Precise Measurements of Charm Meson Lifetimes and Search for D^0-$\overline{D}^0$ Mixing

Junichi Tanaka
Department of Physics, Faculty of Science,
University of Tokyo

A DISSERTATION
Submitted to the Department of Physics,
Faculty of Science, University of Tokyo.

December 2001

Abstract

We present measurements of charm meson lifetimes and search for D^0-$\overline{D}^0$ mixing using data collected near the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric e^+e^- collider.

The lifetimes and ratios of the lifetimes of D^0, D^+ and D_s^+ mesons are measured to be

$$\begin{aligned}
\tau(D^0) &= 415.0 \pm 1.7(\text{ stat.}) \pm 1.9(\text{ syst.})\text{ fs}, \\
\tau(D^+) &= 1037.2 \,^{+\,12.4}_{-\,12.2}(\text{ stat.}) \,^{+\,6.0}_{-\,6.8}(\text{ syst.})\text{ fs}, \\
\tau(D_s^+) &= 485.7 \,^{+\,7.9}_{-\,7.8}(\text{ stat.}) \,^{+\,4.0}_{-\,5.1}(\text{ syst.})\text{ fs}, \\
\tau(D^+)/\tau(D^0) &= 2.50 \pm 0.03(\text{ stat.}) \pm 0.02(\text{ syst.}), \\
\tau(D_s^+)/\tau(D^0) &= 1.17 \pm 0.02(\text{ stat.}) \pm 0.01(\text{ syst.}),
\end{aligned}$$

using 11.1 fb^{-1} of data. Our results represent a substantial improvement over the best previous published measurements.

The D^0-$\overline{D}^0$ mixing parameter y_{CP} is measured using the lifetime difference of D^0 mesons decaying into the $K^-\pi^+$ state and the CP even eigenstate K^-K^+ using 23.4 fb^{-1} of data. We find

$$y_{CP} = -0.005 \pm 0.010(\text{ stat.}) \,^{+\,0.007}_{-\,0.008}(\text{ syst.}),$$

corresponding to a 95% confidence interval $-0.030 < y_{CP} < 0.020$. Our result is consistent with zero within one standard deviation. It indicates no mixing in the D^0-$\overline{D}^0$ system.

Acknowledgments

It has been a great pleasure to work with and to learn from many people at the University of Tokyo and the Belle collaboration. First, I would like to thank my advisor, Prof. Hiroaki Aihara, for his guidance and encouragement. He gave me a chance to work on research in the high energy physics. I would like to thank colleagues at my laboratory: T. Higuchi, M. Yokoyama, T. Nakadaira, T. Tomura, T. Matsubara, H. Kawai, N. Uozaki and Y. Yamashita for their cooperation and for many discussions with them.

I owe thanks to all Belle collaborators, and in particular the members of the tracking, the charm physics, the ICPV and the SVD groups for their advice and suggestions. I would like to thank Y. Iwasaki, J. Suzuki, T. Karim, T. Matsumoto, Y. Sakai, T. Nozaki, S. Uno, K. Inami, T. Hara, T. Hojo, T. Kawasaki, Y. Ohnishi, T. Kuniya, M. Różańska, A. Bożek, and H. W. Zhao for their helpful advice of the developments of the tracking programs and of the SVD softwares. I would like to thank K. Sumisawa, K. Hanagaki, M. Yamaga, S. Vahsen, M. Hazumi and R. Itoh for their helps to measure and to check our $\sin 2\phi_1$ value. I would like to thank M. Tanaka, J. Haba, Y. Yamada and T. Tsuboyama for their leaderships in the installation of the SVD detector. I would like to thank N. Katayama for his answers to many questions about the Belle computer systems and the programing techniques. I would like to thank B. Yabsley, T. Browder, and C. Everton for their suggestions for my analysis and corrections of my English.

I am especially thankful to Hitoshi Ozaki for his advice and for many suggestions for the developments of tracking programs and physics analysis. I have learned methods and strategies of the charged track reconstructions and the fitting algorithms from him.

I would like to thank Hiroyasu Tajima for his continuous and helpful advice. I have learned methods of physics analysis from him. I'm sure that this thesis did not exist without his great support.

Finally I thank my mother Aiko. T, my sister Satomi. T and my fiancée Mayumi. E for their constant support and encouragement to continue my study of the physics. I'm glad that my father Kiyotaka. T in heaven is pleased with my Ph.D.

Jun-ichi Tanaka

Contents

Chapter 1

Introduction

High energy physics has developed a mixture of theoretical and experimental results known as the Standard Model (SM) [1] which is a gauge theory with an $SU(3) \otimes SU(2) \otimes U(1)$ gauge structure, combining the electroweak theory and QCD.

In the last century, the SM has been tested using many experimental results, and is considered to be a reliable theory to predict what happens in particle interactions. In the new century, the main goals of high energy physics are the search for phenomena which have been predicted by the SM but not yet found, and the search for new physics beyond the SM. These two things are correlated to each other.

We have a few important unresolved matters: Kobayashi–Maskawa mechanism [2] and the Higgs particle. The former, that is, CP violation has been observed in the first year of the new century in the B physics experiments [3], however the latter is not discovered yet. The existence of the Higgs particle is important to conclude that the SM is successful since the mass of W and Z gauge bosons is created by the Higgs mechanism [4]. They are based on the electroweak theory. On the other hand, we have a common problem in the precise prediction of the important physics parameters. It is a complexity of QCD effects that should be resolved in the SM to claim signals for new physics in the precise measurements.

Measurements of individual charm meson lifetimes are useful and important information for the theoretical understanding of heavy flavor decay mechanisms in QCD since the same theory is also used to calculate B meson lifetimes, mixing parameters, Cabibbo-Kobayashi-Maskawa (CKM) parameters extracted from the branching fractions and so on. In the 1980s [5] the experimental measurements of the lifetimes of the charm mesons D^0, D^+, F^+[a] shocked the theoretical community. These lifetimes had been considered to be almost equal, based on the spectator mechanism and the helicity suppression of the non-spectator contributions, but the measured lifetime of the D^+ was longer than that of the D^0 by a factor of 10 or so. Our present understanding of heavy flavor decay mechanisms is still too poor to explain experimental results in detail: for example, the lifetime ratio of the D_s^+ and D^0 mesons is measured to be $1.202^{+0.026}_{-0.023}$ [6], whereas theory predicts 1.00–1.07 [7]. The precise measurement of charm meson lifetimes will help to

[a] F^+ is the old name of D_s^+.

discriminate between the many available QCD models.

In the SM, there is no evidence for particle-antiparticle mixing in the charm-quark sector. Since the charm-quark sector has no measurable mixing, it provides an opportunity to probe new physics coupling to the up-type-quark sector, such as supersymmetry particles, Higgs particles, a fourth quark generation, and leptoquarks. In 2000 the FOCUS collaboration reported a D^0-$\overline{D}^0$ mixing result more than two standard deviations from zero [8], and this result has stimulated extensive theoretical discussion. It is very important to check the FOCUS result to see if new physics is present in the charm-quark sector in the B factory experiments Belle, BABAR and CLEO III, and the proposed next experiment CLEO-c [9].

In this thesis, we present measurements of the lifetimes of charm mesons and the D^0-$\overline{D}^0$ mixing parameter y_{CP} using the Belle detector at the KEKB asymmetric e^+e^- collider. D^0, D^+ and D_s^+ mesons are fully reconstructed via the following decay chains:

$$
\begin{aligned}
&D^0 \to K^-\pi^+ \text{ (with and without } D^{*+} \to D^0\pi^+ \text{ requirement)},\\
&D^0 \to K^-K^+,\\
&D^+ \to K^-\pi^+\pi^+ \text{ (with } D^{*+} \to D^+\pi^0 \text{ requirement)},\\
&D^+ \to \phi\pi^+, \quad \phi \to K^+K^-,\\
&D_s^+ \to \phi\pi^+, \quad \phi \to K^+K^-,\\
&D_s^+ \to \overline{K}^{*0}K^+, \quad \overline{K}^{*0} \to K^-\pi^+.
\end{aligned}
$$

Charge-conjugate modes are implied throughout this thesis. In the lifetime measurement the data sample consists of 11.1 fb^{-1} of e^+e^- collisions near the $\Upsilon(4S)$ resonance. The D^0-$\overline{D}^0$ mixing parameter y_{CP} measurement is based on a larger data sample with an integrated luminosity of 23.4 fb^{-1}.

The outline of this thesis is as follows: the physics formalism describing lifetime and D^0-$\overline{D}^0$ mixing is given in Chapter 2. An overview of the KEKB accelerator and the Belle detector is presented in Chapter 3. Event selection and the reconstruction of charm mesons are explained in Chapter 4. The method of the fit to extract lifetimes and y_{CP} is described in Chapter 5. The fit to data is then described in detail in Chapter 6. The systematic uncertainties are studied in Chapter 7. Finally the conclusion of the thesis is given in Chapter 8.

Chapter 2

Physics Formalism for Charm Meson Lifetimes and D^0-$\overline{D}^0$ Mixing

We describe the mechanisms of charm meson decay and D^0-$\overline{D}^0$ mixing in the following sections.

2.1 Lifetime

We are still learning about the QCD aspect of hadron decay mechanisms. There are many theoretical models explaining inclusive and exclusive decays. Precise measurements of lifetimes and ratios of lifetimes can determine whether existing theoretical models are suitable and provide constraints on new theoretical models.

Table 2.1 shows QCD expectations and measurements of charm lifetime ratios. The observed lifetime difference between D^+ and D^0 is explained by the destructive interference in D^+ Cabibbo-allowed decays between the external spectator and internal spectator decay diagrams shown in Figure 2.1(b) and 2.1(c). The D_s^+ and D^0 lifetimes are theoretically expected to be nearly identical.

Table 2.1: QCD predictions and measurements of charm lifetime ratios. The errors of the lifetime ratios are calculated from the errors of the lifetimes [6] by the error propagation formula, that is, $\sigma_r^2/r^2 = \sigma_a^2/a^2 + \sigma_b^2/b^2$ in case of $r = a/b$. The f_D is the decay constant of D meson.

Observable	QCD expectations	Measurements
$\tau(D^+)/\tau(D^0)$	~ 2 (for $f_D \simeq 200$ MeV)	2.547 ± 0.036
$\tau(D_s^+)/\tau(D^0)$	1.00–1.07	$1.202^{+0.026}_{-0.023}$

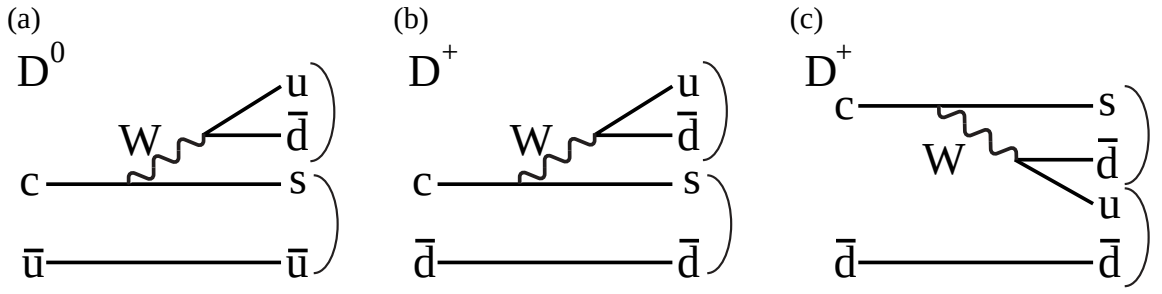


Figure 2.1: Diagrams of D^0 and D^+: Spectator decay of (a) D^0 and (b) D^+, and internal spectator decay (color suppressed) of (c) D^+.

2.1.1 Spectator Model

The spectator quark model describes charm hadron decays by treating the c quark as a free particle that decays via a tree-level charged weak current as shown in Figure 2.2. The light anti-quark ($\overline{u}, \overline{d}, \overline{s}$) in D^0, D^+ and D_s^+ mesons and the diquark in Λ_c baryons are called spectator quarks and are postulated to have a negligible role in the decay process. Therefore, all charm hadrons are expected to have the same lifetime.

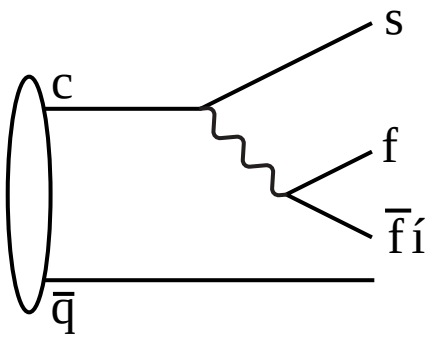


Figure 2.2: Spectator model.

In the spectator model, the total charm decay rate Γ_D has the same form as the muon decay rate $\Gamma(\mu \to e\nu_e\nu_\mu)$ except for a few modifications:

$$\begin{aligned}\Gamma_D &= N_W |V_{cs}|^2 \frac{G_F^2 m_c^5}{192\pi^3} \\ &= \left\{ N_W |V_{cs}|^2 \left(\frac{m_c}{m_\mu}\right)^5 \right\} \cdot \Gamma(\mu \to e\nu_e\nu_\mu),\end{aligned}$$

where G_F is the Fermi constant, m_c and m_μ are the c quark and μ masses, N_W is the number of final states available to the virtual W and V_{cs} is an element of the CKM matrix.

When we consider the color factor, N_W is 5 since the virtual W^+ decays to $u\overline{d}$, $e^+\nu_e$ and $\mu^+\nu_\mu$. Therefore we obtain $\Gamma(\mu \to e\nu_e\nu_\mu) \sim 3.0 \times 10^{-19}$ GeV and $\Gamma_D \sim 8.7 \times 10^{-13}$ GeV[a]. Then the μ and charm lifetimes are calculated to be 2.2 μs and 0.76 ps, respectively, in the spectator model. The former is consistent with the experimental result ($\tau(\mu) = 2.19703 \pm 0.00004$ μs [6]) and the latter is of the same order as the measured charm meson lifetimes.

2.1.2 Non-Spectator Process

It is possible that the light quark participates in the decay. We need to distinguish between charged and neutral mesons: for definiteness, we consider the decay of D^+ and D^0. In case of D^+, the c and the d quarks can annihilate into a virtual W boson which subsequently decays to a quark-antiquark pair or a lepton pair as shown in Figures 2.3. For D^+ this process is Cabibbo suppressed but for the D_s^+ the same process can occur without Cabibbo suppression.

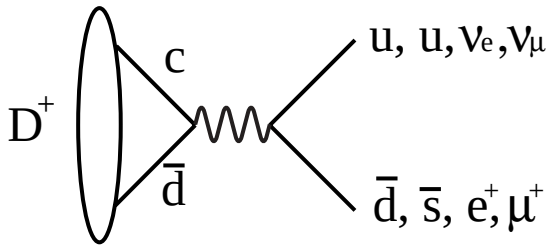


Figure 2.3: Annihilation process.

In the case of D^0, the annihilation of the c and the u quarks occurs by W boson exchange as shown in Figure 2.4. Leptonic final states are not possible.

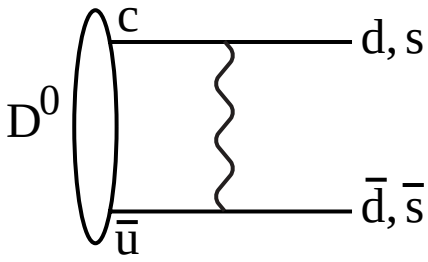


Figure 2.4: Exchange process.

[a]We use the following numbers: $G_F = 1.116 \times 10^{-5}$ GeV^{-2}, $m_\mu = 105.6$ MeV, $m_c = 1.5$ GeV [7], $V_{ud} = V_{cs} = 1$, $V_{us} = 0$.

