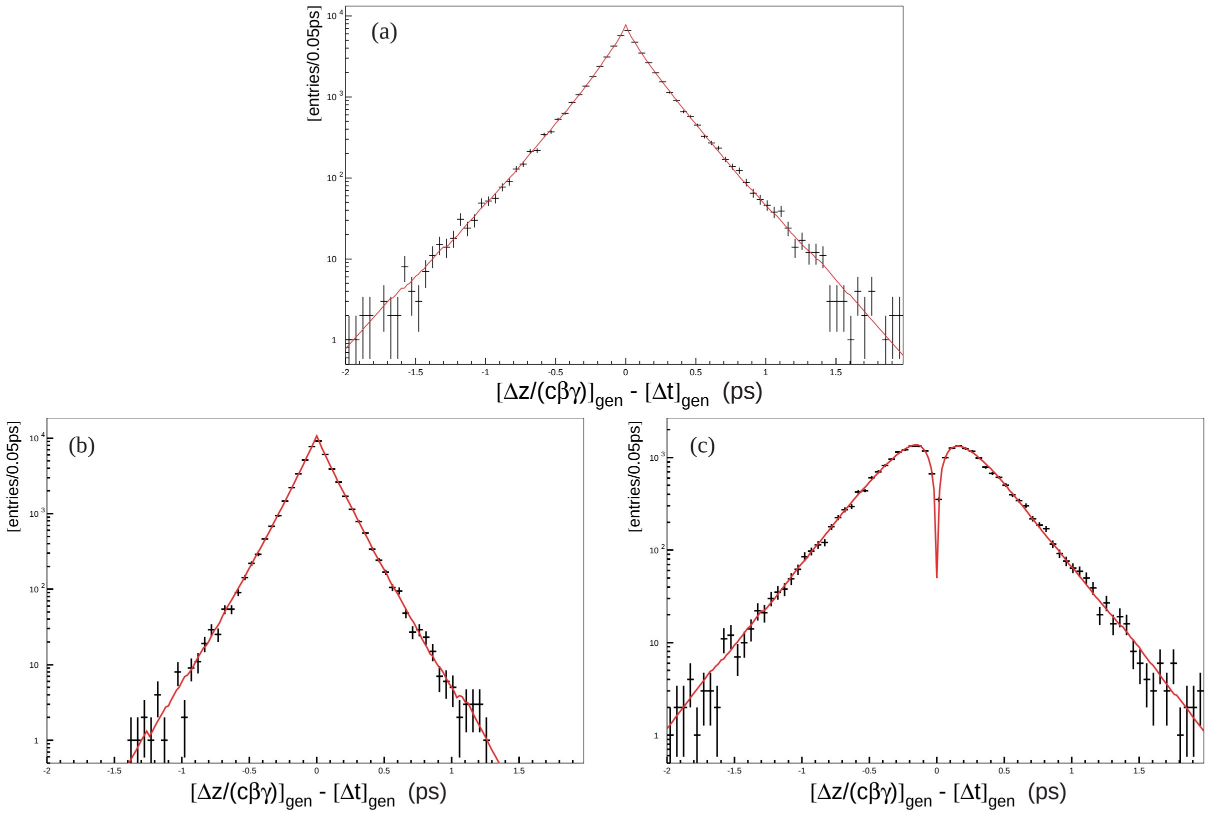


Figure E.13: Distributions of $\Delta z/[c(\beta\gamma)_\Upsilon] - \Delta t$, where (a) for all regions, (b) for $0^\circ \leq \theta_B^{\rm cms} < 60^\circ$, or $120^\circ < \theta_B^{\rm cms} \leq 180^\circ$, and (c) for $60^\circ \leq \theta_B^{\rm cms} \leq 120^\circ$. The superimposed lines show $R_{\rm k}$.

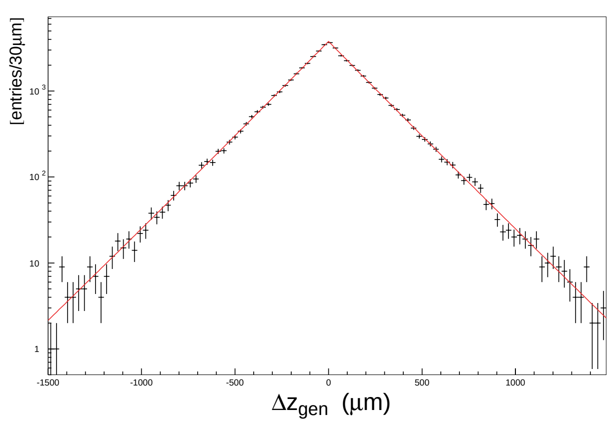


Figure E.14: Distribution of generated Δz. The superimposed line shows convoluted signal probability with $R_{\rm k}$.

from a single track. We redefine the signal probability density function, equation (E.5), with outlier component as follows:

$$\mathcal{P}_{\rm sig}(\Delta t;\ \tau_B) = (1-f_{\rm ol})\cdot\int_{-\infty}^{\infty} {\rm d}(\Delta t')\ \mathcal{R}(\Delta t-\Delta t')\cdot f(\Delta t';\ \tau_B) + f_{\rm ol}\cdot p_{\rm ol}(\Delta t). \qquad \text{(E.35)}$$

E.2.2 Background Shape

The proper-time difference distribution for the background events is obtained from the ΔE-$M_{\rm bc}$ background dominated region. The background shape is fitted to the sum of the main part of the background Δt distribution ($p_{\rm bkg}$) and the outlier part ($p_{\rm ol}^{\rm bkg}$):

$$\mathcal{P}_{\rm bkg}(\Delta t) = (1-f_{\rm ol}^{\rm bkg})\cdot p_{\rm bkg}(\Delta t) + f_{\rm ol}^{\rm bkg}\cdot p_{\rm ol}^{\rm bkg}(\Delta t). \qquad \text{(E.36)}$$

It is natural to introduce the outlier component for the background probability density function because this is originated from failure in the vertex reconstruction. The main part, $p_{\rm bkg}$, is represented by the convolution of a sum of two Gaussians and a sum of δ-function and lifetime distribution:

$$\begin{aligned} p_{\rm bkg}(\Delta t) \quad = \quad & \int_{-\infty}^{\infty} {\rm d}(\Delta z')\ \Big[(1-f_{\rm tail}^{\rm bkg})\, G\left(\Delta t-\Delta z';\ s_{\rm main}^{\rm bkg}\sqrt{\sigma_{\rm rec}^2+\sigma_{\rm asc}^2}\right) \\ & \qquad\qquad + f_{\rm tail}^{\rm bkg}\, G\left(\Delta t-\Delta z';\ s_{\rm tail}^{\rm bkg}\sqrt{\sigma_{\rm rec}^2+\sigma_{\rm asc}^2}\right)\Big] \\ & \times \left[f_\delta^{\rm bkg}\ \delta(\Delta z'-\mu_\delta^{\rm bkg}) + (1-f_\delta^{\rm bkg})\frac{1}{2\tau_{\rm bkg}}\exp\left(-\frac{|\Delta z'-\mu_\tau^{\rm bkg}|}{\tau_{\rm bkg}}\right)\right]. \end{aligned} \qquad \text{(E.37)}$$

Different values are used for $s_{\rm main}^{\rm bkg}$, $s_{\rm tail}^{\rm bkg}$, $f_{\rm tail}^{\rm bkg}$, and $f_\delta^{\rm bkg}$ in the case that both $z_{\rm rec}$ and $z_{\rm asc}$ are reconstructed with multiple tracks and in the case that at least one of $z_{\rm rec}$ and $z_{\rm asc}$ is reconstructed with a single track. The outlier part, $p_{\rm ol}^{\rm bkg}$, has the same function form as equation (E.34). The parameters for background shapes obtained from the data in ΔE-$M_{\rm bc}$ sideband regions are listed in Table E.2. The Δt distribution in background dominated region is shown in Figure E.15 mode by mode, with superimposed fitted curves to the background shapes obtained as Table E.2. After the fit, we found discrepancy in $f_\delta^{\rm bkg}$ between the background events in ΔE-$M_{\rm bc}$ sideband region and in the signal region for $B\to DX$ Monte Carlo sample. We calculate the ratio of $f_\delta^{\rm bkg}$ in the signal region to $f_\delta^{\rm bkg}$ in the sideband region, $(f_\delta^{\rm bkg})^{\rm signal}/(f_\delta^{\rm bkg})^{\rm sideband}$, from the Monte Carlo sample, and multiply $f_\delta^{\rm bkg}$ obtained from the data by this factor when we fit the lifetimes. The ratio obtained from Monte Carlo sample for each mode is listed in Table E.3. Since we do not observe such discrepancy in $B\to J/\psi X$ modes, no correction of $f_\delta^{\rm bkg}$ is applied for those modes.