The decay processes in which the light quark takes an active role are called non-spectator processes. Because the mesons are spinless these non-spectator contributions are suppressed by the square of the mass of the final state quarks or leptons. This is known as helicity suppression [10]. For instance, in case of D_s^+ the ratio of non-spectator (NS) to spectator (S) processes is calculated to be

$$\frac{\Gamma_{\rm NS}(D_s^+ \to \mu^+\nu_\mu)}{\Gamma_{\rm S}(D_s^+ \to \overline{s}s\mu^+\nu_\mu)} \sim 1\%.$$

Before the measurement of the charm meson lifetimes, the helicity suppression led one to completely neglect non-spectator interactions and to assume that the spectator model was sufficient to explain the D [11] meson decay.

2.1.3 Heavy Quark Expansion

After the measurement of the charm meson lifetimes showing that the D^+ lifetime is considerably longer than the D^0 lifetime, several models have been constructed, in which the contributions of the non-spectator processes are taken into account, to explain why the two lifetimes are different. One of them is the Heavy Quark Expansion [7] which is based on QCD and consists of an operator product expansion in the heavy quark mass.

The total decay width of the charm meson is given by

$$\Gamma(D) = \Gamma_{\rm NL} + \Gamma_{\rm SL} + \Gamma_{\rm PL},$$

where $\Gamma_{\rm NL}$, $\Gamma_{\rm SL}$ and $\Gamma_{\rm PL}$ denote nonleptonic (hadronic), semileptonic and pure leptonic decay widths, respectively. Compared to the total rate, the pure leptonic decay width is normally very small due to helicity suppression[b]. The theory predicts that the semileptonic decay widths of all charm mesons are nearly identical. This is supported by the current data [6]:

$$\begin{aligned}
\Gamma(D^0 \to e^+X) &= {\rm BR}(D^0 \to e^+X)/\tau(D^0) = 0.164 \pm 0.007 \text{ ps}^{-1} \\
\Gamma(D^+ \to e^+X) &= {\rm BR}(D^+ \to e^+X)/\tau(D^+) = 0.164 \pm 0.018 \text{ ps}^{-1} \\
\Gamma(D_s^+ \to e^+X) &= {\rm BR}(D_s^+ \to e^+X)/\tau(D_s^+) = 0.161^{+0.121}_{-0.101} \text{ ps}^{-1},
\end{aligned}$$

where ${\rm BR}(D^0 \to e^+X)$ represents the branching fraction of D^0 decaying e^+ plus anything, that is, inclusive decay. This fact indicates that lifetime differences are mainly due to the nonleptonic decays.

For the nonleptonic decays, the width of the heavy-flavor hadron H_Q decaying to the final state f is described by Eq. (2.1) in the Heavy Quark Expansion, treated to order $1/m_Q^3$. Detailed descriptions of the Heavy Quark Expansion for charm meson lifetimes can be found in [7, 12].

[b] ${\rm BR}(D^+ \to \mu^+\nu_\mu) = (8^{+17}_{-5})\times10^{-4}$, ${\rm BR}(D_s^+ \to \mu^+\nu_\mu) = (5.1 \pm 1.9)\times10^{-3}$, and ${\rm BR}(D_s^+ \to \tau^+\nu_\tau) = 7 \pm 4$ [6], where BR represents a branching fraction.

$$\Gamma(H_Q \to f) = \frac{G_{\rm F}^2 m_Q^5}{192\pi^3}|V_{\rm CKM}|^2[c_3^f \frac{\langle H_Q|\overline{Q}Q|H_Q\rangle}{2M_{H_Q}} + c_5^f \frac{\langle H_Q|\overline{Q}i\sigma\cdot GQ|H_Q\rangle}{2M_{H_Q}m_Q^2} + \sum_i c_{6,i}^f \frac{\langle H_Q|(\overline{Q}\Gamma_i q)(\overline{q}\Gamma_i Q)|H_Q\rangle}{2M_{H_Q}m_Q^3} + O(1/m_Q^4)] \quad (2.1)$$

Here, m_Q and M_{H_Q} are masses of the heavy quark Q and the heavy-flavor hadron, and Q between $\langle$ and $\rangle$ is the heavy quark field. The dimensionless coefficients c^f take into account QCD radiative corrections which are determined by perturbative QCD. The $G_{\rm F}$ is the Fermi constant and $V_{\rm CKM}$ is the appropriate combination of CKM matrix elements. The σ, G and Γ_i are the Pauli matrix, the chromomagnetic field operator and combinations of Dirac and color (=Gell–Mann) matrices.

Since $\langle H_Q|\overline{Q}Q|H_Q\rangle/2M_{H_Q}$ in the Eq. (2.1) can be written as[c]

$$\langle H_Q|\overline{Q}Q|H_Q\rangle_{\rm norm} = 1 + \frac{\langle H_Q|\overline{Q}(i/2)\sigma\cdot GQ|H_Q\rangle_{\rm norm}}{2m_Q^2} - \frac{\langle(\boldsymbol{p}_Q)^2\rangle_{H_Q}}{2m_Q^2} + O(1/m_Q^3),$$

$\langle H_Q|\overline{Q}Q|H_Q\rangle_{\rm norm}$ becomes 1 for $m_Q \to \infty$. In this limit all c hadrons have the same width, which is the spectator model.

The $\langle(\boldsymbol{p}_Q)^2\rangle_{H_Q}$ term dose not generate a significant difference among charm meson lifetimes since it is the square of the spatial momentum of the heavy quark Q moving in the soft gluon background and describes the kinetic energy of the heavy quark.

The $\langle H_Q|\overline{Q}i\sigma\cdot GQ|H_Q\rangle$ describes the interaction of the heavy quark spin σ with the color "magnetic" field G induced by light quarks inside the hadron. The expectation value for pseudoscalar mesons P_Q is given by the observed hyperfine splitting between the masses of the vector V_Q and pseudoscalar mesons:

$$\langle H_Q|\overline{Q}i\sigma\cdot GQ|H_Q\rangle_{\rm norm} \simeq \frac{3}{2}(M_{V_Q}^2 - M_{P_Q}^2).$$

Using the world average values for the D^0, D^+, D_s^+ and corresponding vector meson masses [6], we obtain

$$\frac{3}{2}(M_{D^{*0}}^2 - M_{D^0}^2) = 0.826\ ({\rm GeV}/c^2)^2$$
$$\frac{3}{2}(M_{D^{*+}}^2 - M_{D^+}^2) = 0.819\ ({\rm GeV}/c^2)^2$$
$$\frac{3}{2}(M_{D_s^{*+}}^2 - M_{D_s^+}^2) = 0.880\ ({\rm GeV}/c^2)^2.$$

Therefore the differences in the first and second terms of Eq. (2.1) are small.

The $\langle H_Q|(\overline{Q}\Gamma_i q)(\overline{q}\Gamma_i Q)|H_Q\rangle$ contains the contributions of Pauli Interference (PI) and Weak Annihilation (WA)[d]. This term contains the light quark $q = u, d, s$ and generates

[c] We use the definition $\langle H_Q|O_i|H_Q\rangle_{\rm norm} \equiv \langle H_Q|O_i|H_Q\rangle/2M_{H_Q}$.

[d] Weak Annihilation (WA) includes both the "annihilation" process of Figure 2.3 and the "exchange" process of Figure 2.4.

lifetime differences in the charm meson. Figure 2.5 shows PI and WA contributions for D^0, D^+ and D_s^+.

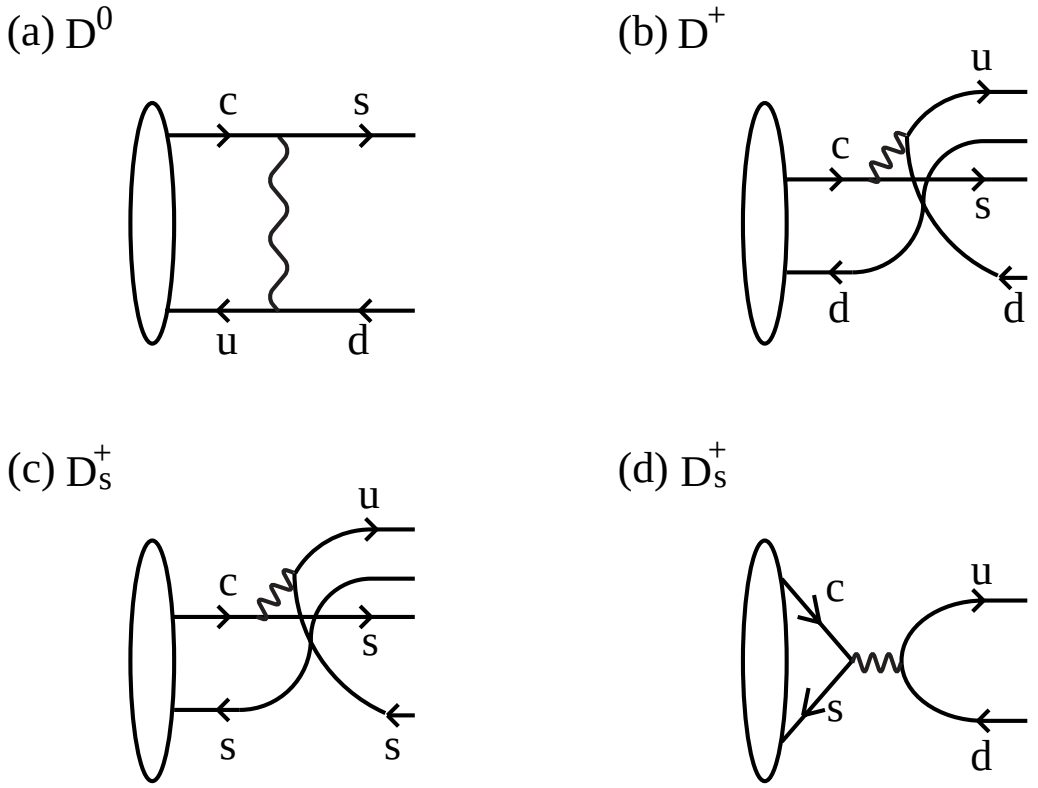


Figure 2.5: Nonspectator effects: (a) weak exchange, (b,c) Pauli interference, (d) weak annihilation.

The D^+ and D^0 lifetimes differ at the $1/m_c^3$ level due to both WA (Figure 2.5(a)) and PI (Figure 2.5(b)). Both affect the lifetime ratio in the same direction, that is, they enhance $\tau(D^+)$ over $\tau(D^0)$; destructive interference due to PI is the dominant effect. The ratio is calculated [7, 13] to be

$$\frac{\tau(D^+)}{\tau(D^0)} \simeq 1 + \left(\frac{f_D}{200\text{MeV}}\right)^2 \sim 2,$$

where f_D is the decay constant of D meson. The expectation does not coincide numerically with the measurement as shown in Table 2.1. This deviation could be ascribed to (a) uncertainty of $f_D \sim 200 \pm 30$ MeV, (b) $1/m_Q^4$ contributions which have been ignored and (c) the validity of the Heavy Quark Expansion for charm decays.

The Heavy Quark Expansion strongly suggests that the two lifetimes $\tau(D_s^+)$ and $\tau(D^0)$ are nearly identical. The ratio is estimated to be 1.00–1.07 from WA and PI at the $1/m_c^3$ level, i.e., including diagrams such as those shown in Figures 2.5(c) and 2.5(d)[e]. This

[e]There are other contributions such as the pure leptonic decay $D_s^+ \to \tau^+\nu_\tau$.

prediction is contradicted by the experimental measurement (Table 2.1). A possible reason could be the uncertainty of the size of the helicity suppression [14] since non-perturbative effects (e.g. soft gluon emission [15]) are able to decrease the effect of the helicity suppression.

2.2 D^0-$\overline{D}^0$ Mixing

"Mixing" is the phenomenon of neutral particle-antiparticle exchange. In D^0-$\overline{D}^0$ mixing, D^0 and $\overline{D}^0$ mesons evolve with time and mix into each other. In the Standard Model, mixing is described with a weak interaction involving two W bosons as shown in Figures 2.6 which is called a "box diagram".

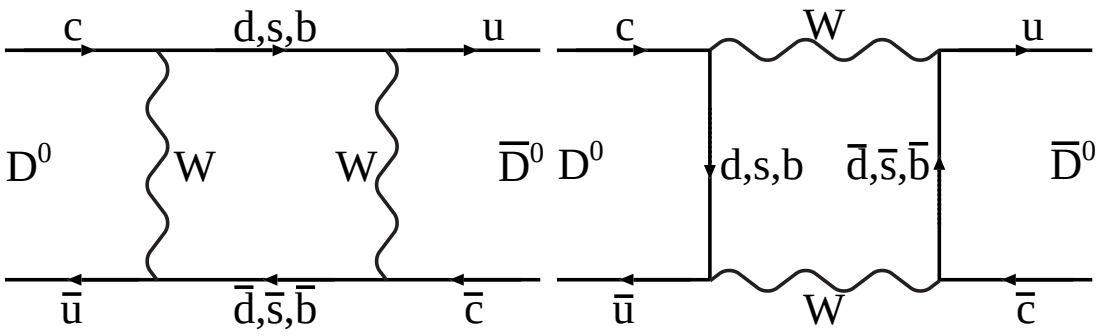


Figure 2.6: Box diagrams of D^0-$\overline{D}^0$ mixing.

First we describe the size of D^0-$\overline{D}^0$ mixing expected in the Standard Model. Next we show exact expressions for the time-dependent decay rates and then evaluate approximate formulas assuming that D^0-$\overline{D}^0$ mixing and CP violation are small but non-zero values in general.

2.2.1 Expectation in the Standard Model

The mixing amplitude calculated from the box diagrams is proportional to:

$$\sum_{i,j=d,s,b} V_{ci}^* V_{ui} V_{cj} V_{uj}^* S(m_i^2, m_j^2)$$

where V_{ij}s are the elements of the CKM matrix. If the quark masses m_i were all equal, the loop functions $S(m_i^2, m_j^2)$ would be equal and the amplitude would be zero due to the unitarity ($\sum_i V_{ci}^* V_{ui} = 0$) of the CKM matrix. This is known as GIM (Glashow-Iliopoulos-Maiani) mechanism. If the mass differences are small, the GIM mechanism is effective and mixing is small.

The analytic formulas for the box diagram contributions to Δm_{D^0} and $\Delta\Gamma_{D^0}$ are given in the approximate form [16]:

$$\Delta m_{D^0} = \frac{2}{3\pi^2} X_D \frac{(m_s^2 - m_d^2)^2}{m_c^2} \left[1 - \frac{5}{4}\frac{B'_D}{B_D}\frac{m_D^2}{(m_u + m_c)^2}\right]$$
$$\Delta\Gamma_{D^0} = \frac{4}{3\pi} X_D \frac{(m_s^2 - m_d^2)^2}{m_c^2}\frac{m_d^2 + m_s^2}{m_c^2} \left[1 - \frac{5}{2}\frac{B'_D}{B_D}\frac{m_D^2}{(m_u + m_c)^2}\right],$$

where the b quark contribution has been neglected, no QCD corrections are yet included, X_D is given by $X_D \equiv |V_{ud}V_{cd}^*|^2 B_D G_{\rm F}^2 m_D F_D^2$, $G_{\rm F}$ is the Fermi constant, m_u, m_d, m_s, m_c and m_D are masses of u, d, s, c quarks and D^0, F_D is the decay constant, and B_D and B'_D are bag parameters. From these formulas, $\Delta\Gamma_{D^0}$ is suppressed by a factor of $\frac{m_d^2+m_s^2}{m_c^2}$ comparing to Δm_{D^0}. The values of $\Delta m_{D^0}/\Gamma_{D^0}$ and $\Delta\Gamma_{D^0}/2\Gamma_{D^0}$ are calculated to be $\sim 10^{-5}$ and $\sim 10^{-6}$, respectively [16]. The D^0-$\overline{D}^0$ mixing [f] is so small that we cannot measure the D^0-$\overline{D}^0$ mixing in the current experiments unless there is physics beyond the Standard Model.

2.2.2 Basic Formulas: Phenomenology of Mixing Process

The two mass eigenstates, $|D_1\rangle$ of mass m_1 and width Γ_1 and $|D_2\rangle$ of mass m_2 and width Γ_2 are linear combinations of the interaction eigenstates [17, 18]:

$$|D_1\rangle = p|D^0\rangle + q|\overline{D}^0\rangle, \qquad |D_2\rangle = p|D^0\rangle - q|\overline{D}^0\rangle. \tag{2.2}$$

The coefficients are normalized as $|p|^2 + |q|^2 = 1$.

The two physical states evolve according to

$$|D_i(t)\rangle = e^{-im_i t - \frac{1}{2}\Gamma_i t}|D_i\rangle, \tag{2.3}$$

with the physical mass and width given by m_i and Γ_i $(i = 1, 2)$.

The differences of mass and width are parametrized by

$$x \equiv \frac{m_2 - m_1}{\Gamma}, \qquad y \equiv \frac{\Gamma_2 - \Gamma_1}{2\Gamma},$$

where $\Gamma \equiv \frac{\Gamma_1+\Gamma_2}{2}$.

Other useful parameters are defined as

$$\gamma = \frac{\Gamma_2 - \Gamma_1}{2}, \qquad \Delta_{p,q} = \frac{|p|^2 - |q|^2}{|p|^2 + |q|^2}.$$

Decay amplitudes for a final state f are defined by

$$A_f \equiv \langle f|H_d|D^0\rangle, \qquad \overline{A}_f \equiv \langle f|H_d|\overline{D}^0\rangle,$$
$$A_{\overline{f}} \equiv \langle \overline{f}|H_d|D^0\rangle, \qquad \overline{A}_{\overline{f}} \equiv \langle \overline{f}|H_d|\overline{D}^0\rangle.$$

[f] We call the box diagram contributions "short distance" contributions. There are also long distance (dispersive) contributions to the D^0-$\overline{D}^0$ mixing. They are considered to be larger than the short distance contributions but to be smaller than the current experimental limit [16].

The complex parameters λ_f and $\lambda_{\overline{f}}$ are defined by

$$\lambda_f \equiv \frac{q}{p}\frac{\overline{A}_f}{A_f}, \qquad \lambda_{\overline{f}} \equiv \frac{q}{p}\frac{\overline{A}_{\overline{f}}}{A_{\overline{f}}}.$$

For the observation of mixing, we are interested in the evolution of the state $|D^0(t)\rangle$ that starts out as a pure $|D^0\rangle$ at $t = 0$ and of the state $|\overline{D}^0(t)\rangle$ that is initially pure $|\overline{D}^0\rangle$. Solving Eqs. (2.2) and (2.3) we obtain

$$\begin{aligned}
|D^0(t)\rangle &= f_+(t)|D^0\rangle + \frac{q}{p}f_-(t)|\overline{D}^0\rangle, \\
|\overline{D}^0(t)\rangle &= \frac{p}{q}f_-(t)|D^0\rangle + f_+(t)|\overline{D}^0\rangle,
\end{aligned}$$

where

$$\begin{aligned}
f_+(t) &\equiv e^{-imt-\frac{1}{2}\Gamma t}\cos\left(\frac{1}{2}x\Gamma t - \frac{i}{2}y\Gamma t\right), \\
f_-(t) &\equiv e^{-imt-\frac{1}{2}\Gamma t} i\sin\left(\frac{1}{2}x\Gamma t - \frac{i}{2}y\Gamma t\right), \\
m &\equiv \frac{m_1 + m_2}{2}.
\end{aligned}$$

Then the decay amplitudes are written as

$$\begin{aligned}
\langle f|H_d|D^0(t)\rangle &= f_+(t)\langle f|H_d|D^0\rangle + \frac{q}{p}f_-(t)\langle f|H_d|\overline{D}^0\rangle \\
&= f_+(t)A_f + \frac{q}{p}f_-(t)\overline{A}_f, \\
\langle \overline{f}|H_d|\overline{D}^0(t)\rangle &= \frac{p}{q}f_-(t)A_{\overline{f}} + f_+(t)\overline{A}_{\overline{f}}, \\
\langle \overline{f}|H_d|D^0(t)\rangle &= f_+(t)A_{\overline{f}} + \frac{q}{p}f_-(t)\overline{A}_{\overline{f}}, \\
\langle f|H_d|\overline{D}^0(t)\rangle &= \frac{p}{q}f_-(t)A_f + f_+(t)\overline{A}_f.
\end{aligned} \tag{2.4}$$

Squaring these amplitudes gives the corresponding time-dependent decay rates:

$$\begin{aligned}
\Gamma[D^0(t) \to f] &= |\langle f|H_d|D^0(t)\rangle|^2 \\
&= |A_f|^2 e^{-\Gamma t}\left\{\frac{1+|\lambda_f|^2}{2}\cosh y\Gamma t + \Re(\lambda_f)\sinh y\Gamma t\right. \\
&\left. + \frac{1-|\lambda_f|^2}{2}\cos x\Gamma t - \Im(\lambda_f)\sin x\Gamma t\right\},
\end{aligned} \tag{2.5}$$

$$\begin{aligned}
\Gamma[\overline{D}^0(t) \to \overline{f}] &= |\langle \overline{f}|H_d|\overline{D}^0(t)\rangle|^2 \\
&= |\overline{A}_{\overline{f}}|^2 e^{-\Gamma t}\left\{\frac{1+|\lambda_{\overline{f}}^{-1}|^2}{2}\cosh y\Gamma t + \Re(\lambda_{\overline{f}}^{-1})\sinh y\Gamma t\right. \\
&\left. + \frac{1-|\lambda_{\overline{f}}^{-1}|^2}{2}\cos x\Gamma t - \Im(\lambda_{\overline{f}}^{-1})\sin x\Gamma t\right\},
\end{aligned} \tag{2.6}$$

for the "favored" modes and

$$\begin{aligned}
\Gamma[\overline{D}^0(t) \to f] &= |\langle f|H_d|\overline{D}^0(t)\rangle|^2 \\
&= |A_f|^2 \left|\frac{p}{q}\right|^2 e^{-\Gamma t}\left\{\frac{1+|\lambda_f|^2}{2}\cosh y\Gamma t + \Re(\lambda_f)\sinh y\Gamma t\right. \\
&\left. - \frac{1-|\lambda_f|^2}{2}\cos x\Gamma t + \Im(\lambda_f)\sin x\Gamma t\right\}, \qquad (2.7)\\
\Gamma[D^0(t) \to \overline{f}] &= |\langle \overline{f}|H_d|D^0(t)\rangle|^2 \\
&= |\overline{A}_{\overline{f}}| \left|\frac{q}{p}\right|^2 e^{-\Gamma t}\left\{\frac{1+|\lambda_{\overline{f}}^{-1}|^2}{2}\cosh y\Gamma t + \Re(\lambda_{\overline{f}}^{-1})\sinh y\Gamma t\right. \\
&\left. - \frac{1-|\lambda_{\overline{f}}^{-1}|^2}{2}\cos x\Gamma t + \Im(\lambda_{\overline{f}}^{-1})\sin x\Gamma t\right\}, \qquad (2.8)
\end{aligned}$$

for the "suppressed" modes using

$$\begin{aligned}
|f_+(t)|^2 &= \frac{e^{-\Gamma t}}{2}(\cosh y\Gamma t + \cos x\Gamma t),\\
|f_-(t)|^2 &= \frac{e^{-\Gamma t}}{2}(\cosh y\Gamma t - \cos x\Gamma t),\\
f_+(t)f_-(t)^* &= \frac{e^{-\Gamma t}}{2}(\sinh y\Gamma t - i\sin x\Gamma t).
\end{aligned}$$

Assuming the equal numbers of D^0 and $\overline{D}^0$ are produced, the flavor-untagged time-dependent decay rates are

$$\begin{aligned}
\Gamma[D^0(t),\overline{D}^0(t) \to f] &= \Gamma[D^0(t) \to f] + \Gamma[\overline{D}^0(t) \to f] \\
&= |A_f|^2\left(1+\left|\frac{p}{q}\right|^2\right)e^{-\Gamma t}\left\{\frac{1+|\lambda_f|^2}{2}\cosh y\Gamma t + \Re(\lambda_f)\sinh y\Gamma t\right. \\
&\left. - \Delta_{p,q}\left(\frac{1-|\lambda_f|^2}{2}\cos x\Gamma t - \Im(\lambda_f)\sin x\Gamma t\right)\right\},\\
\Gamma[D^0(t),\overline{D}^0(t) \to \overline{f}] &= \Gamma[D^0(t) \to \overline{f}] + \Gamma[\overline{D}^0(t) \to \overline{f}] \\
&= |\overline{A}_{\overline{f}}|^2\left(1+\left|\frac{q}{p}\right|^2\right)e^{-\Gamma t}\left\{\frac{1+|\lambda_{\overline{f}}^{-1}|^2}{2}\cosh y\Gamma t + \Re(\lambda_{\overline{f}}^{-1})\sinh y\Gamma t\right. \\
&\left. + \Delta_{p,q}\left(\frac{1-|\lambda_{\overline{f}}^{-1}|^2}{2}\cos x\Gamma t - \Im(\lambda_{\overline{f}}^{-1})\sin x\Gamma t\right)\right\},
\end{aligned}$$

where "flavor-untagged" means that we do not distinguish the initial state, that is, D^0 or $\overline{D}^0$ decay to the same final state f. These expressions for time-dependent decay rates are exact in the sense that they do not use the approximations that x, y, or $\Delta_{p,q}$ are small.

Since we are interested in the slope at $t \sim 0$, we express the time-dependent decay rates keeping terms only to the first order in xt and yt for the "favored" modes and to

the second order in xt and yt for the "suppressed" modes:

$$\begin{aligned}
\Gamma[D^0(t) \to f] &= |A_f|^2 e^{-\Gamma t}\{1 + \Re(\lambda_f) y\Gamma t - \Im(\lambda_f) x\Gamma t\} \\
\Gamma[\overline{D}^0(t) \to \overline{f}] &= |\overline{A}_{\overline{f}}|^2 e^{-\Gamma t}\{1 + \Re(\lambda_{\overline{f}}^{-1}) y\Gamma t - \Im(\lambda_{\overline{f}}^{-1}) x\Gamma t\} \\
\Gamma[\overline{D}^0(t) \to f] &= |A_f|^2 \left|\frac{p}{q}\right|^2 e^{-\Gamma t}\Bigg\{ |\lambda_f|^2 + \Re(\lambda_f) y\Gamma t + \Im(\lambda_f) x\Gamma t \\
&\quad + \frac{1}{4}(x^2 + y^2)\Gamma^2 t^2 - \frac{|\lambda_f|^2}{4}(x^2 - y^2)\Gamma^2 t^2 \Bigg\}, \qquad (2.9) \\
\Gamma[D^0(t) \to \overline{f}] &= |\overline{A}_{\overline{f}}|^2 \left|\frac{q}{p}\right|^2 e^{-\Gamma t}\Bigg\{ |\lambda_{\overline{f}}^{-1}|^2 + \Re(\lambda_{\overline{f}}^{-1}) y\Gamma t + \Im(\lambda_{\overline{f}}^{-1}) x\Gamma t \\
&\quad + \frac{1}{4}(x^2 + y^2)\Gamma^2 t^2 - \frac{|\lambda_{\overline{f}}^{-1}|^2}{4}(x^2 - y^2)\Gamma^2 t^2 \Bigg\}.
\end{aligned}$$

By adding $D^0(t)$ and $\overline{D}^0(t)$ amplitudes, we obtain the flavor-untagged time-dependent decay rates

$$\begin{aligned}
\Gamma[D^0(t), \overline{D}^0(t) \to f] &= |A_f|^2 \left(1 + \left|\frac{p}{q}\right|^2\right) e^{-\Gamma t} \\
&\quad \cdot \Bigg\{ \frac{1 + |\lambda_f|^2}{2} - \Delta_{p,q} \frac{1 - |\lambda_f|^2}{2} \\
&\quad + \left(\Re(\lambda_f) y + \Delta_{p,q} \Im(\lambda_f) x\right) \Gamma t \Bigg\}, \qquad (2.10) \\
\Gamma[D^0(t), \overline{D}^0(t) \to \overline{f}] &= |\overline{A}_{\overline{f}}|^2 \left(1 + \left|\frac{q}{p}\right|^2\right) e^{-\Gamma t} \\
&\quad \cdot \Bigg\{ \frac{1 + |\lambda_{\overline{f}}^{-1}|^2}{2} + \Delta_{p,q} \frac{1 - |\lambda_{\overline{f}}^{-1}|^2}{2} \\
&\quad + \left(\Re(\lambda_{\overline{f}}^{-1}) y - \Delta_{p,q} \Im(\lambda_{\overline{f}}^{-1}) x\right) \Gamma t \Bigg\}, \qquad (2.11)
\end{aligned}$$

where terms only to the first order in xt and yt are retained.

2.2.3 Time-Dependent Decay Rates of $D^0 \to K^-\pi^+, K^+K^-$ and $K^+\pi^-$

We consider the Cabibbo-favored $D^0 \to K^-\pi^+$, the singly-Cabibbo-suppressed $D^0 \to K^+K^-$, and the doubly-Cabibbo-suppressed $D^0 \to K^+\pi^-$ decays as shown in Figure 2.7.

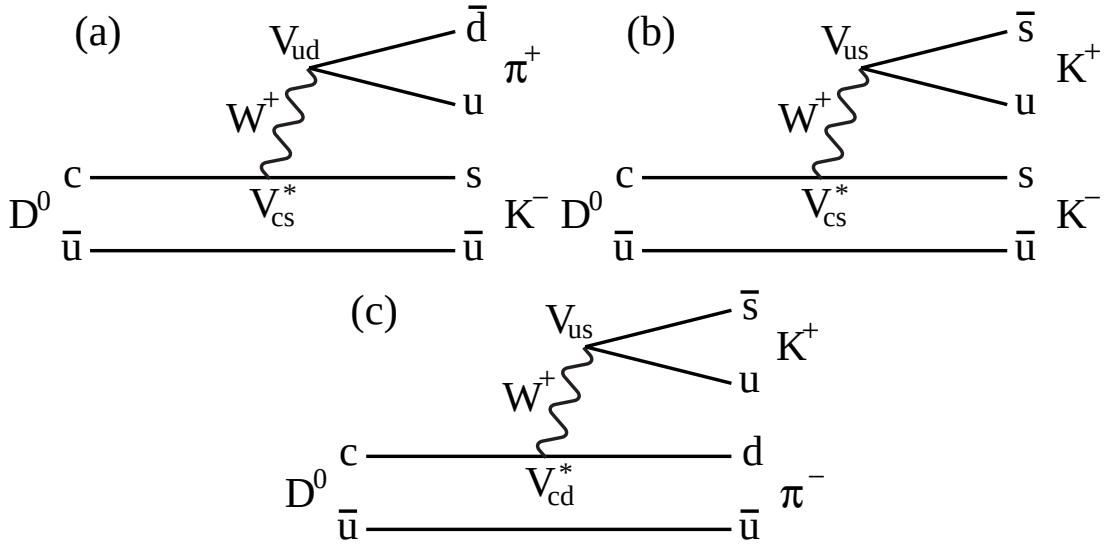


Figure 2.7: Diagrams of (a) $D^0 \to K^-\pi^+$, (b) $D^0 \to K^+K^-$ and (c) $D^0 \to K^+\pi^-$.