Table E.2: Background shape parameters. The units of τ_{bkg}, $\mu_\delta^{\mathrm{bkg}}$, and μ_τ^{bkg} in the list are ps.

(a) Parameters for $B \to J/\psi X$ modes.

	$B^0 \to J/\psi K_S$	$B^0 \to J/\psi K^{*0}$	$B^+ \to J/\psi K^+$
$(s_{\mathrm{main}}^{\mathrm{bkg}})_{\mathrm{multiple}}$	$0.40^{+0.13}_{-0.10}$	$1.09^{+0.24}_{-0.26}$	$0.79^{+0.22}_{-0.24}$
$(s_{\mathrm{tail}}^{\mathrm{bkg}})_{\mathrm{multiple}}$	$9.46^{+4.26}_{-2.57}$	$6.97^{+6.39}_{-2.30}$	$1.90^{+0.64}_{-0.37}$
$(f_{\mathrm{tail}}^{\mathrm{bkg}})_{\mathrm{multiple}}$	$0.29^{+0.11}_{-0.12}$	$0.03^{+0.05}_{-0.03}$	$0.66^{+0.21}_{-0.28}$
$(f_\delta^{\mathrm{bkg}})_{\mathrm{multiple}}$	$0.39^{+0.18}_{-0.17}$	0.08 ± 0.08	$0.85^{+0.04}_{-0.05}$
$(s_{\mathrm{main}}^{\mathrm{bkg}})_{\mathrm{single}}$	$0.96^{+0.19}_{-0.23}$	$0.82^{+0.15}_{-0.14}$	$1.03^{+0.09}_{-0.10}$
$(s_{\mathrm{tail}}^{\mathrm{bkg}})_{\mathrm{single}}$	$4.90^{+2.15}_{-1.29}$	$8.33^{+2.96}_{-2.13}$	$11.2^{+4.7}_{-3.0}$
$(f_{\mathrm{tail}}^{\mathrm{bkg}})_{\mathrm{single}}$	$0.16^{+0.16}_{-0.07}$	$0.09^{+0.04}_{-0.03}$	$0.05^{+0.03}_{-0.02}$
$(f_\delta^{\mathrm{bkg}})_{\mathrm{single}}$	$0.38^{+0.28}_{-0.34}$	$0.18^{+0.14}_{-0.12}$	$0.65^{+0.09}_{-0.11}$
τ_{bkg}	$0.39^{+0.24}_{-0.25}$	1.43 ± 0.16	$2.14^{+0.41}_{-0.33}$
$\mu_\delta^{\mathrm{bkg}}$	-0.56 ± 0.09	$-0.70^{+0.28}_{-0.27}$	-0.00 ± 0.05
μ_τ^{bkg}	$0.23^{+0.17}_{-0.19}$	$-0.00^{+0.14}_{-0.13}$	$-0.21^{+0.33}_{-0.37}$

(b) Parameters for $B \to DX$ modes.

	$B^0 \to D^-\pi^+$	$B^0 \to D^{*-}\pi^+$	$B^0 \to D^{*-}\rho^+$	$B^+ \to \overline{D}^0\pi^+$
$(s_{\mathrm{main}}^{\mathrm{bkg}})_{\mathrm{multiple}}$	1.03 ± 0.04	$0.69^{+0.18}_{-0.14}$	0.90 ± 0.07	1.02 ± 0.02
$(s_{\mathrm{tail}}^{\mathrm{bkg}})_{\mathrm{multiple}}$	$3.03^{+0.68}_{-0.37}$	$2.33^{+0.39}_{-0.32}$	$5.19^{+0.74}_{-0.58}$	$6.27^{+0.38}_{-0.34}$
$(f_{\mathrm{tail}}^{\mathrm{bkg}})_{\mathrm{multiple}}$	$0.14^{+0.06}_{-0.05}$	$0.67^{+0.12}_{-0.17}$	$0.13^{+0.04}_{-0.03}$	0.060 ± 0.008
$(f_\delta^{\mathrm{bkg}})_{\mathrm{multiple}}$	$0.67^{+0.08}_{-0.09}$	$0.70^{+0.07}_{-0.08}$	0.34 ± 0.09	0.52 ± 0.03
$(s_{\mathrm{main}}^{\mathrm{bkg}})_{\mathrm{single}}$	0.73 ± 0.07	$0.87^{+0.09}_{-0.10}$	0.93 ± 0.08	0.79 ± 0.03
$(s_{\mathrm{tail}}^{\mathrm{bkg}})_{\mathrm{single}}$	$4.69^{+0.90}_{-0.76}$	$4.54^{+4.05}_{-1.34}$	$3.52^{+0.47}_{-0.39}$	$5.99^{+0.55}_{-0.52}$
$(f_{\mathrm{tail}}^{\mathrm{bkg}})_{\mathrm{single}}$	0.12 ± 0.04	$0.08^{+0.06}_{-0.04}$	0.17 ± 0.05	0.09 ± 0.01
$(f_\delta^{\mathrm{bkg}})_{\mathrm{single}}$	0.35 ± 0.10	0.54 ± 0.11	0.34 ± 0.13	0.33 ± 0.05
τ_{bkg}	$1.10^{+0.14}_{-0.13}$	$1.68^{+0.26}_{-0.20}$	$0.87^{+0.11}_{-0.10}$	0.98 ± 0.05
$\mu_\delta^{\mathrm{bkg}}$	-0.03 ± 0.03	0.00 ± 0.03	$0.11^{+0.07}_{-0.06}$	-0.02 ± 0.01
μ_τ^{bkg}	$0.00^{+0.08}_{0.07}$	$-0.11^{+0.13}_{-0.15}$	$-0.13^{+0.06}_{-0.07}$	-0.11 ± 0.02

Table E.3: Ratio of f_δ^{bkg} in the signal region to the one in the sideband region, $(f_\delta^{\mathrm{bkg}})^{\mathrm{signal}}/(f_\delta^{\mathrm{bkg}})^{\mathrm{sideband}}$.

	$B^0 \to D^-\pi^+$	$B^0 \to D^{*-}\pi^+$	$B^0 \to D^{*-}\rho^+$	$B^+ \to \overline{D}^0\pi^+$
$(f_\delta^{\mathrm{bkg}})^{\mathrm{signal}}_{\mathrm{multiple}}/(f_\delta^{\mathrm{bkg}})^{\mathrm{sideband}}_{\mathrm{multiple}}$	0.59 ± 0.10	0.54 ± 0.10	0.65 ± 0.08	0.82 ± 0.02
$(f_\delta^{\mathrm{bkg}})^{\mathrm{signal}}_{\mathrm{single}}/(f_\delta^{\mathrm{bkg}})^{\mathrm{sideband}}_{\mathrm{single}}$	$0.82^{+0.28}_{-0.30}$	0.59 ± 0.31	$0.55^{+0.20}_{-0.21}$	0.89 ± 0.07

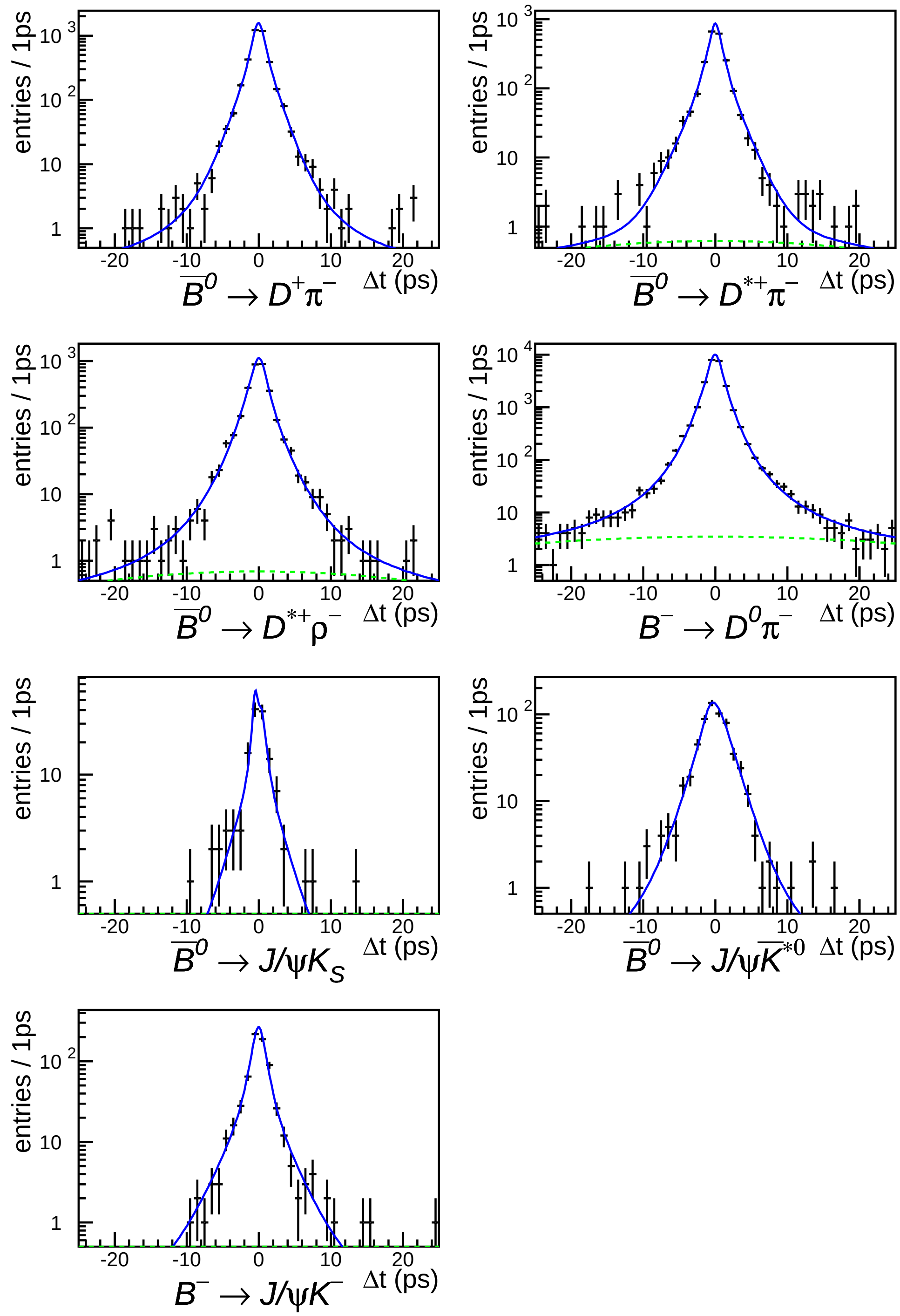


Figure E.15: Δt distribution in background dominated region for each mode. Fitted curves to the background shapes obtained as Table E.2 are superimposed. The broken lines represent the outlier distribution of Δt.

Table E.4: Number of events and for each mode used for the lifetime reconstruction. Purities are also listed. Total number as a sum of B^0 and B^+ mesons is 19910.

(a) Numbers of reconstructed B^0 mesons and their purities.

Mode	Number of events
$B^0 \to J/\psi K_S$	386
$B^0 \to J/\psi K^{*0}$	811
$B^0 \to D^-\pi^+$	2269
$B^0 \to D^{*-}\pi^+$	2495
$B^0 \to D^{*-}\rho^+$	1902
Total	7863

(b) Numbers of reconstructed B^+ mesons and their purities.

Mode	Number of events
$B^+ \to J/\psi K^+$	1804
$B^+ \to \overline{D}^0\pi^+$	10243
Total	12047

E.3 Lifetime Fit

We use 29.1 fb^{-1} data collected by the Belle detector. The numbers of the events fed to the lifetime fit is listed in Table E.4 mode by mode. Number of total events for the fit is 19910.

In order to reduce the relying on the Monte Carlo sample, we determine the detector resolution during fitting of the B meson lifetimes. Moreover, this can reduce systematic uncertainty associated to the resolution function, because The detector resolution is one of the most dominating source for systematic uncertainty. In summary, following parameters are determined from the data simultaneously:

- Lifetimes of B mesons: τ_{B^0} and τ_{B^+}.
- Detector resolution for multiple track vertices: $s^0_{\rm rec}$, $s^1_{\rm rec}$, $s^0_{\rm asc}$ and $s^1_{\rm asc}$.
- Detector resolution for single track vertices: $s_{\rm main}$, $s_{\rm tail}$ and $f_{\rm tail}$.
- Outlier components: $\sigma_{\rm ol}$, $f_{\rm ol}^{\rm single}$ and $f_{\rm ol}^{\rm multi}$.

12 parameters are determined in the final fit. Common parameters for signal resolution function is used for either of B^0 and B^+, and $B \to J/\psi X$ and $B \to DX$. Values for $R_{\rm np}$ are the ones obtained from signal Monte Carlo study as listed in Table E.1. We use the

Table E.5: List of fitted parameters. The input values for the Monte Carlo simulation are $\tau_{\overline{B}^0} = 1.548$ ps and $\tau_{B^+} = 1.653$ ps. Units of τ_{B^0}, τ_{B^+}, and $\sigma_{\rm ol}$ are ps.

(a) Fitted values from Monte Carlo samples.

Fit parameters	Fitted values
τ_{B^0}	1.549 ± 0.006
τ_{B^+}	1.653 ± 0.007
$s^0_{\rm rec}$	0.990 ± 0.055
$s^1_{\rm rec}$	0.054 ± 0.004
$s^0_{\rm asc}$	0.972 ± 0.019
$s^1_{\rm asc}$	0.028 ± 0.001
$s_{\rm main}$	$1.052^{+0.020}_{-0.021}$
$s_{\rm tail}$	$4.436^{+0.870}_{-0.607}$
$f_{\rm tail}$	$0.038^{+0.011}_{-0.009}$
$\sigma_{\rm ol}$	$43.63^{+2.51}_{-2.19}$
$f^{\rm multi}_{\rm ol}$	$(9.9^{+3.8}_{-3.1}) \times 10^{-5}$
$f^{\rm single}_{\rm ol}$	0.022 ± 0.001

(b) Fitted values from 29.1 fb^{-1} data.

Fit parameters	Fitted values
τ_{B^0}	1.554 ± 0.030
τ_{B^+}	1.695 ± 0.026
$s^0_{\rm rec}$	$0.809^{+0.146}_{-0.150}$
$s^1_{\rm rec}$	0.154 ± 0.013
$s^0_{\rm asc}$	$0.753^{+0.064}_{-0.065}$
$s^1_{\rm asc}$	0.064 ± 0.005
$s_{\rm main}$	$0.647^{+0.074}_{-0.083}$
$s_{\rm tail}$	$3.00^{+2.23}_{-0.99}$
$f_{\rm tail}$	$0.083^{+0.083}_{-0.045}$
$\sigma_{\rm ol}$	$36.2^{+5.0}_{-3.5}$
$f^{\rm multi}_{\rm ol}$	$(5.83^{+3.02}_{-2.25}) \times 10^{-4}$
$f^{\rm single}_{\rm ol}$	0.0306 ± 0.0036

same outlier width for multiple track events and single track events, and use same outlier fractions, $f^{\rm multi}_{\rm ol}$ and $f^{\rm single}_{\rm ol}$, for signal events and background events. As for the signal probabilities and the background shapes, we use the values obtained from the fits onto the real data samples, except for the correction factors of $f^{\rm bkg}_{\delta}$.

The reconstructed numbers from the signal Monte Carlo and from 29.1 fb^{-1} real data are listed in Tables E.5 (a) and (b), respectively. The distributions of Δt with the fitted curves are shown in Figure E.16. The fit result of lifetime for each decay mode is listed in Table E.6. All lifetimes are determined simultaneously with sharing resolution function.