The $D^0 \to K^+\pi^-$ decay process can proceed either through direct doubly-Cabibbo-suppressed decay, or through state-mixing followed by the Cabibbo-favored decay, $D^0 \to \overline{D}^0 \to K^+\pi^-$,as shown in Figure 2.8. Both processes are included in the decay amplitudes as expressed by Eqs. (2.4). The $D^0 \to K^-\pi^+$ decay and the singly-Cabbibo-suppressed decay $D^0 \to K^+K^-$ proceed via two paths in the same way.

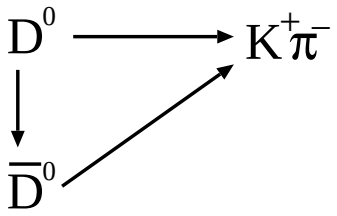


Figure 2.8: Two paths of decay mode $D^0 \to K^+\pi^-$. One is a doubly-Cabibbo-suppressed decay and the other is through state-mixing followed by the Cabibbo-favored decay.

From the experimental information we know that

$$|\lambda_{K^-\pi^+}| = \left|\frac{q}{p}\frac{\langle K^-\pi^+|H_d|\overline{D}^0\rangle}{\langle K^-\pi^+|H_d|D^0\rangle}\right| \ll 1,$$

$$|\lambda_{K^+\pi^-}| = \left|\frac{q}{p}\frac{\langle K^+\pi^-|H_d|\overline{D}^0\rangle}{\langle K^+\pi^-|H_d|D^0\rangle}\right| \gg 1,$$

$$|\lambda_{K^+K^-}| = \left|\frac{q}{p}\frac{\langle K^+K^-|H_d|\overline{D}^0\rangle}{\langle K^+K^-|H_d|D^0\rangle}\right| \sim 1.$$

Then from Eqs. (2.9) we write down approximate expressions for the time-dependent decay rates:

$$\begin{aligned}
\Gamma[D^0(t) \to K^-\pi^+] &= |A_{K^-\pi^+}|^2 e^{-\Gamma t}, \\
\Gamma[\overline{D}^0(t) \to K^+\pi^-] &= |\overline{A}_{K^+\pi^-}|^2 e^{-\Gamma t}, \\
\Gamma[D^0(t) \to K^-K^+] &= |A_{K^-K^+}|^2 e^{-\Gamma t}\{1 + \Re(\lambda_{K^-K^+})y\Gamma t - \Im(\lambda_{K^-K^+})x\Gamma t\}, \\
\Gamma[\overline{D}^0(t) \to K^-K^+] &= |\overline{A}_{K^-K^+}|^2 e^{-\Gamma t}\{1 + \Re(\lambda^{-1}_{K^-K^+})y\Gamma t - \Im(\lambda^{-1}_{K^-K^+})x\Gamma t\}, \\
\Gamma[D^0(t) \to K^+\pi^-] &= |\overline{A}_{K^+\pi^-}|^2 \left|\frac{q}{p}\right|^2 e^{-\Gamma t} \\
&\quad \cdot\left\{|\lambda^{-1}_{K^+\pi^-}|^2 + \Re(\lambda^{-1}_{K^+\pi^-})y\Gamma t + \Im(\lambda^{-1}_{K^+\pi^-})x\Gamma t + \frac{1}{4}(x^2+y^2)\Gamma^2 t^2\right\}, \\
\Gamma[\overline{D}^0(t) \to K^-\pi^+] &= |A_{K^-\pi^+}|^2 \left|\frac{p}{q}\right|^2 e^{-\Gamma t} \\
&\quad \cdot\left\{|\lambda_{K^-\pi^+}|^2 + \Re(\lambda_{K^-\pi^+})y\Gamma t + \Im(\lambda_{K^-\pi^+})x\Gamma t + \frac{1}{4}(x^2+y^2)\Gamma^2 t^2\right\}.
\end{aligned} \tag{2.12}$$

When we assume that there is no direct CP violation, the effects of indirect CP violation can be parametrized in the following way:

$$\begin{aligned}
|q/p| &\equiv R_{\rm m}, \\
\lambda^{-1}_{K^+\pi^-} &\equiv \sqrt{R}R_{\rm m}^{-1}e^{-i(\delta+\phi)}, \\
\lambda_{K^-\pi^+} &\equiv \sqrt{R}R_{\rm m}e^{-i(\delta-\phi)}, \\
\lambda_{K^+K^-} &\equiv -R_{\rm m}e^{i\phi}.
\end{aligned} \tag{2.13}$$

Here R and $R_{\rm m}$ are real and positive dimensionless numbers. ϕ and δ are the weak and strong phase differences between the doubly-Cabibbo-suppressed decay $D^0 \to K^+\pi^-$ and the Cabibbo-flavored decay $\overline{D}^0 \to K^+\pi^-$, respectively. CP violation in mixing is related to $R_{\rm m} \neq 1$ while CP violation in the interference of decays with and without mixing is related to $\sin\phi \neq 0$. The choice of phases and signs in Eq. (2.13) is consistent with having $\phi = 0$ in the Standard Model and $\delta = 0$ in the SU(3) flavor symmetry limit [19].

Using Eqs (2.13) we rewrite Eq. (2.12) as

$$\begin{aligned}
\Gamma[D^0(t) \to K^-\pi^+] &= \Gamma[\overline{D}^0(t) \to K^+\pi^-] \\
&= |A_{K^-\pi^+}|^2 e^{-\Gamma t}, \\
\Gamma[D^0(t) \to K^+K^-] &= |A_{K^+K^-}|^2 e^{-\Gamma t}[1 - R_{\rm m}(y\cos\phi - x\sin\phi)\Gamma t], \\
\Gamma[\overline{D}^0(t) \to K^+K^-] &= |A_{K^+K^-}|^2 e^{-\Gamma t}[1 - R_{\rm m}^{-1}(y\cos\phi + x\sin\phi)\Gamma t], \\
\Gamma[D^0(t) \to K^+\pi^-] &= |A_{K^-\pi^+}|^2 e^{-\Gamma t} \\
&\quad \cdot [R + \sqrt{R}R_{\rm m}(y'\cos\phi - x'\sin\phi)\Gamma t + \frac{R_{\rm m}^2}{4}(x^2+y^2)(\Gamma t)^2], \\
\Gamma[\overline{D}^0(t) \to K^-\pi^+] &= |A_{K^-\pi^+}|^2 e^{-\Gamma t} \\
&\quad \cdot [R + \sqrt{R}R_{\rm m}^{-1}(y'\cos\phi + x'\sin\phi)\Gamma t + \frac{R_{\rm m}^{-2}}{4}(x^2+y^2)(\Gamma t)^2],
\end{aligned} \tag{2.14}$$

where x' and y' are defined as

$$\begin{aligned} x' &\equiv x\cos\delta + y\sin\delta, \\ y' &\equiv y\cos\delta - x\sin\delta. \end{aligned} \tag{2.15}$$

Using Eq. (2.14) we express the flavor-untagged decay rate for $D^0 \to K^-K^+$ as

$$\begin{aligned} \Gamma[D^0(t),\overline{D}^0(t) \to K^-K^+] &= \Gamma[D^0(t) \to K^-K^+] + \Gamma[\overline{D}^0(t) \to K^-K^+] \\ &= 2|A_{K^-K^+}|^2 e^{-\Gamma t} \\ &\cdot \left[1 - \left\{ \frac{R_{\rm m} + R_{\rm m}^{-1}}{2} y\cos\phi - \frac{R_{\rm m} - R_{\rm m}^{-1}}{2} x\sin\phi \right\} \Gamma t \right]. \end{aligned} \tag{2.16}$$

Using Eqs. (2.10) and (2.11) to keep terms only to the first order in xt and yt we express the flavor-untagged decay rates for $D^0 \to K^-\pi^+$ as:

$$\begin{aligned} \Gamma[D^0(t),\overline{D}^0(t) \to K^-\pi^+] &= |A_{K^-\pi^+}|^2 \left(1 + \left|\frac{p}{q}\right|^2\right) \frac{1 + |\lambda_{K^-\pi^+}|^2 - \Delta_{p,q}(1 - |\lambda_{K^-\pi^+}|^2)}{2} e^{-\Gamma t} \\ &\cdot \left\{ 1 + \frac{2\left(\Re(\lambda_{K^-\pi^+})y + \Delta_{p,q}\Im(\lambda_{K^-\pi^+})x\right)}{1 + |\lambda_{K^-\pi^+}|^2 - \Delta_{p,q}(1 - |\lambda_{K^-\pi^+}|^2)} \Gamma t \right\}, \end{aligned} \tag{2.17}$$

$$\begin{aligned} \Gamma[D^0(t),\overline{D}^0(t) \to K^+\pi^-] &= |\overline{A}_{K^+\pi^-}|^2 \left(1 + \left|\frac{q}{p}\right|^2\right) \frac{1 + |\lambda^{-1}_{K^+\pi^-}|^2 + \Delta_{p,q}(1 - |\lambda^{-1}_{K^+\pi^-}|^2)}{2} e^{-\Gamma t} \\ &\cdot \left\{ 1 + \frac{2\left(\Re(\lambda^{-1}_{K^+\pi^-})y - \Delta_{p,q}\Im(\lambda^{-1}_{K^+\pi^-})x\right)}{1 + |\lambda_{K^+\pi^-}|^2 + \Delta_{p,q}(1 - |\lambda^{-1}_{K^+\pi^-}|^2)} \Gamma t \right\}. \end{aligned}$$

2.3 Relation between Theory and Experiment

2.3.1 Effective Decay Rates

In general the lifetime is unique for each particle. However we often refer to a "lifetime" for each decay mode such as a lifetime of $D^0 \to K^-\pi^+$, a lifetime of $D^0 \to K^-K^+$. This "lifetime" is defined to be the inverse of the effective partial decay rate. The effective decay rate is obtained by fitting a single exponential to the measured decay rate distribution for each decay mode.

Since there are two separate exponentials in the decay rates, we suppose that we have a function

$$f(t) = e^{-\Gamma_1 t} + e^{-\Gamma_2 t}.$$

The relative slope at $t = 0$ $f'(0)/f(0)$ is the same as that of another function given by

$$g(t) = 2e^{-\frac{\Gamma_1+\Gamma_2}{2}t} = 2e^{-\Gamma t}.$$

These two functions are quite similar when $\Gamma_1 \sim \Gamma_2$ and γt are small,

$$\frac{f(t)}{g(t)} = \cosh\gamma t \sim 1 + \frac{(\gamma t)^2}{2}.$$

For $\gamma t < 0.1$ such as $y = 0.01$ and less than 10 lifetimes, this deviation is less than 0.005 and can be ignored.

Moreover a toy Monte Carlo study [18] has shown that when a single exponential is fit to the function $f(t)$, the slope is close to the average of Γ_1 and Γ_2, but it is always smaller by $\sim y^2$ relatively. When we measure y by comparing Γ_1 with the effective decay rate of $f(t)$, the measured value of y is shifted by a fraction y of y itself. As the experimental limit on y becomes smaller, the bias caused by fitting a single exponential would become smaller.

2.3.2 Lifetime Measurement using $D^0 \to K^-\pi^+$

The average decay rate Γ is often measured as the effective decay rate of the $K^{\mp}\pi^{\pm}$ final states. The parameters $\lambda_{K^-\pi^+}$ and $\lambda^{-1}_{K^+\pi^-}$ are both of order 0.06 [20].

Then we write Eqs. (2.17) in exponential form,

$$\begin{aligned}
\Gamma[D^0(t), \overline{D}^0(t) \to K^-\pi^+] &\propto e^{-(1+y_{K^-\pi^+})\Gamma t}, \\
y_{K^-\pi^+} &\equiv -\frac{2\left(\Re(\lambda_{K^-\pi^+})y + \Delta_{p,q}\Im(\lambda_{K^-\pi^+})x\right)}{1 + |\lambda_{K^-\pi^+}|^2 - \Delta_{p,q}(1 - |\lambda_{K^-\pi^+}|^2)} \\
&\approx -\frac{2\left(\Re(\lambda_{K^-\pi^+})y + \Delta_{p,q}\Im(\lambda_{K^-\pi^+})x\right)}{1 - \Delta_{p,q}}, \\
\Gamma[D^0(t), \overline{D}^0(t) \to K^+\pi^-] &\propto e^{-(1+y_{K^+\pi^-})\Gamma t}, \\
y_{K^+\pi^-} &\equiv -\frac{2\left(\Re(\lambda^{-1}_{K^+\pi^-})y - \Delta_{p,q}\Im(\lambda^{-1}_{K^+\pi^-})x\right)}{1 + |\lambda_{K^+\pi^-}|^2 + \Delta_{p,q}(1 - |\lambda^{-1}_{K^+\pi^-}|^2)} \\
&\approx -\frac{2\left(\Re(\lambda^{-1}_{K^+\pi^-})y - \Delta_{p,q}\Im(\lambda^{-1}_{K^+\pi^-})x\right)}{1 + \Delta_{p,q}}.
\end{aligned}$$

The parameters $y_{K^-\pi^+}$ and $y_{K^+\pi^-}$ are the fractional shifts of the effective decay rates from Γ. If $\Delta_{p,q}$ is $\sim$0, they are roughly

$$y_{K^-\pi^+} \sim -2\Re(\lambda_{K^-\pi^+})y, \quad y_{K^+\pi^-} \sim -2\Re(\lambda^{-1}_{K^+\pi^-})y.$$

We measure the effective decay rate of the sum of the $K^-\pi^+$ and $K^+\pi^-$. Since there is no automatic cancellation between $y_{K^-\pi^+}$ and $y_{K^+\pi^-}$, the measured lifetime is different from the inverse of the average decay rate Γ. Since $\Re(\lambda_{K^-\pi^+})$ and $\Re(\lambda^{-1}_{K^+\pi^-})$ are of order 0.06, we express the shift value $y_{K^{\mp}\pi^{\pm}}$ as

$$y_{K^{\mp}\pi^{\pm}} \equiv \delta_\lambda \cdot y.$$

Using the effective decay rate $\Gamma_{K^{\mp}\pi^{\pm}}$ the measured lifetime $\tau_{K^{\mp}\pi^{\pm}}$ is expressed as

$$\tau_{K^{\mp}\pi^{\pm}} = \frac{1}{\Gamma_{K^{\mp}\pi^{\pm}}} = \frac{1}{(1 + \delta_\lambda \cdot y)\Gamma}.$$

2.3.3 Definition and Measurement of y_{CP} Parameter

We write Eq. (2.16) in exponential form,

$$\Gamma[D^0(t), \overline{D}^0(t) \to K^-K^+] \propto e^{-(1+y_{K^-K^+})\Gamma t},$$
$$y_{K^-K^+} \equiv \frac{R_{\rm m} + R_{\rm m}^{-1}}{2} y\cos\phi - \frac{R_{\rm m} - R_{\rm m}^{-1}}{2} x\sin\phi.$$

Since the Standard Model suggests that $R_{\rm m}$ is not very different from 1, we define $A_{\rm m}$ as $R_{\rm m}^{\pm 2} \equiv 1 \pm A_{\rm m}$. Then $y_{K^-K^+}$ is expressed as

$$y_{K^-K^+} \sim y\cos\phi - \frac{A_{\rm m}}{2} x\sin\phi.$$

Using the effective decay rate $\Gamma_{K^-K^+}$ the measured lifetime $\tau_{K^-K^+}$ is expressed as

$$\tau_{K^-K^+} = \frac{1}{\Gamma_{K^-K^+}} = \frac{1}{(1+y_{K^-K^+})\Gamma},$$

and then $y_{K^-K^+}$ is

$$y_{K^-K^+} = \frac{\Gamma_{K^-K^+}}{\Gamma} - 1.$$

We define the parameter y_{CP} as

$$\begin{aligned} y_{CP} &\equiv \frac{\Gamma_{K^-K^+}}{\Gamma_{K^\mp\pi^\pm}} - 1 \\ &= \frac{\tau_{K^\mp\pi^\pm}}{\tau_{K^-K^+}} - 1 \\ &= \frac{1+y_{K^-K^+}}{1+\delta_\lambda \cdot y} - 1 \\ &\sim y_{K^-K^+} - \delta_\lambda \cdot y \\ &\sim y\cos\phi - \frac{A_{\rm m}}{2} x\sin\phi - \delta_\lambda \cdot y. \end{aligned} \tag{2.18}$$

We ignore a $\delta_\lambda \cdot y$ term in the rest of this thesis [g]. The y_{CP} is then expressed as

$$y_{CP} \sim y\cos\phi - \frac{A_{\rm m}}{2} x\sin\phi. \tag{2.19}$$

If we assume that a $\Gamma_{K^\mp\pi^\pm}$ is the average of CP-odd state and CP-even state decay rates, we can express y_{CP} as

$$\begin{aligned} y_{CP} &= \frac{\Gamma_{K^-K^+} - \Gamma_{K^\mp\pi^\pm}}{\Gamma_{K^\mp\pi^\pm}} \\ &\sim \frac{\Gamma_{CP\rm even} - (\Gamma_{CP\rm even} + \Gamma_{CP\rm odd})/2}{(\Gamma_{CP\rm even} + \Gamma_{CP\rm odd})/2} \\ &= \frac{\Gamma_{CP\rm even} - \Gamma_{CP\rm odd}}{\Gamma_{CP\rm even} + \Gamma_{CP\rm odd}}, \end{aligned}$$

where $\Gamma_{CP\rm even}$ and $\Gamma_{CP\rm odd}$ are decay rates of the CP even and CP odd eigenstates, respectively. This expression shows the origin of the name of the y_{CP} parameter, that is, the lifetime difference of the CP eigenstates.