We also obtained the lifetime ratio of charged B meson against neutral B meson by replacement of τ_{B^+} with $r_\tau \cdot \tau_{B^0}$. The obtained result is

$$r_\tau = 1.091 \pm 0.023,$$

where the result from Monte Carlo simulation is $r_\tau = 1.067 \pm 0.005$. The input value of r_τ for the Monte Carlo simulation is 1.068.

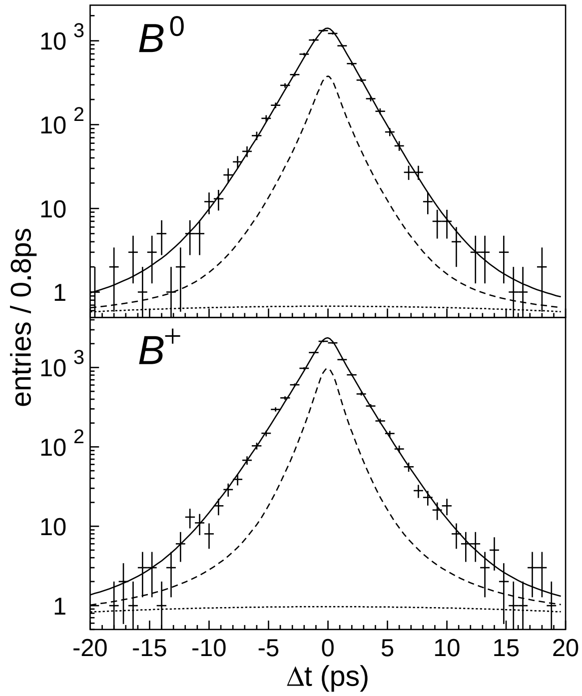


Figure E.16: Δt distribution of neutral B (top) and charged B mesons (bottom), with fitted curves to probability density function of lifetime fit. The broken lines represent the background distribution of Δt and the dotted lines represent the outlier distribution.

Table E.6: Fit result of lifetime for each decay mode. All lifetimes are determined simultaneously with sharing resolution function.

Mode	Lifetime (ps)
$B^0 \to J/\psi K_S$	$1.778^{+0.125}_{-0.116}$
$B^0 \to J/\psi K^{*0}$	$1.368^{+0.068}_{-0.065}$
$B^0 \to D^- \pi^+$	$1.535^{+0.046}_{-0.045}$
$B^0 \to D^{*-} \pi^+$	$1.576^{+0.050}_{-0.048}$
$B^0 \to D^{*-} \rho^+$	$1.618^{+0.065}_{-0.063}$
$B^+ \to J/\psi K^+$	$1.704^{+0.053}_{-0.051}$
$B^+ \to \overline{D}^0 \pi^+$	1.694 ± 0.029

E.4 Systematic Uncertainties

In the following subsections we consider systematic effects from the reconstruction of the B decay distance, and from the fit function. The results are summarized in Tables E.7. All systematic errors are combined in quadrature. The total systematic uncertainties are 0.019 ps for τ_{B^0}, 0.015 ps for τ_{B^+}, and 0.014 for r_τ.

- Δt dependence of reconstruction efficiency

  The Δt dependence of the reconstruction efficiency is assessed by performing binned maximum likelihood fits on the generated Δt distribution for all generated events, and then on the reconstructed events, with a pure exponential function. The difference of lifetime obtained from the two fits is considered to be due to the bias in the reconstruction efficiency. Since we measured no difference between two fits beyond the statistical error, no systematic uncertainty is listed below.

- B decay position

  We take into account the uncertainty of lifetimes due to the discrepancy of interaction point profile between $B\overline{B}$ events and Bhabha events, as described in 4.4. The difference is measured to be quite small. The uncertainty in the evaluation of the B meson flight length is also considered as we did in 4.4. The uncertainty is varied by ± 10 μm from the nominal value, 21 μm and repeat the fits to estimate the error associated with assuming Gaussian function instead of exponential function.

- $|\Delta t|$ tail

  Possible systematic effects due to the track quality selection for the associated B decay vertices have also been studied. Each track selection criteria has been varied by 10% to estimate the systematic error associated with the estimation of fraction for badly measured events.

  The criteria for the reconstructed vertices with the goodness of fit, ξ, is varied from 50 to 200, and repeated the lifetime fit to estimate the systematic dependence on ξ selection criteria. Each change of ξ limit is corresponding to the change of number of events for the fit by $\simeq$ 1%. If the dependence of scaling factors on ξ is not actual linearity, $s_{\rm rec}$ and $s_{\rm asc}$ given as equation (E.11) may be changed by changing the upper limit of ξ. The systematic effect of the changes of $s_{\rm rec}$ and $s_{\rm asc}$ is also estimated by fixing $s_{\rm rec}$ and $s_{\rm asc}$ to the nominal reconstructed values at the criterion of $\xi < 100$. Then the lifetime fitting is repeated by varying ξ selection, $\xi < 50$ and $\xi < 200$. The systematic dependence on ξ criteria varying is observed to be small and no tendency was seen.

  The reconstructed proper-time difference should satisfy $|\Delta t| < 70$ ps. The selection criteria is varied by ± 30 ps. We only observed very small systematic dependence on $|\Delta t|$ selection. It indicates the validity of the resolution representation mainly for the tail region.

- Fit bias and Monte Carlo statistics
  The possible fit bias is studied with Monte Carlo samples. The bias is not observed. Therefore, we do not introduce correction, and we quote the statistical error of the lifetime fit shown in Table E.5 (a) as the uncertainty due to the Monte Carlo statistics.

- Dependence on B meson mass
  B meson mass may affect the lifetimes through R_{k} calculation. The lifetime dependence of B meson mass is measured by varying the mass by $\pm 1\sigma$ of world average [49], but the difference is measured to be negligible. Therefore, no systematic uncertainty on B meson mass is listed below.

- Signal resolution representation
  We use a single Gaussian for the signal resolution function for the multiple track vertices, which is basically equivalent to ignore the small tail fraction observed in each of pull distributions in Figures E.3. We introduce one more Gaussian to the signal resolution function to see the systematic uncertainty due to a choice of resolution function form. We determine the fraction of tail Gaussian and a scaling factor for the tail part with the signal Monte Carlo simulation, and the parameters for R_{np} are re-fitted. The systematic effect due to the choice is estimated to be 0.008 ps for both the τ_{B^0} and τ_{B^+} lifetimes. The systematic effect for the r_τ cancels.

- R_{np} parameterization
  The lifetime fit is repeated with varying the R_{np} parameterization by $\pm 2\sigma$ to estimate the dependence of the lifetimes on the R_{np} parameters.

- Signal probability
  The event by event signal probability is determined based on fit to ΔE-M_{bc} distribution in the data. The systematic error associated to the fit is estimated by varying the parameters determining signal probabilities by $\pm 1\sigma$.

- Background fraction
  ΔE and M_{bc} signal regions are varied from by ± 10 MeV for ΔE and by ± 3 MeV$/c^2$ for M_{bc} to estimate the systematic dependence on the background fraction and the fits are repeated.

- Background shape
  The systematic error associated to the determination of the background shape is estimated by varying the parameters by $\pm 1\sigma$ except for f_δ^{bkg}. The systematic uncertainty due to the f_δ^{bkg} is estimated by varying $\pm 2\sigma$.

Table E.7: Summary of systematic errors for neutral and charged B lifetimes, and their ratio. The errors are combined in quadrature. The units for lifetime uncertainties are ps.

	Uncertainties		
Fitted parameters	τ_{B^0}	τ_{B^+}	r_τ
B decay position	± 0.004	± 0.003	± 0.001
$\lvert\Delta t\rvert$ tail	± 0.007	± 0.005	± 0.002
Monte Carlo statistics	± 0.006	± 0.007	± 0.005
Signal resolution	± 0.008	± 0.008	cancels
$R_{\rm np}$ parameterization	± 0.006	± 0.004	± 0.006
Signal probability	± 0.001	± 0.001	± 0.001
Background fraction	± 0.003	± 0.004	± 0.003
Background shape	± 0.012	± 0.007	± 0.011
Total	± 0.019	± 0.015	± 0.014

E.5 Conclusion

We presented new measurement of B^0 and B^+ meson lifetimes using 29.1 fb^{-1} data sample collected with the Belle detector near $\Upsilon(4S)$ energy. Unbinned-maximum-likelihood fits to proper-time difference distributions of two B meson decay vertices yield following results for the B meson lifetimes and their ratio [71]:

$$\begin{aligned}
\tau_{B^0} &= (1.554 \pm 0.030\ (\text{stat})\ \pm 0.019\ (\text{syst}))\ \text{ps}, \\
\tau_{B^+} &= (1.695 \pm 0.026\ (\text{stat}) \pm 0.015\ (\text{syst}))\ \text{ps}, \\
r_\tau &= 1.091 \pm 0.023\ (\text{stat}) \pm 0.014\ (\text{syst}).
\end{aligned}$$

These are currently the most precise measurements. A value of unity for r_τ is ruled out at a level greater than 3σ and the measured value is consistent with the theoretical prediction [69].

Appendix F

SVD Data Acquisition System

The data acquisition of the SVD is specialized than other DAQ sub-systems mainly due to its enormous number of channels. In this chapter, we give a detail description of the SVD DAQ system.

F.1 Overview

Figure F.1 is an illustration of a readout scheme of the SVD. The SVD comprises three

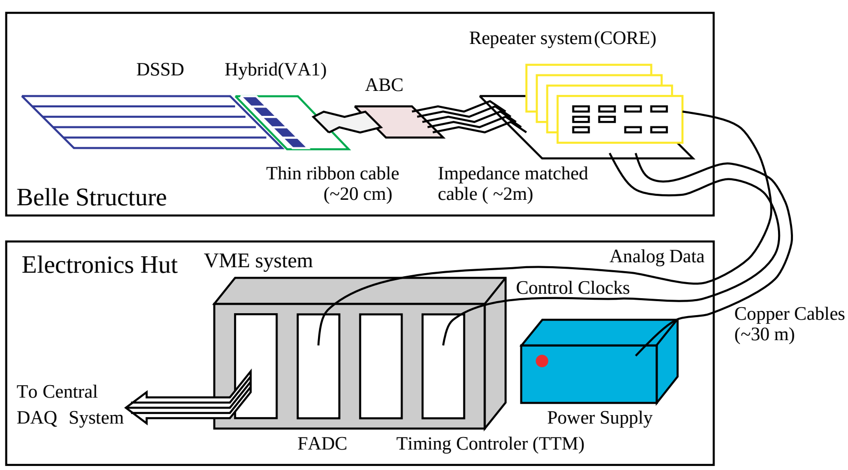


Figure F.1: Readout scheme of the SVD.

layers of full ladders, each of which is composed of either two, three, or four DSSDs. A pair of 640-channel hybrid cards is mounted at both ends of each full ladder. The total number of full ladder is 32 and the number of readout channels is 81920. Hit signals on the silicon strips are scanned by 10 readout chips embedded in each end of the ladder. We use VA1 for the readout chips commercially available from the IDEAs company in Oslo, Norway. Each VA1 includes preamplifiers, shapers, and track-and-hold circuits that

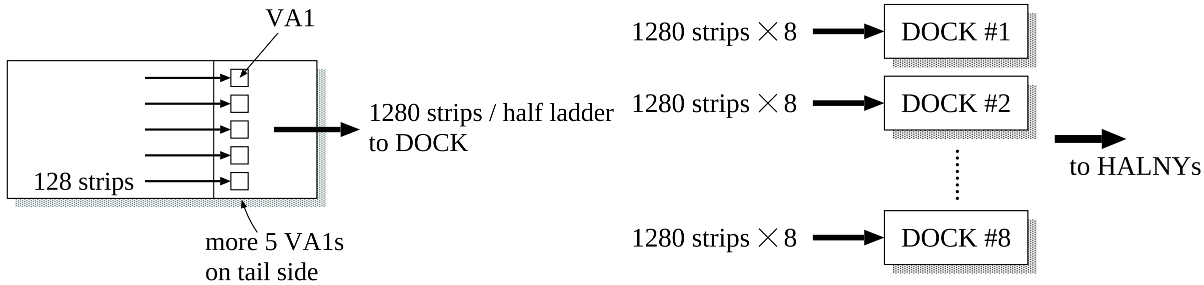


Figure F.2: Illustrations of the SVD data flow from VA1 readout chips to fast A/D converters. The left illustration shows a data path from one half ladder to VA1s. The right illustration shows the data path through the repeater modules to the fast A/D converters. The buffered and amplified data are transfered to each of 32 fast A/D converters.

capture the analogue pulse-hight information on 128 channels. Upon the receipt of the TSC trigger, the system is switched from track mode to hold mode and the stored analog information for each channel is sequentially routed to fast A/D converters via repeater module (Figure F.2 (left)).

Since the cable length between the detector and the electronics hut is about 30 m, the repeater modules are installed near the interaction region ($\sim$ 2m), which are comprised in "CORE system". To meet the readout speed requirements of the Belle DAQ system, the repeater module is capable of supporting multiplex scan rates up to 5 MHz. We have eight repeater modules, where each of them receives 1280×8 strip signals (Figure F.2 (right)). The CORE system is involved in other roles than analog buffering and amplification: analog and digital controls of the VA1 chips, and monitoring of analog levels.

At the electronics hut, the fast A/D conversion is performed for all of scanned data by 32 "HALNY" modules. Figure F.3 shows a block diagram of a HALNY module. Each HALNY module is equipped with four digital signal processors (DSPs), Motorola DSP56302. for elaborated data processings: pedestal subtraction, zero suppression on a channel-by-channel basis, and low-level formatting. The data size is compressed to typically 6% by the algorithm. The DSPs also calculate pedestals and hit thresholds using a dynamic algorithm that automatically adjusts for pedestal shifts and changes in noise level. One DSP receives 640 strip signals from one input FIFO of the HALNY and after A/D conversion and data processing, it puts the processed data to one output FIFO.

The HALNYs are installed in four VME crates, and each VME crate holds eight HALNY

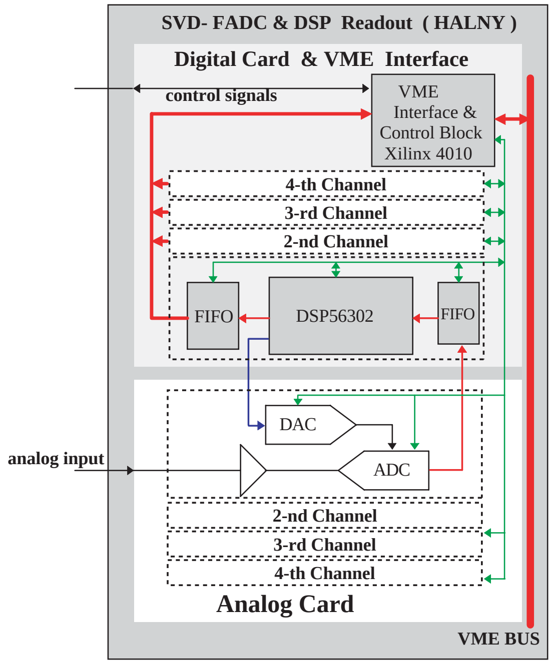


Figure F.3: Illustration of data path within one HALNY module. Data are fed to each of four input FIFOs. The processed data by the DSPs are put to output FIFOs. The SVD DAQ system is equipped with 32 HALNYs.

modules. We divide the SVD DAQ system into four so that the data size from the HALNYs is fitted to the one event builder capability. A SPARC module equipped with 171 MHz SPARC CPU is installed in every VME crate. It works as a bridge between the HALNYs and the event builder. The processed data are read out by the SPARC module from the eight HALNYs via VME bus, then they are transfered to the event builder (Figure F.4), together with data formatting for the CDAQ. A single process on the SPARC module is involved in both a data extraction from the HALNYs and a data transfer to the event builder, so that we minimize the overhead of copying data. The bus arbitration is also made by the SPARC module.