We measure y_{CP} using the lifetime ratio between $D^0 \to K^-\pi^+$ and $D^0 \to K^-K^+$.

[g] Since the current experimental values of y_{CP} and these uncertainties are order of 0.01 [8, 21, 22, 23], $\delta_\lambda \cdot y$ is calculated to be $\sim 0.1 \times 0.01 = 0.001$ which is smaller than the experimental uncertainty.

2.3.4 Experimental Status

The mixing parameter y_{CP} has been measured to be

- $+0.008 \pm 0.029 \pm 0.010$ (E791) [21],
- $+0.0342 \pm 0.0139 \pm 0.0074$ (FOCUS) [8],
- $-0.011 \pm 0.025 \pm 0.014$ (CLEO) [22], and
- $-0.010 \pm 0.022 \pm 0.017$ (BABAR) [23].

CLEO has also measured the D^0-$\overline{D}^0$ mixing contribution to the $D^0 \to K^+\pi^-$ decay, finding $x' = 0.000 \pm 0.015 \pm 0.002$, $y' = -0.025^{+0.014}_{-0.016} \pm 0.003$ and $A_{\rm m} = 0.23^{+0.63}_{-0.80} \pm 0.01$ [24]. The relations among $y_{CP}, x', y', A_{\rm m}$ are represented by Eq. (2.15) and (2.19).

The strong phase difference δ vanishes in the SU(3) flavor symmetry limit. However when we examine the consistency of the FOCUS (y_{CP}) and CLEO ($x', y', A_{\rm m}$) results, the following relation is obtained.

$$\cos\delta - (x/y)\sin\delta = -0.73 \pm 0.55. \tag{2.20}$$

The difference in strong phases between the $D^0 \to K^+\pi^-$ and $D^0 \to K^-\pi^+$ decays is large if the mixing parameters are within one sigma of the FOCUS (y_{CP}) and CLEO ($x', y', A_{\rm m}$) measurements. From Eq. (2.20), the upper limits of $\cos\delta$ become

$$\cos\delta \lesssim \begin{cases} +0.61 & |x| \sim |y| \\ -0.18 & |x| \ll |y|, \end{cases}$$

implying a large SU(3) flavor symmetry breaking effect in the strong phase.

Moreover if CP violation does not occur or is negligible ($\phi \sim 0$) we obtain

$$y_{CP} \simeq y.$$

In the case of small or vanishing CP violation, the FOCUS result indicates a 2.2σ signal in y. Most theoretical estimates [25] of x and y are well below the one percent level ($x, y \sim O(10^{-3})$). There are several models [16] beyond the Standard Model in which the mixing rate is enhanced. Figure 2.9 shows additional D^0-$\overline{D}^0$ mixing box and tree diagrams: supersymmetry, a fourth quark generation, leptoquarks and so on. However, since all these models enhance x but not y, a large value for y is unexpected and could be interpreted as a failure of the Standard Model prediction[h] [26]. It is therefore very important to measure y_{CP} precisely to check the FOCUS result and to see if new physics is present in the D^0-$\overline{D}^0$ system.

[h] Since limits on the mixing parameter x are weak ($|x| < 0.03 \sim 0.06$), and the asymmetry parameter $A_{\rm m}$ is not constrained by existing measurements [24], a large y_{CP} could be accommodated if CP violating effects were large (see Eq. 2.19).

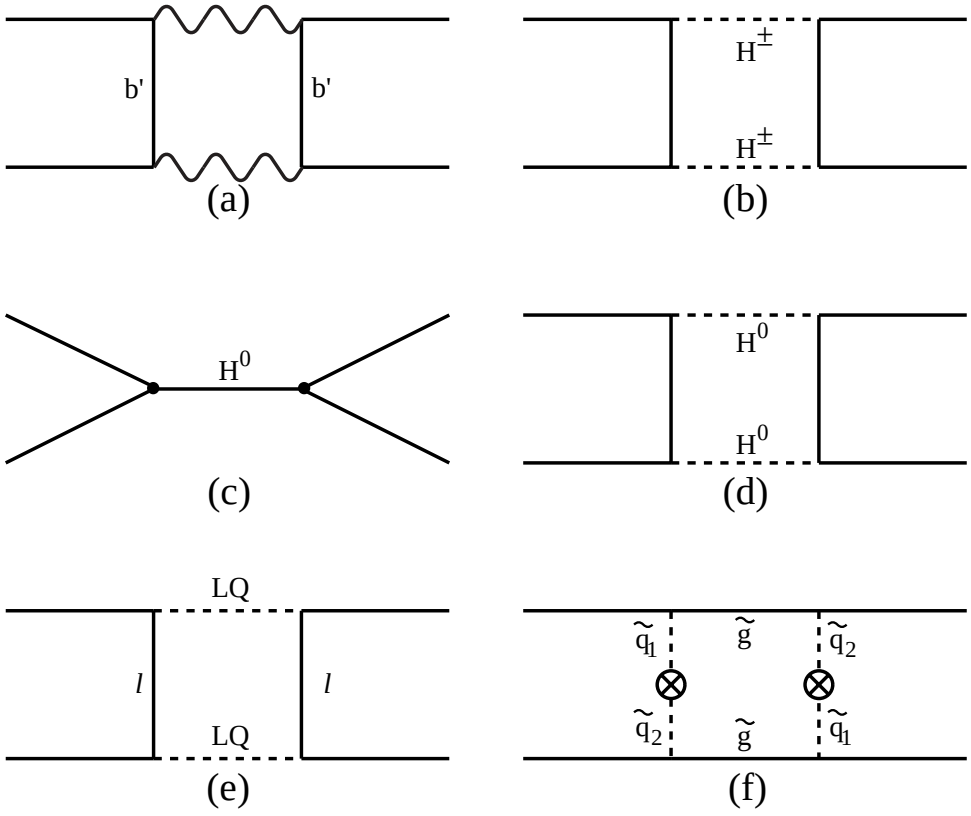


Figure 2.9: Effects of new physics on D^0-$\overline{D}^0$ mixing: (a) extra $Q = -1/3$ quark b', (b) charged higgs scalars $H^{\pm}$, (c,d) tree, box contributions of flavor-changing neutral higgs scalars H^0, (e) leptoquark LQ, (f) gluino $\tilde{g}$ and squarks $\tilde{q}_{1,2}$.

Chapter 3

Experimental Apparatus

The Belle detector is primarily designed to observe and measure CP violation [2] in B decay. In Section 3.1, the KEKB accelerator is described briefly and in Section 3.2, the overview of the Belle detector and the descriptions of its principal components are given. We describe the analysis tools in Section 3.3.

3.1 KEKB Accelerator

The configuration of the KEKB accelerator [27] is shown in Figure 3.1. KEKB is an asymmetric energy electron-positron collider designed to boost B mesons. The electron and positron beam energies are 8 GeV and 3.5 GeV respectively. The resulting energy in the center-of-mass system (cms), 10.58 GeV, corresponds to the mass of the $\Upsilon(4S)$ resonance. The Lorentz boost factor $\beta\gamma \simeq 0.425$ corresponds to the average distance of the two B meson decay vertices of approximately 200 μm. The design luminosity is 10^{34} cm^{-2}s^{-1} which corresponds to 10^8 $\Upsilon(4S)$ produced per year. The continuum events ($q\overline{q} = u\overline{u}, d\overline{d}, s\overline{s}, c\overline{c}$) are also produced, in the ratio $q\overline{q}$:$\Upsilon(4S) \simeq 2.5$:1 [28].

Electron and positron rings are built side by side in the existing TRISTAN [29] tunnel, which has a circumference of $\sim$3 km. KEKB has only one interaction point in the Tsukuba experimental hall, where the electron and positron beams collide at a finite angle of 22 mrad to reduce parasitic collisions. The Belle detector is installed around the interaction region.

3.2 Belle Detector

The structure of the detectors making up the Belle detector [30, 31] is shown in Figure 3.2. Because of the asymmetry of the beam energies, the detector itself is made asymmetric as shown in Figure 3.3. It has a larger acceptance toward the direction of the electron beam (the forward direction). The definition of the xyz-axes is shown in Figure 3.4 and the z-axis is parallel with the beampipe and its positive direction is the direction of the electron beam. We use two angles ϕ and θ, which are the azimuth angle around z-axis

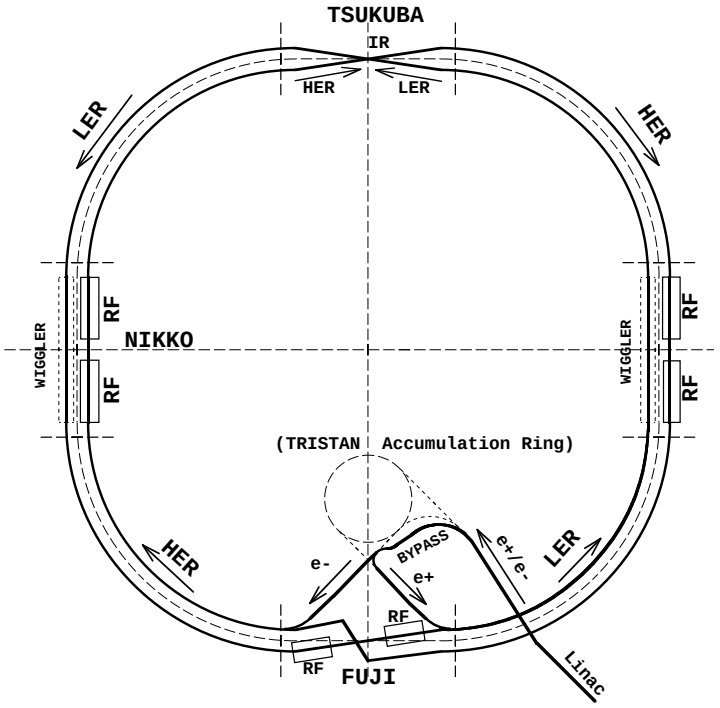


Figure 3.1: KEKB beamlines. The interaction region (IR) is in the Tsukuba experimental hall.

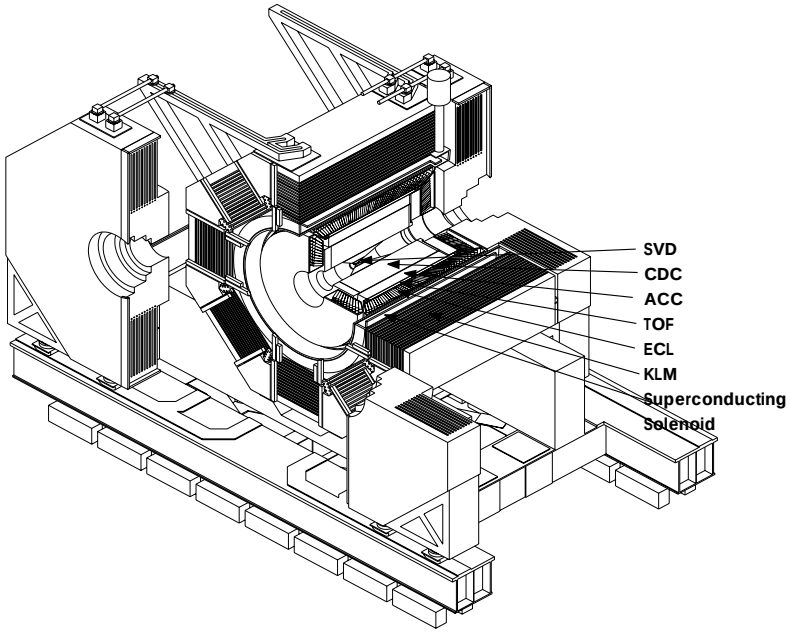


Figure 3.2: Belle detector.

and the polar angle from z-axis, respectively. The xy plane is often called $r\phi$ plane as shown in Figure 3.5.

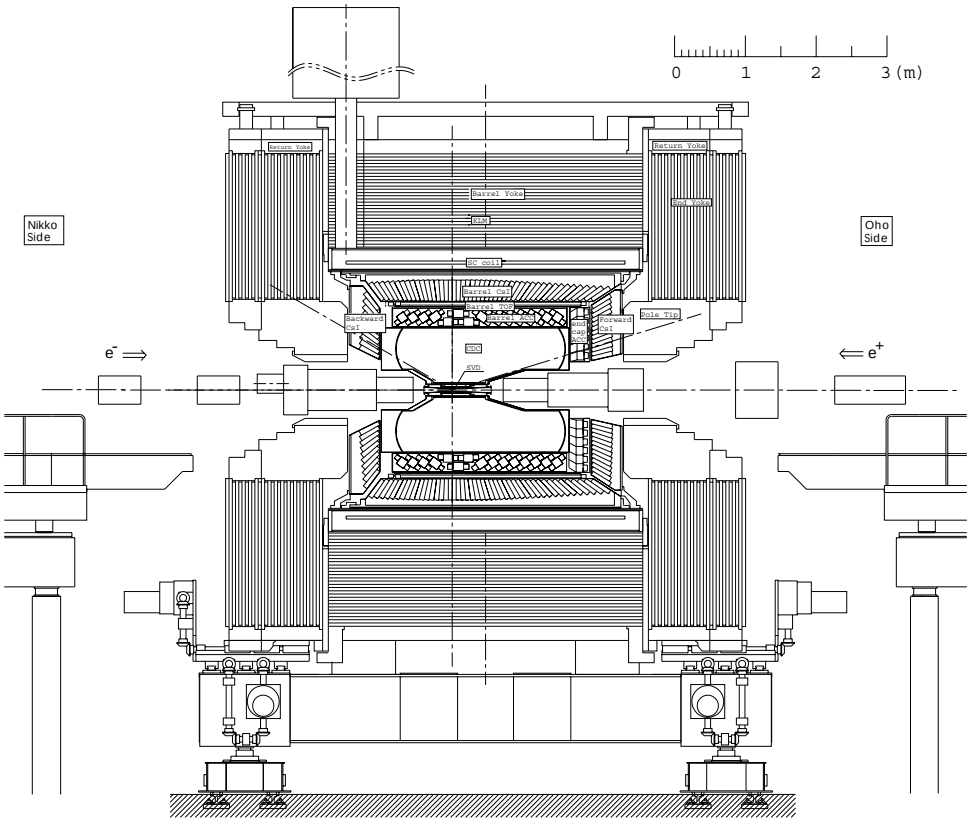


Figure 3.3: Side view of the Belle detector. The detector is asymmetric, that is, it has a larger acceptance toward the direction of the electron beam.