Trigger timing modules (TTMs) are used to control the timing of the front-end readout electronics and to control the HALNY modules. TTMs are driven by the central data acquisition (CDAQ) system. The TTM system consists of nine modules: one master and eight slaves. The two types of modules use same hardware but achieve different functionality through differently configured programmable logic. As we described, the data taking procedure is initiated by the TSC signal. The TSC signal is once fed to master TTM, then distributed to eight slave TTMs. The slave TTMs send the trigger signal for the VA1 chips through the repeater modules so that they hold the signals on the silicon strips. The held signals are sampled and A/D converted by the HALNYs, when

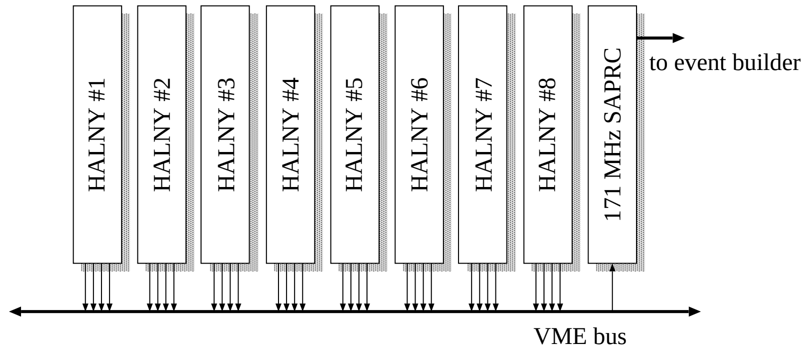


Figure F.4: Illustration of the data path within one VME crate. Each SPARC reads data from eight HALNYs via VME bus. The SVD DAQ system is divided into four such VME crates.

the GDL, which is the final trigger decision logic, is fired. The decision by the GDL is supplied from the CDAQ via timing distributor module (TDM) to the master TTM, and then it is routed to the slave TTMs to generate the sampling timing of the pulse-height for the HALNYs. In the data processing by the HALNYs, TTM forbids the next trigger to be delivered by the SEQ of through the TDM. The signal of ready-for-next-event is notified from HALNYs to the slave TTM.

The SVD run manager controls DAQ VME modules (TDM, TTMs and HALNYs). It initiates or terminated trigger sequence, downloads firmwares to the TTMs and HALNYs, and issues resets at the beginning of each run and when requested by the CDAQ. It also controls the run mode of the VA1 and CORE front-end electronics, via TTMs. Finally, it initiates recovery procedure upon request from the interlock.

The SVD interlock system monitors running condition of the SVD. The data integrity is checked by the consistency of the event size and begin and end event markers. The condition of VA1s is monitored by checking the digital outputs from the hybrids and by tracking their power-supply currents. Temperatures of the hybrids, heat-sinks, and end-rings and water circulation are monitored to make sure that the cooling system is in working order. The beam abort signal from the radiation monitor is also fed into the interlock. When the system detects a fault condition on one of its inputs, it sends abort signal to the run manager. The power supplies for the front-end electronics are shut down by the run manager when such signal is issued.

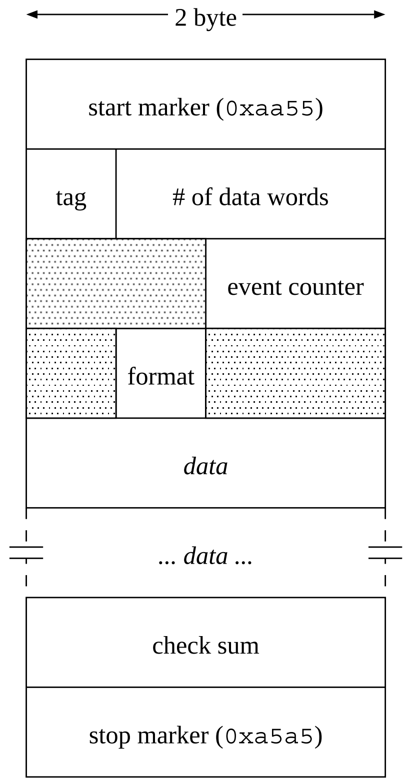


Figure F.5: Data format from one HALNY FIFO.

F.2 Data Transfer via VME Bus

In this section, we describe a detail of the data transfer from the HALNY modules to the SPARC via VME bus. The processed event data by the HALNYs are read by the SPARC module and transfered to the event builder. Figure F.5 shows the data format from one HALNY FIFO. Each data sequence must be started with a start marker and must be ended with a stop marker, otherwise, the fault signal is issued by the readout process on the SPARC to the CDAQ to halt the run. The “tag” number is used to check the synchronization among HALNY FIFOs. When the synchronization is violated, the fault signal is also issued. Through several months, the faults due to HALNYs are seldom issued. The HALNY provides good stability.

The SPARC module polls a one-byte “status register” on the HALNY module via VME bus, and after the register indicates “HALNY-event-ready” it starts the reading procedure of HALNY FIFOs. A width of the VME bus address for one FIFO is 256 bytes. The readout process checks number of words in the data header and repeats reading FIFO in word-by-word mode according to the number. In the method of data transfer via VME bus, we have following choices:

- Programmed I/O – PIO
  The PIO method is most naive method to access the VME bus, because the data on the VME bus are fetched by the CPU in word-by-word. Each action accessing to the VME bus is initiated by the CPU in the VME master. The fetched data are stored to the memory embedded in the VME master module also by the CPU.

- Direct memory access – DMA
  This method is usually faster than PIO method, because the data on the VME bus goes bypass route from the bus to the memory in the VME master. The bypass route is provided by a specialized processor for the VME bus access. The processor takes three parameters for the data transfer: source address on the VME bus, destination address on the master memory, and data size. Setting those parameters and the initialization of the VME processor is the dominating overhead of this method. The data transfer by the VME processor is still performed in word-by-word manner.

- Block transfer – BLT
  The block transfer is the fastest data transfer technique. In the BLT mode, some negotiation protocols on the VME bus are omitted. Multiple data fetching instead of word-by-word manner, the fastest data transfer speed can be achieved. We can combine the BLT with the DMA to realize the highest performance. Because the BLT is not an required functionality by the VME feature, some VME modules do not support this method.

Among those three options, we adopt the PIO method to transfer the data from the HALNY modules to the SPARC. Since the BLT is not supported by the HALNY module, we cannot utilize this method. In the SVD DAQ system, the DMA method is much slower than the PIO method, due to the following reason. Figure F.6 shows the comparison of the data transfer speed between the PIO method and the DMA method from a VME memory module (MVME 202-2F) to the SPARC module. The latency due to the VME memory module can be negligible. The data transfer speed of the PIO is $\sim$ 4 MB/s independent of data size. The DMA realize more than 5 MB/s in the limit of larger data size, however, in the method shows poor performance for the fragmented data, because the overhead for the preparation of the VME bus processor becomes dominant. The maximum data size per one fetch by the SPARC module is by the width of the FIFO address, 256 bytes. Therefore, the PIO is most appropriate method in the data transfer.

When the CPU on the VME master module supports the multiple-word-transfer from its registers to the memory with a single instruction, there is a chance to earn the speed of the data transfer. However, since the SPARC CPU does not have such functionality, reading data in word-by-word manner is still fastest method.

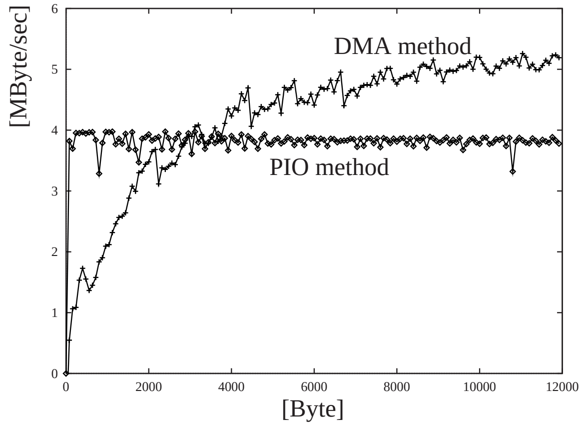


Figure F.6: Data transfer speed from the VME memory module to the SPARC module as a function of data size. The comparison of the PIO method and the DMA method is shown.

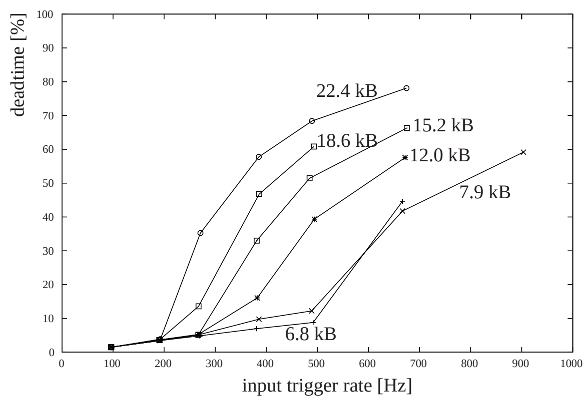


Figure F.7: Dead time of the DAQ as a function of input trigger rate.

F.3 Performance

We studied the performance of the SVD DAQ by checking the dead time of the DAQ system as a function of the input trigger rate while varying the data size. The TDM has functionality to generate a various trigger timing by itself independently of the GDL trigger. The data size from the SVD DAQ is varied by modifying the threshold of noise level supplied to the HALNY DSPs. Figure F.7 shows the result of the performance study. The estimated data size for the luminosity of 5×10^{33} cm^{-2}s^{-1} is $1.0 - 1.2$ kB/event in a sum of four crates. The nominal trigger rate for physics events and background events around the luminosity is ~ 200 Hz. The DAQ system requires less than 10% dead time. The results of the study states that the SVD DAQ drops only small fraction of the events at the typical trigger rate.

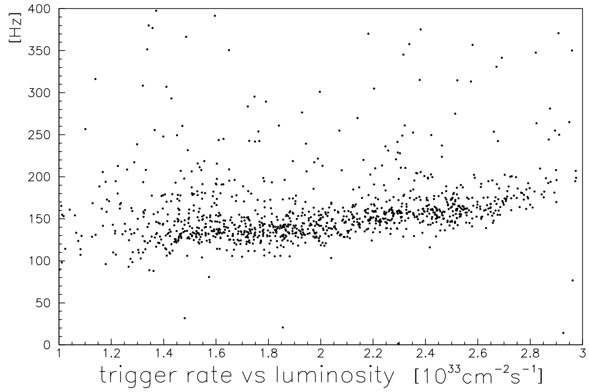


Figure F.8: Scatter plot for the average trigger rate and the average luminosity in recent luminosity runs. The luminosities are measured by the ECL.

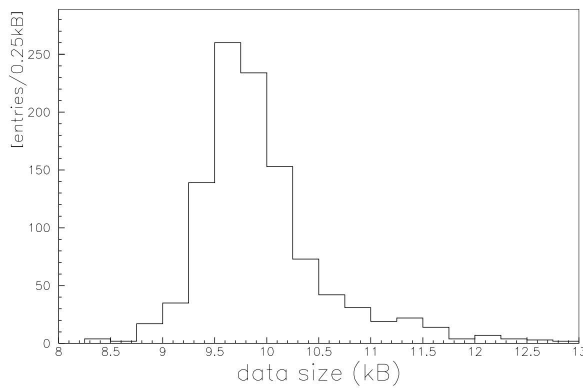


Figure F.9: Distribution of the typical data size from the SVD. A sun of the outputs from the four crates is taken.

Figure F.8 shows the scatter plot for the average trigger rate and the average luminosity for recent luminosity runs. The luminosity average is calculated from the ones at the run start and run stop timing measured by the ECL. A slight dependence of the trigger rate on the luminosity is observed, and the typical trigger rate is measured to be up to 200 Hz for the recent luminosity. The typical data sizes are measured to be independent of the luminosity for $\lesssim 3 \times 10^{33}$ cm^{-2}s^{-1}. It is measured to be $\sim$ 10 kB/event (Figure F.9). The SVD DAQ system shows fine performance up to recent luminosities from the view point of the dead time requirement, and it is still tolerable up to $\sim$ 300 Hz trigger rate, which corresponds to the average luminosity of $\sim 4 \times 10^{33}$ cm^{-2}s^{-1}.

F.4 Future Upgrade

Toward coming increase of accelerator luminosity, the DAQ system should catch up more trigger rate, where the final aim is $>$ 500 Hz. In this chapter, we discuss the future

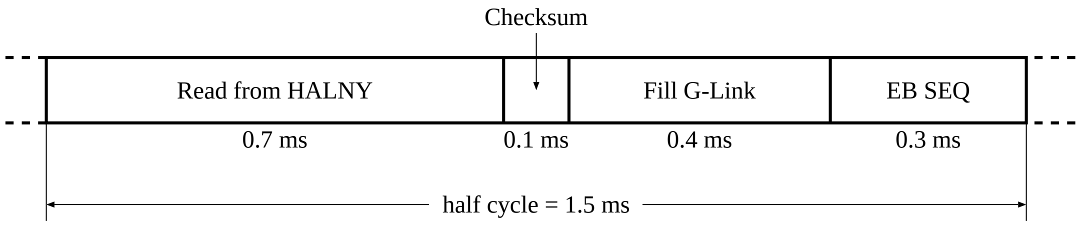


Figure F.10: Timing profile of half event cycle for the SVD data transfer.

upgrade plan while viewing the timing profile of the data transfer cycle.

Four SVD DAQ crates as well as other DAQ sub-systems run in parallel, as we described in Section 2.4.3. Because we construct the event builder as 12×6 barrel shifter, one DAQ sub-system can use only half of one event cycle, and the remaining half cycle is occupied by other sub-system than itself [72]. For the typical event size, 2.5 kB/crate/event, a duration of 1.5 ms is consumed by one of four SVD sub-systems, which corresponds to up to $\sim$ 330 Hz trigger rate. Figure F.10 shows the timing profile for one accessible event cycle for one SVD sub-system to transfer the data for the typical event size. The readout from HALNY takes 0.7 ms, checksum calculation on the SPARC takes 0.1 ms, filling the event onto the output FIFO for the event builder takes 0.4 ms, and event building sequence takes 0.3 ms. Besides the re-configuration of the event builder, we can gain the speed by compressing the read-out time from the HALNYs to the SPARC, which takes almost half of one accessible event cycle. We have a plan to replace the VME buses with PCI buses, which provides much faster data transfer than the VME bus. The data transfer speed of single PCI card is measured to be 40 MB/s [73], which is 10 times of the VME bus speed.

Acknowledgement

First of all, I would like to express my great acknowledgement to Prof. H. Aihara for his supervising my education for high-energy physics.
I greatly appreciate Prof. H. Tajima and Prof. Y. Sakai. They gave me so much advice throughout my study. Moreover, my thesis could not have reached the goal unless their careful readings and corrections.
I also appreciate the all Belle members. I am especially thankful for ICPV, DAQ, and SVD members: Prof. S. Schrenk, Prof. F. Fang, Prof. M. Hazumi, Prof. H. Ozaki, Prof. N. Katayama, Prof. R. Itoh, Prof. M. Nakao, Prof. J. Haba, Prof. T. Tsuboyama, Prof. M. Tanaka, Prof. Y. Yamada, Prof. Y. Ushiroda, Prof. S. Y. Suzuki, Prof. Y. Igarashi, Dr. K. Sumisawa, Dr. J. Yashima, Dr. K. Suzuki, Dr. S. Ichizawa, and Mr. S. Nishida.
I owe a great deal to my colleagues, with whom I shared a life in Aihara laboratory, Dr. J. Tanaka, Dr. M. Yokoyama, Mr. T. Nakadaira, Mr. T. Tomura, Mr. T. Matsubara, Mr. H. Kawai, Mr. N. Uozaki, and Mr. Y. Yamashita. Discussions with them were so valuable for my understanding the physics.
I would like to thank three secretaries: Ms. K. Kono, Ms. M. S. Hikami, and Ms. K. Enomoto for their help of my university life.
At last, I am especially indebted to all of my family. Without their kind helps, I could never finish my work. Before closing my thesis, I would like to express my thank to Ms. N. Koide for her greatest supports.

References

[1] L. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, Phys. Rev. Lett. **13** 138 (1964).

[2] M. Kobayashi and T. Maskawa, Prog. Theor. Phys, **49**, 652 (1973).

[3] J. J. Aubert *et al.*, Phys. Rev. Lett. **33**, 1404 (1974).