The Belle detector consists of a three-layer silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of 1188 aerogel Čerenkov counters (ACC), 128 time-of-flight (TOF) scintillation counters, an electromagnetic calorimeter containing 8736 CsI(Tl) crystals (ECL), 14 layers of 4.7-cm-thick iron plates interleaved with a system of resistive plate counters (KLM) and an extra calorimeter made of BGO crystals (EFC). All subdetectors, apart from the KLM, are located inside a 3.4-m-diameter superconducting solenoid that provides a 1.5 Tesla magnetic field. Table 3.1 summarizes the achieved performance of the subdetectors, which are described in the following sections.

3.2.1 SVD – Silicon Vertex Detector

The main task of the Silicon Vertex Detector (SVD) [31, 32] is to reconstruct the decay vertices of the two primary B mesons in order to observe time-dependent CP asymmetries in the decay of B mesons. It requires the measurement of the difference in z-vertex positions for B mesons with a precision of $\lesssim$200 μm. In addition, the vertex detector is useful for identifying and measuring the decay vertices of D and τ particles.

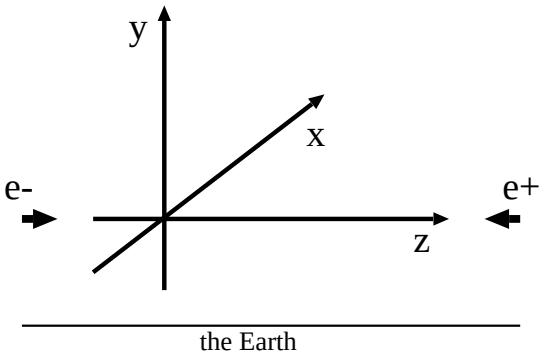


Figure 3.4: Definition of the coordinate in the Belle.

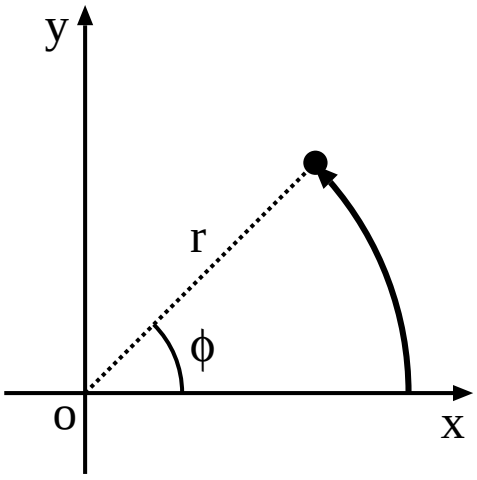


Figure 3.5: Definition of r and ϕ in the xy plane.

Table 3.1: Performance of subdetectors. p and $p_{\rm t}$ in GeV/c, E in GeV.

Detector	Type	Configuration	Readout	Performance
SVD	Double sided Si-strip	300 μm thick 3-layers r = 3.0–6.05 cm	ϕ: 40.96k θ: 40.96k	$\sigma_{\Delta z} \lesssim 200$ μm
CDC	Small-cell drift chamber	Wire: 50 layers Cathode: 3 layers r = 8–88 cm	W: 8.4k C: 1.8k	$\sigma_{r\phi}$ = 100–150 μm $\sigma_{p_{\rm t}}/p_{\rm t} = (0.2p_{\rm t} \oplus 0.3)\%$ $\sigma_{dE/dx} = 8\%$
ACC	Threshold Čerenkov $n = 1.01 - 1.03$ Silica aerogel	960 barrel 228 end-cap FM-PMT readout	1788	K/π separation: $1.0 < p < 3.6$ GeV/c
TOF/TSC	Plastic scintillator	ϕ segmentation: 128/64 r = 120 cm	128×2/64	σ_t = 100 ps K/π separation: $p < 1.2$ GeV/c
ECL	CsI(Tl) crystal	r = 125–162 cm $z = -102/+196$ cm	r: 6624 z: 1152+960	$\sigma_E/E = (1.34$ $\oplus 0.07/E$ $\oplus 0.81\%/E^{1/4})\%$
KLM	Glass resistive plate counter	14 layers	barrel: 21856 end-cap: 16128	Angular resolution ~ 30 mrad for K_L 1% hadron fake rate for μ
EFC	BGO crystal	Segmentation: 32 in ϕ, 5 in θ	160×2	Energy resolution: 7.3% at 8.0 GeV 5.8% at 3.5 GeV
Magnet	Superconducting	Inner radius 170 cm		$B = 1.5$ Tesla

Since the majority of the decay particles of B mesons have momenta lower than 1 GeV/c, the vertex resolution is dominated by the multiple-Coulomb scattering. This imposes strict constraints on detector design and mechanical layout. The innermost layer of the vertex detector must be placed as close to the interaction point as possible. The support structure must be low in mass but rigid. The readout electronics must be placed outside the tracking volume.

The SVD must also withstand large beam backgrounds. With the current operation of KEKB, the radiation dose to the detector due to beam background is measured to be 10 kRad/month. Radiation doses of this level both degrade the noise performance of the electronics and induce leakage currents in the silicon detectors. In addition, the beam backgrounds induce large single-hit count rates. The electronic shaping time is determined by a tradeoff between the design to minimize count-rate and leakage current effects, which argue for short shaping times, and input-FET noise of front-end integrated circuits, which is minimized with longer shaping times.

The SVD has three cylindrical layers consisting of units of the silicon sensors as shown in Figure 3.6. It covers a solid angle $23° < \theta < 139°$ where θ is the polar angle from the beam axis. This corresponds to 86% of the full solid angle. The radii of the three layers are 30 mm, 45.5 mm and 60.5 mm. Each ladder is constructed from independent ladders. Each ladder consists of double-sided silicon strip detectors (DSSDs) reinforced by support ribs which are made of boron-nitride sandwiched by carbon-fiber reinforced plastic.

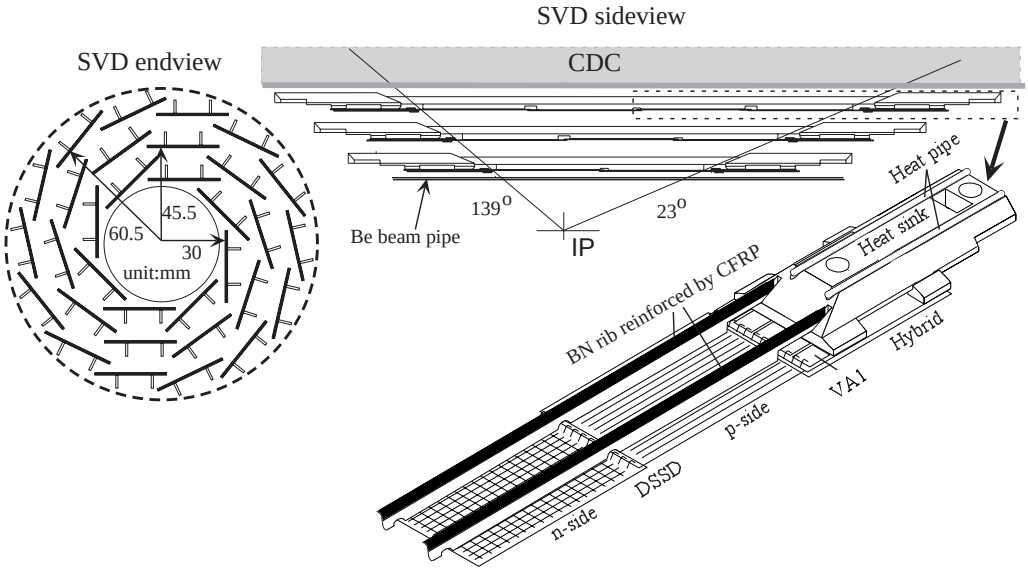


Figure 3.6: Detector configuration of the SVD.

The S6936 DSSDs are fabricated by Hamamatsu Photonics. Each DSSD with a thickness of 300 μm consists of 1280 sense strips and 640 readout pads on opposite sides. One side (n-side) of the DSSDs has n^+ strips oriented perpendicular to the beam direction to measure the z coordinate. The other side (p-side) has longitudinal p^+ strips for ϕ coordi-

nate measurements. The z-strip pitch is 42 μm and the ϕ-strip pitch is 25 μm. Adjacent strips are connected to a single readout trace on the z-side which gives an effective strip pitch of 84 μm. Every other sense-strip is connected to readout electronics on the ϕ-side. Signals collected by floating strips are read from adjacent strips by means of capacitive charge division. In total 102 DSSDs are used with a total number of readout channels of 81920.

The VA1 chips fabricated with the 1.2-μm CMOS process by Austrian Micro Systems are used as the readout LSI for the DSSDs. The VA1 chip consists of 128-channel pre-amplifiers, shapers, sample/hold circuits and analog multiplier. Radiation hardness tests of VA1 indicate that it is radiation tolerant to levels of $\sim$ 200 kRad.

The momentum and angular dependence of the impact parameter resolution are shown in Figure 3.7 and well represented by the following formulas.

$$\sigma_{xy} = 19 \oplus 50/(p\beta \sin^{3/2}\theta) \ \mu\mathrm{m}$$
$$\sigma_{z} = 36 \oplus 42/(p\beta \sin^{5/2}\theta) \ \mu\mathrm{m}$$

The resolution $\sigma_{\Delta z}$ of the difference in z-vertex positions for B mesons is obtained to be $\sim$200 μm in the measurement of the $\sin 2\phi_1$ [3].

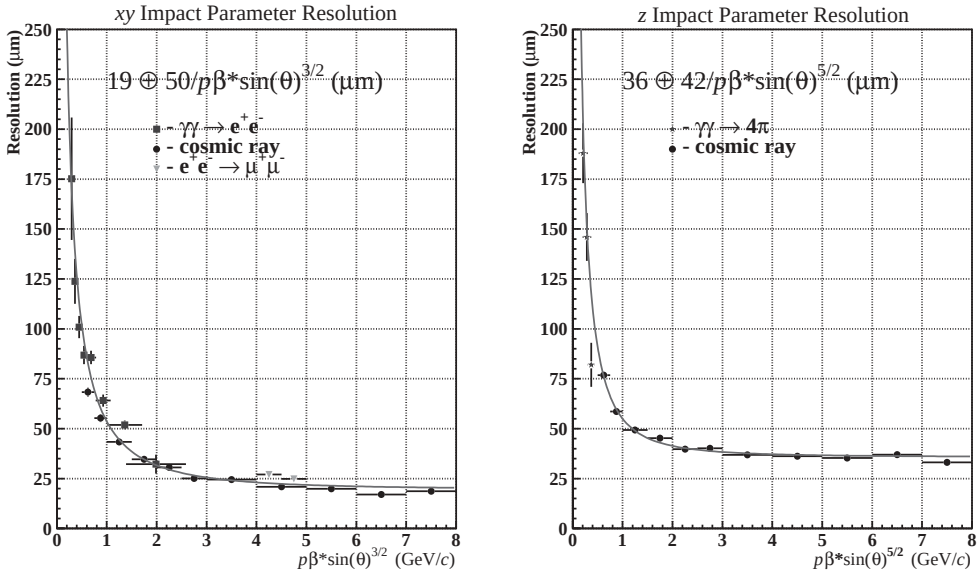


Figure 3.7: Impact parameter resolution.

3.2.2 CDC – Central Drift Chamber

The main role of Central Drift Chamber (CDC) [31, 33] is the efficient reconstruction of charged particle tracks and precise determination of their momenta. Specifically, the physics goals of the Belle experiment require a momentum resolution better than

$$\sigma_{p_t}/p_t \sim 0.5\% \cdot \sqrt{1 + p_t^2} \ (p_t \text{ in GeV}/c)$$

for all charged particles with $p_t \geq 100$ MeV/c, where p_t is the transverse momentum of a track. In addition, the charged particle tracking system is expected to provide important information for the trigger system. It also provides information on particle identification (PID) in the form of precise dE/dx measurements for charged particles.

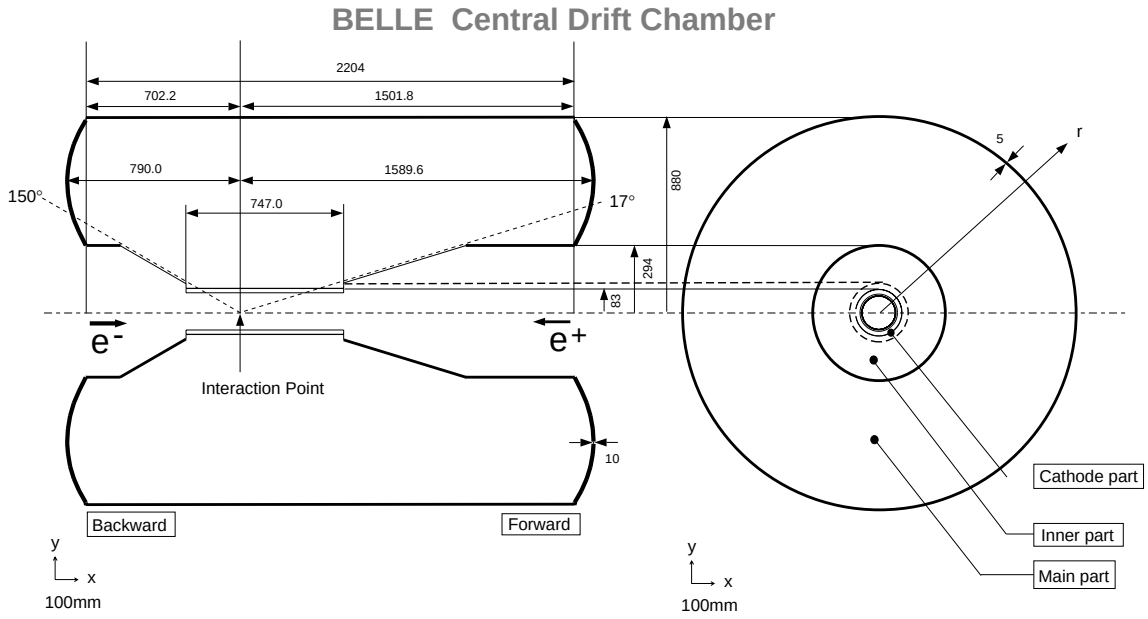


Figure 3.8: Overview of the CDC structure. The dimensions in the figure are in units of mm.

The structure of the CDC is shown in Figure 3.8. It is asymmetric in the z direction in order to provide an angular coverage of $17° < \theta < 150°$ which corresponds to 92% of the full solid angle. The longest wires are 2400 mm long. The inner and outer radii are 83 mm and 880 mm, respectively. There is no wall in the radius region from 103.5 mm to 875 mm in order to obtain good tracking efficiency for low-p_t tracks by minimizing the material thickness. The forward and backward small-radius regions have conical shapes in order to clear the accelerator components while maximizing the acceptance.

The CDC consists of 50 sense wire layers and 3 cathode strip layers. The sense wire layers are grouped into 11 superlayers, where 6 of them are axial and 5 are stereo superlayers as shown in Figure 3.9. An axial superlayer consists of only axial wires which do not slant to z-axis and a stereo superlayer only stereo wires which slant to z-axis. Stereo wires are necessary to reconstruct z component of the momentum of charged tracks. The number of sense wires is 8400 in total (axial: 5280, stereo: 3120) as shown in Table 3.2. The cathode strips provide z coordinate measurements at the inner-most radii. They are divided into 8 segments in the ϕ direction and have an 8.2 mm pitch in the z direction. The number of cathode channels is 1792.

A mixture of helium (50%) and ethane (50%) is used in order to minimize the multiple scattering. Since low-Z gases have a smaller photo-electric cross section than argon-based gases, they have the additional advantage of reduced background from synchrotron

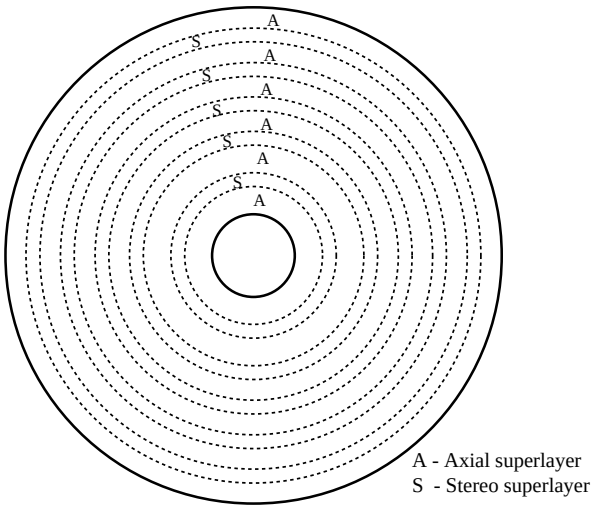


Figure 3.9: $r\phi$ view of the CDC. There are 6 axial and 5 stereo superlayers.

Table 3.2: Sense wires of the superlayers. The first and last rows of this table are the most inner and outer superlayers, respectively.

Layer type	No. of the layer	Wires per one layer	Total wires
axial	6	64	384
stereo	3	80	240
axial	6	96	576
stereo	3	128	384
axial	5	144	720
stereo	4	160	640
axial	5	192	960
stereo	4	208	832
axial	5	240	1200
stereo	4	256	1024
axial	5	288	1440

radiation. This mixture has a long radiation length (640 m) and a drift velocity that saturates at 4 cm/μs at a relatively low electric field ($>\sim$1.8 kV/cm·atm). The use of a saturated gas makes calibrations simpler and helps to ensure reliable and stable performance. Even though the gas mixture has a low Z, a good dE/dx resolution is provided by the large ethane component.

The spatial resolution is shown in Figure 3.10 as a function of the drift distance. Near the sense wire and near the cell boundary the spatial resolution is significantly poorer. The spatial resolution for tracks passing near the middle of the drift space is approximately 100 μm.

The reconstruction of the charged particle tracks is described in Appendix A. The $p_{\rm t}$ resolution as a function of $p_{\rm t}$ is shown in Figure 3.11. The solid curve indicates the result of a fit to the data points:

$$(0.201\pm0.003)p_{\rm t}\oplus(0.290\pm0.006)/\beta\% \ (p_{\rm t} \text{ in GeV}/c).$$

where β is the velocity of a particle normalized by the light velocity. The dashed curve shows the ideal expectation for the charged particles ($\beta = 1$):

$$0.188p_{\rm t}\oplus0.195\%.$$

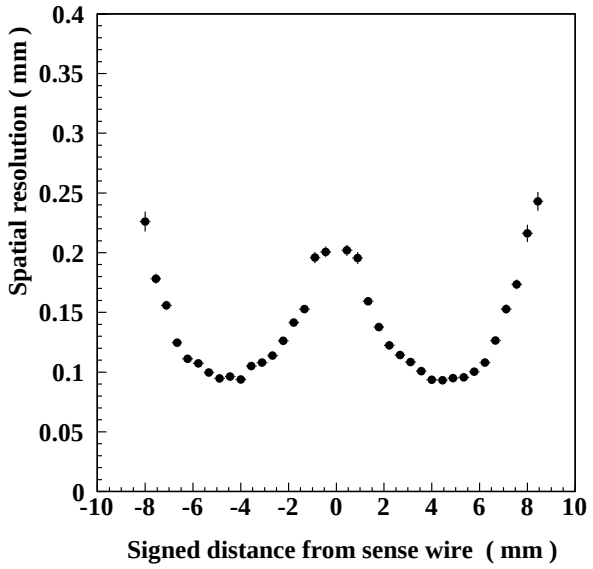


Figure 3.10: Spatial resolution as a function of the drift distance.

The dE/dx measurement in the CDC provides the particle identification information in the lower ($<$0.8 GeV/c) and the higher ($>$2.0 GeV/c) momentum ranges. Since dE/dx is a function of β of the particle[a], it can distinguish particles with the same momentum

[a]The Bethe-Bloch formula [6] for the mean rate of energy loss of a charged particle is given by

$$-\frac{dE}{dx} = \frac{4\pi N_{\rm A} z^2 \alpha^2}{m_e c^2}\frac{Z}{A}\frac{1}{\beta^2}\left\{\ln\left[\frac{2m_e c^2\beta^2}{I(1-\beta^2)}\right] - \beta^2\right\},$$

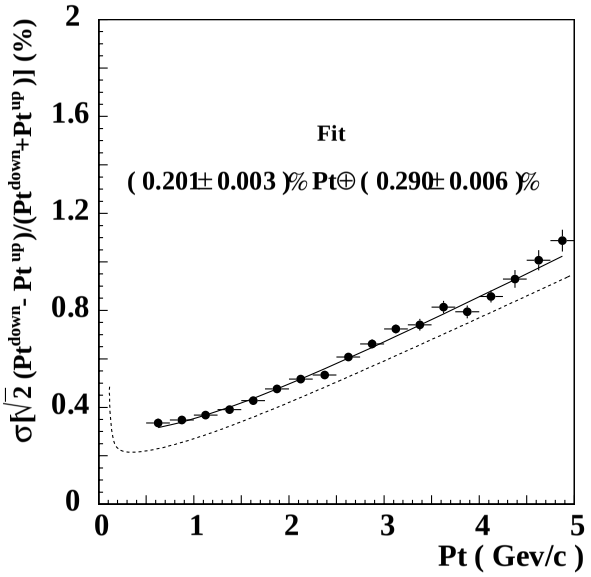


Figure 3.11: $p_{\rm t}$ dependence of $p_{\rm t}$ resolution for cosmic rays. The detail is described in the text.

but different masses. The truncated-mean method is used to estimate the most probable energy loss. The largest 20% of measured dE/dx values for each track are discarded and the remaining data are averaged in order to minimize the effect of occasional large fluctuations in the dE/dx distribution. A scatter plot of measured $\langle dE/dx \rangle$[b] and particle momentum is shown in Figure 3.12, together with the expected mean energy losses for different particle species. Populations of pions, kaons, protons and electrons can be clearly seen. The normalized $\langle dE/dx \rangle$ distribution for minimum ionizing pions from K_S decays is shown in Figure 3.13. The $\langle dE/dx \rangle$ resolution is measured to be 7.8% in the momentum range from 0.4 to 0.6 GeV/c.

3.2.3 ACC – Aerogel Čerenkov Counter

Particle identification, specifically the ability to distinguish $\pi^{\pm}$ from $K^{\pm}$, plays a key role in the Belle experiment. An array of the Aerogel Čerenkov Counter (ACC) system [31, 34] has been selected as a part of the Belle particle identification system to extend the momentum coverage beyond the reach of dE/dx measurements by the CDC and time-of-flight measurements by the TOF.

The ACC makes use of the character of Čerenkov light that the threshold of Čerenkov

where α is the fine structure constant ($\sim 1/137$), m_e is the electron mass, z and β the charge (in units of e) and velocity (in units of c, $\beta = v/c$) of the particle, $N_{\rm A}$ is Avogadro's number, Z and A are the atomic umber and mass number of the atoms of the medium, and x is the path length in the matter transversed. The quantity I is an effective ionization potential, averaged over all electrons, with approximate magnitude $I = 10Z$ eV. In the small β, $-\frac{dE}{dx}$ varies as $1/\beta^2$ and after passing through the minimum value $\sim$1 MeV·cm^2·g^{-1} at $\beta\gamma = 3 - 4$ ($\gamma = 1/\sqrt{1-\beta^2}$), it increases slowly.

[b] $\langle\rangle$ indicates the value obtained by the truncated-mean method.

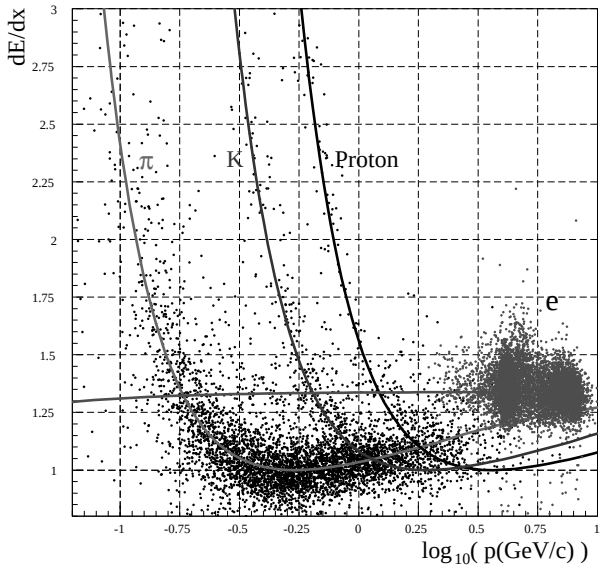


Figure 3.12: Truncated mean of dE/dx versus momentum observed in collision data. The units of dE/dx and p are MeV·cm^2·g^{-1} and GeV/c. Other PID detectors are used for the area around the intersection of the lines for kaon and pion ($\log_{10}(p = 1.2) \sim 0.08$).

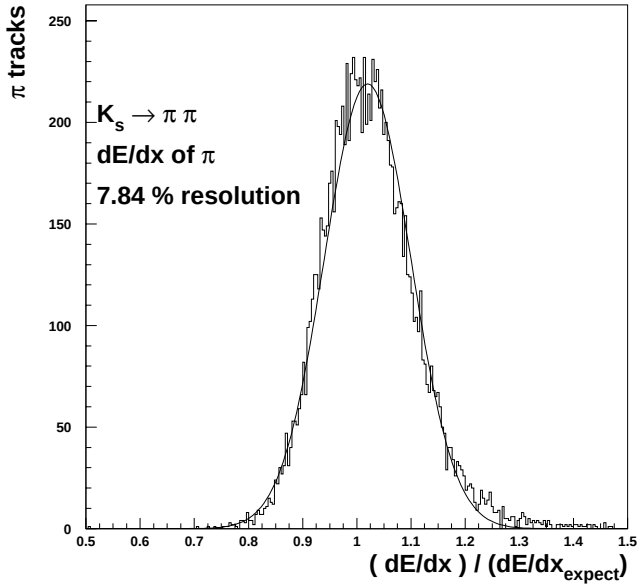


Figure 3.13: Distribution of $\langle dE/dx \rangle$ for pions from K_S decays. The $\langle dE/dx \rangle$ is normalized by its expected value.

light emission in matter with refractive index n is dependent on the velocity β of a particle as follows:

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

where p is the particle momentum. There are regions where the pion produces Čerenkov light in the aerogel while the kaon does not. As the value of the refractive index becomes larger, the width of this region becomes narrower and the momentum range lower.

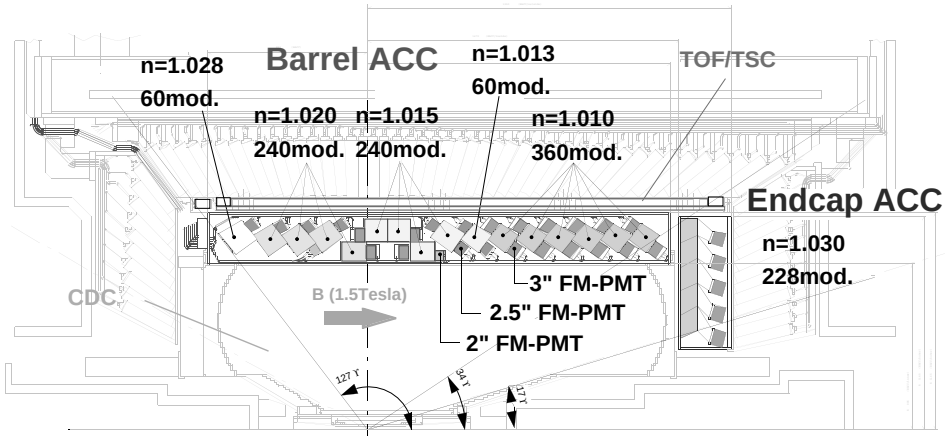


Figure 3.14: Arrangement of the ACC at the central part of the Belle detector.

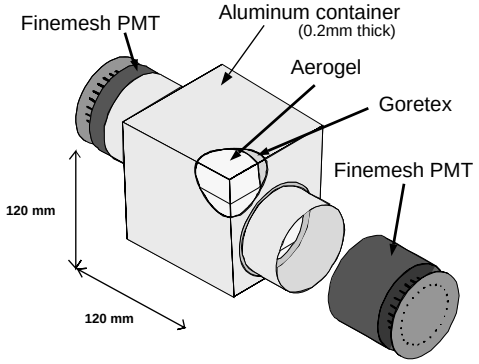
The configuration of the ACC is shown in Figure 3.14. The ACC consists of 960 counter modules segmented into 60 cells in the ϕ direction for the barrel part (BACC) and 228 modules arranged in 5 concentric layers for the forward end-cap part (EACC) of the detector. The BACC and EACC cover angular ranges of $34^\circ < \theta < 127^\circ$ and $17^\circ < \theta < 34^\circ$ respectively. In the BACC, five different refractive indices, $n = 1.010$, 1.013, 1.015, 1.020, 1.028, are used depending on the polar angle in order to provide 3σ K/π separation in the momentum region $1.0 < p_K < 3.6$ GeV/c, where p_K is a momentum of kaon. In the EACC, only one refractive index is used $n = 1.030$ in order to provide 3σ K/π separation in the momentum region $0.7 < p_K < 2.4$ GeV/c.

All the counters are arranged in a semi-tower geometry, pointing to the interaction point. A typical single ACC module is shown in Figures 3.15 for the barrel and the end-cap ACC, respectively. Five aerogel tiles are stacked in a 0.2-mm-thick aluminum box of approximate dimensions 12×12×12 cm^3. In order to detect Čerenkov lights effectively, one or two fine mesh-type photomultiplier tubes (one: 360 of BACC and all EACC, two: 600 of BACC), which are operated in a magnetic field of 1.5 T, are attached directly to the aerogels at the sides of the box.

Figure 3.16 shows the measured pulse height distribution for the BACC for $e^\pm$ tracks in Bhabha events and also $K^\pm$ candidates in hadronic events, which are selected by the TOF and dE/dx measurements. The figure demonstrates a clear separation between high energy electrons and particles below the threshold. It also indicates good agreement between the data and Monte Carlo simulations.

The particle identification of charged kaons is described in Appendix B.

(a) Barrel ACC Module



(b) Endcap ACC Module

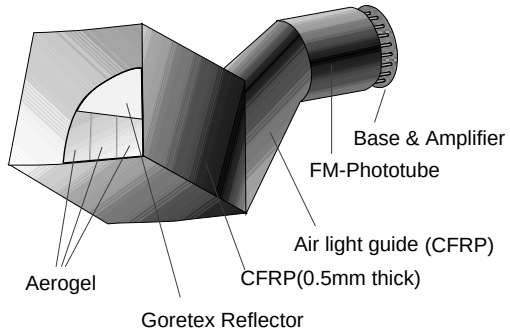


Figure 3.15: Schematic drawing of a typical ACC counter module: (a) barrel and (b) end-cap ACC.