[4] "KEKB B-Factory Design Report", KEK Report 95-7, (1995).

[5] KEKB Commissioning Group, KEK Preprint 99-8, (1999);
Y. Funakoshi *et al.*, Proc. 2000 european Particle Accelerator Conference, Vienna, (2000).

[6] A. Abashian *et al.* [BELLE Collaboration], KEK Progress Report 2000-4, (2000).

[7] W. Pauli, In Niels Bohr and the Development of Physics, Oxford, Pergamon 2^{nd} edition, (1995).

[8] T. D. Lee and C. N. Yang, Phys. Rev. **104** 254 (1956).

[9] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, Phys. Rev. **105** 1413 (1957).

[10] R. L. Garwin, L. M. Lederman and M. Weinrich, Phys. Rev. **105**, 1415 (1957).

[11] J. M. Gaillard *et al.*, Phys. Rev. Lett. **18** 20 (1967);
J. W. Cronin *et al.*, Phys. Rev. Lett. **18** 25 (1967).

[12] S. Bennett *et al.*, Phys. Rev. Lett. **19** 993 (1967).

[13] D. E. Dorfan, J. Enstrom, D. Raymond, M. Schwartz, S. G. Wojcicki, D. H. Miller and M. Paciotti, Phys. Rev. Lett. **19**, 987 (1967).

[14] M. Gell-Mann, Phys. Lett. **8**, 214 (1964).

[15] N. Cabbibo, Phys. Rev. Lett. **10**, 531 (1963).

[16] S. L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. **D2**, 1585 (1970).

[17] L. Wolfenstein, Phys. Rev. Lett. **13** 362 (1964).

[18] A. Carter and A. I. Sanda, Phys. Rev. Lett. **45** 952 (1980).

[19] T. Inami and C. S. Lim, Prog. Theor. Phys. **65**, 297 (1981) [Erratum-ibid. **65**, 1772 (1981)];
F. J. Gilman and M. B. Wise, Phys. Rev. **D27**, 1128 (1983);
A. J. Buras, M. Jamin and P. H. Weisz, Nucl. Phys. B **347**, 491 (1990).

[20] J. S. Hagelin, Nucl. Phys. B **193**, 123 (1981).

[21] G. Buchalla, A. J. Buras, M. E. Lautenbacher, Rev. Mod. Phys. **68**, 1125 (1996).

[22] CKM Fitter:
`http://ckmfitter.in2p3.fr/`,
A. Hocker, H. Lacker, S. Laplace, and F. Le Diberder, Eur. Phys. J. C **21**, 225 (2001).

[23] M. Ciuchini *et al.*, JHEP **0107**, 013 (2001) [arXiv:hep-ph/0012308].

[24] K. Ackerstaff *et al.* [OPAL collaboration], Eur. Phys. J. C **5**, 379 (1998).

[25] R. Barate *et al.* [ALEPH Collaboration], Phys. Lett. B **492**, 259 (2000).

[26] T. Affolder *et al.* [CDF Collaboration], Phys. Rev. D **61**, 072005 (2000).

[27] V. Papadimitriou, The Fifth KEK Topical Conference, (2001).

[28] G. Alimonti *et al.* [BELLE Collaboration], Nucl. Instrum. Meth. A **453**, 71 (2000).

[29] H. Hirano *et al.*, Nucl. Instrum. Meth. A **455**, 294 (2000).

[30] T. Iijima *et al.*, Nucl. Instrum. Meth. A **453**, 321 (2000);
K. Suzuki, Ph. D thesis, Graduate School of Science and Technology, Chiba University, (2001).

[31] H. Kichimi *et al.*, Nucl. Instrum. Meth. A **453**, 315 (2000).

[32] H. Ikeda *et al.*, Nucl. Instrum. Meth. A **441**, 401 (2000).

[33] A. Abashian *et al.* [BELLE Collaboration], Nucl. Instrum. Meth. A **449**, 112 (2000).

[34] E. Nygård *et al.*, Nucl. Instrum. Meth. A **301**, 506 (1991);
O. Toker *el at.*, Nucl. Instrum. Meth. A **340**, 572 (1994).

[35] R. E. Kalman, Trans. ASME, J. Bas. Eng. **82D**, 35 (1960);
R. E. Kalman and R. S. Bucy, Trans. ASME, J. Bas. Eng. **83D**, 95 (1961).

[36] K. Hanagaki *et al.*, Belle note #312, (2000), unpublished.

[37] R. Santonico and R. Cardarelli, Nucl. Instrum. Meth. **187**, 377 (1981).

[38] L. Piilonen *et al.*, Belle note #338, (2000), unpublished.

[39] Y. Ushiroda, A. Mohapatra, H. Sakamoto, Y. Sakai, M. Nakao, Q. An and Y. F. Wang, Nucl. Instrum. Meth. A **438**, 460 (1999).

[40] T. Higuchi *et al.*, Proc. Intl. Conf. on Computing in High Energy Physics (1998); T. Higuchi, "Development of The Online Event Selection Program for The BELLE Experiment", Master thesis, (1999), unpublished.

[41] K. Hanagaki, M. Hazumi, and H. Kakuno, Belle note #299, (2000), unpublished.

[42] S. Y. Suzuki *et al.*, "The Belle Event Building System," IEEE Transactions on Nuclear Science, Vol.47, No.2, pp.61-64 (2000).

[43] R. Itoh, Belle note #161, (1996), unpublished.

[44] R. Itoh and S. Ichizawa, Belle note #97, (1996), unpublished.

[45] R. Brun and D. Lienart, CERN-Y250.

[46] `http://www.lns.cornell.edu/public/CLEO/soft/QQ`

[47] R. Itoh, "QQ quick reference guide for BELLE":
`http://bsunsrv1.kek.jp/~software/qq/html/belle_qq.html`

[48] GEANT manual, CERN Program Library Long Writeup W0513, (1993).

[49] D. E. Groom *et al.* [Particle Data Group Collaboration], Eur. Phys. J. C **15**, 1 (2000), and partial update for edition 2002 (`http://pdg.lbl.gov`).

[50] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

[51] B. Casey, Belle note #390, (2001), unpublished.

[52] F. Fang, Belle note #323, (2000), unpublished.

[53] H. Kakuno, Belle note #384, (2001), unpublished.

[54] J. Tanaka, Belle note #194, (2000), unpublished.

[55] M. Hazumi and T. Kawasaki, Belle note #314, (2000), unpublished.

[56] H. Cramer, Mathematical Methods of Statistics, Princeton University Press, (1946).

[57] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).

[58] H. Albrecht *et al.* [ARGUS Collaboration], Phys. Lett. B **241** 278 (1990)

[59] T. Nakadaira, T. Tomura, T. Higuchi, H. Tajima, and H. Aihara, Belle note #290, (2000), unpublished.

[60] J. Tanaka, Ph. D thesis, Department of Physics, Faculty of Science, University of Tokyo, (2001).

[61] A. G. Frodesen *et al.*, Probability and Statistics in Particle Physics, Columbia University Press, (1979).

[62] G. J. Feldman and R. D. Cousins. Phys. Rev. D **57** 3873 (1998).

[63] D. Jackson, Belle note #391, (2001), unpublished.

[64] K. Abe *et al.* [Belle Collaboration], "Measurements of Polarization and CP Asymmetry in $B \to J/\psi + K^*$ decays", BELLE-CONF-0105, (2001).

[65] T. Higuchi, Belle note #220, (2000), unpublished.

[66] W. Gautschi, Comm. ACM. **12**, 635 (1969).

[67] F. Hoshino, private communication.

[68] M. Neubert, C. T. Sachrajda, Nucl. Phys. B **483**, 339 (1997).

[69] D. Becirevic, arXiv:hep-ph/0110124.

[70] A recent and precise measurement is published in: B. Aubert *et al.* [Babar Collaboration], Phys. Rev. Lett. **87** 201803 (2001).

[71] K. Abe *et al.* [Belle Collaboration], KEK Preprint 2001-165, Belle Preprint 2002-2 (2002);
submitted to Phys. Rev. Lett.

[72] S. Y. Suzuki, Ph. D thesis, National High Energy Accelerator Research Organization, (2001).

[73] Y. Yamashita, private communication.