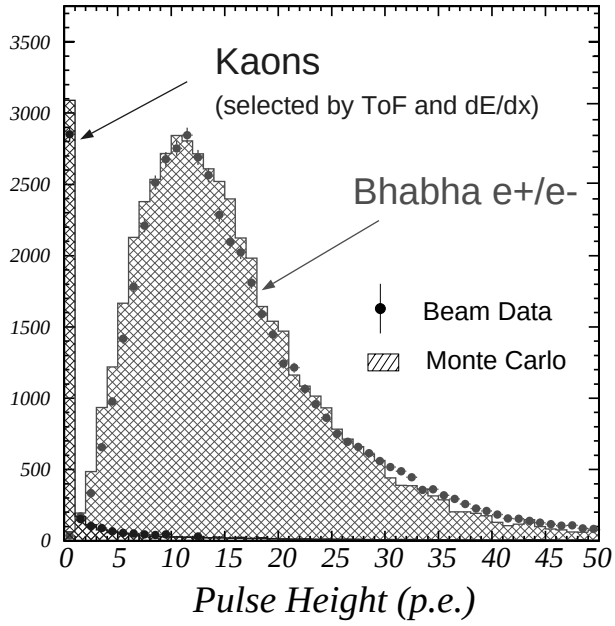


Figure 3.16: Pulse-height spectra in units of photoelectrons observed by barrel ACC for electrons and kaons. Kaon candidates are obtained by the dE/dx and TOF measurements. The Monte Carlo expectations are superimposed.

3.2.4 TOF – Time/Trigger of Flight Counter

The Time of Flight (TOF) [31, 35] system using plastic scintillation counters is powerful for particle identification. For a 1.2 m flight path, the TOF system with 100 ps time resolution is effective for particle momenta below $\sim$1.2 GeV/c. Most ($\sim$90%) of the charged particles produced in $\Upsilon(4S)$ decays have momenta below $\sim$1.2 GeV/c. In addition, the TOF counters provide fast timing signals for the trigger system to generate gate signals for ADCs and stop signals for TDCs.

The relation between the measured flight time T and the measured particle momentum p is expressed as

$$T = L/c\sqrt{1+(m/p)^2},$$

where L is the flight length depends on the TOF geometry and m is the particle mass. In case of $L = 1.2$ m, the flight time differences between kaon and pion are calculated to be 1470, 299 and 194 ps at $p = 0.5$, 1.2 and 1.5 GeV/c, respectively, which corresponds to approximately 15σ, 3σ and 2σ K/π separation for 100 ps time resolution.

The TOF system consists of 128 TOF counters and 64 Thin Trigger Scintillation Counters (TSC). Two trapezoidally shaped 4-cm-thick TOF counters and one 5-mm-thick TSC counter, with a 1.5-cm intervening radial gap, form one module as shown in Figure 3.17. In total 64 TOF/TSC modules located at a radius of 1.2 m from the interaction point cover a polar angle range from 34° to 120°. The 1.5-cm gap between the TOF and TSC counters is introduced to isolate the TOF from photon conversion

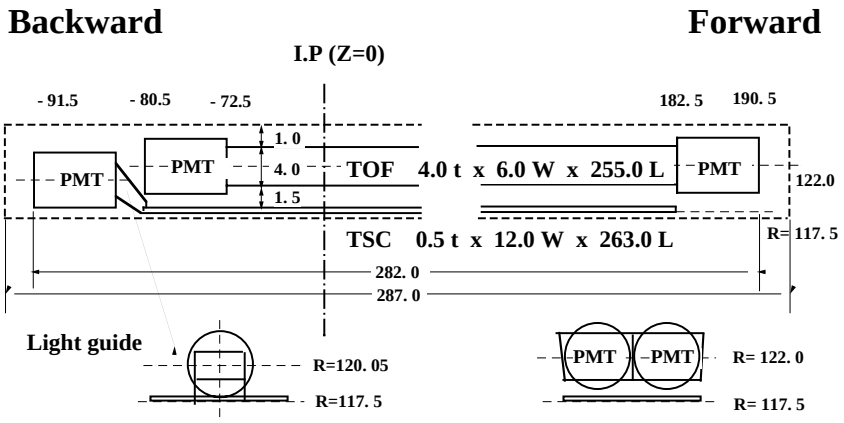


Figure 3.17: Dimensions of a TOF/TSC module.

backgrounds by taking the coincidence between the TOF and TSC counters. Electrons and positrons created in TSC are impeded from reaching the TOF counters due to this gap in the 1.5 T magnetic field.

Figure 3.18 shows the time resolutions for forward and backward PMTs and for the weighted average time as a function of z. These resolutions are obtained from μ-pair events. The resolution for the weighted average time is $\sim$100 ps with a small z dependence.

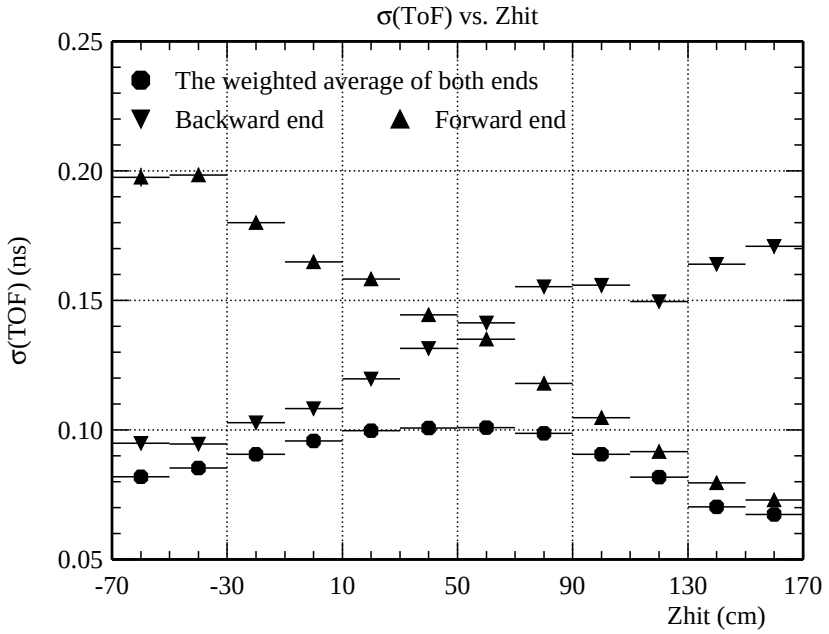


Figure 3.18: Time resolution of the TOF for μ-pair events.

3.2.5 ECL – Electromagnetic Calorimeter

The main purpose of the Electromagnetic Calorimeter (ECL) [31, 36] is the detection of electromagnetic showers caused by photons and electrons from B meson decays with high efficiency and good resolutions in energy and position.

Since most of these photons are end-products of cascade decays, they have relatively low energies and, thus, good performance below 500 MeV is especially important. On the other hand, important 2-body decay modes such as $B \to K^*\gamma$ and $B^0 \to \pi^0\pi^0$ produce photons energies up to 4 GeV and good energy resolution is necessary to reduce backgrounds for these modes. High momentum π^0 detection requires the separation of two nearby photons and a precise determination of their opening angle.

Electron identification in the Belle experiment relies primarily on a comparison of the charged particle momentum and the energy deposit in the electromagnetic calorimeter. Good electromagnetic energy resolution results in better hadron rejection.

In order to satisfy these requirements, we choose a highly segmented array of CsI(Tl) crystals with silicon photodiode readout installed in a 1.5 T field. CsI(Tl) crystals have many nice features such as a large photon yield, weak hygroscopicity and mechanical stability.

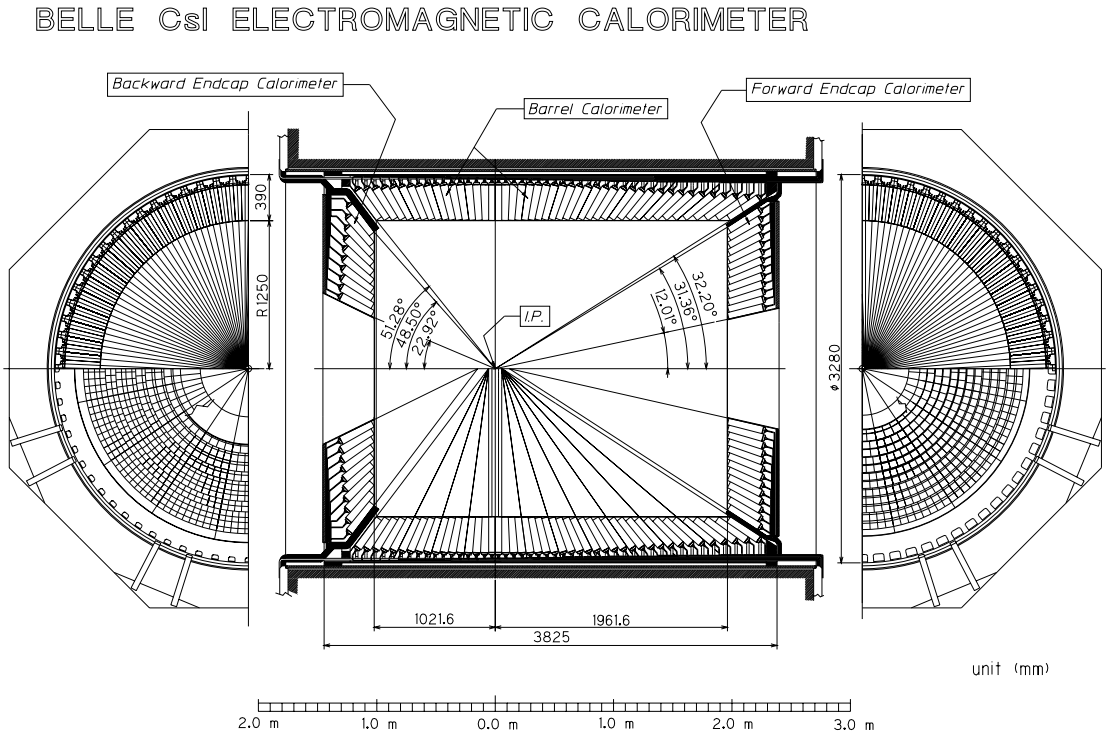


Figure 3.19: Configuration of the ECL.

The configuration of the ECL is shown in Figure 3.19. The ECL consists of the barrel section of ∼3.0 m in length with an inner radius of 1.25 m and the annular end-caps at $z \simeq +2.0$ m and $z \simeq -1.0$ m from the interaction point. The geometrical parameters of each section are given in Table 3.3, for instance, the barrel section has

46-fold segmentation in θ and 144-fold segmentation in ϕ with a total of 6624 crystals.

Table 3.3: Geometrical parameters of the ECL. The description is in text.

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

Each crystal has a tower-like shape and is arranged so that it points almost to the interaction point. A typical dimension of a crystal is 55×55 mm^2 (front face), 65×65 mm^2 (rear face) and 300 mm (length) of the barrel part. The 30 cm length ($16.2X_0$) is chosen to avoid deterioration of the energy resolution at high energies due to the fluctuations of shower leakages out of the rear of the counter. The ECL covers the polar angle region of $17.0° < \theta < 150.0°$ which corresponds to a total solid-angle coverage of 91% of the full solid angle. Small gaps between the barrel and end-cap crystals provide a pathway for cables and a room for supporting the inner detectors. The loss of solid angle associated with these gaps is ∼3% of the total acceptance. The entire ECL system contains 8736 CsI(Tl) counters and weights 43 tons.

Figure 3.20 shows the energy resolutions which are obtained from Bhabha events. The resolution is achieved to be 1.70% for barrel ECL, and 1.74% and 2.85% for the forward and backward ECL, respectively. Also the photon energy resolution was measured to be

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

with the beam test before the installation into the Belle structure.

The efficiency of electron identification is greater than 90% and a hadron rate (the probability to misidentify hadron as electron) is ∼0.3% for $p > 1$ GeV/c as shown in Figure 3.21.

3.2.6 KLM – K_L/μ Detector

The main purpose of K_L/μ Detector (KLM) [31, 37] is identification of K_L's and muons with high efficiency over a broad momentum range greater than 600 MeV/c. The detection of K_L is necessary to reconstruct $B \to J/\psi K_L$. Muons are used to identify the flavor of B and to reconstruct $J/\psi \to \mu^+\mu^-$ in the CP violation measurements.

Figure 3.22 shows a schematic view 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 the forward and backward end-caps. The barrel region around the interaction point covers an angular range from 45° to 125° in the polar angle and the end-caps in the forward and backward directions extend this range to 20° and 155°, respectively. The

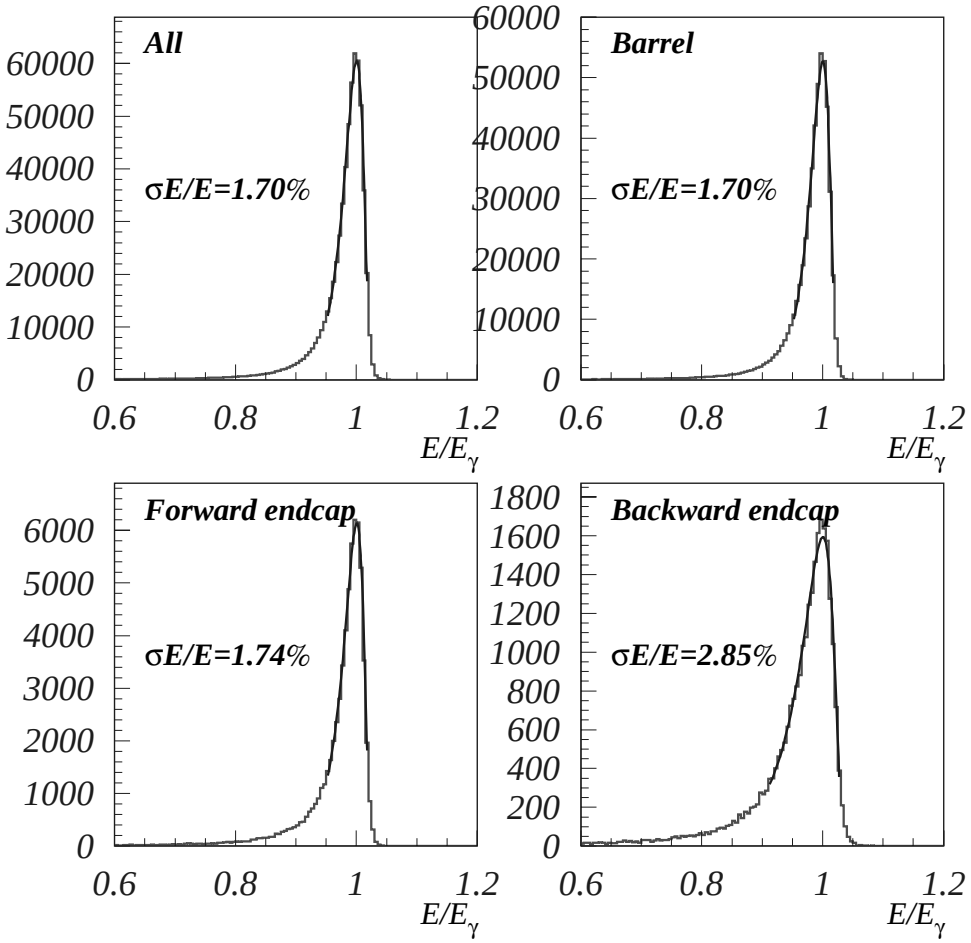


Figure 3.20: Energy resolutions measured from Bhabha events: overall, barrel, forward end-cap and backward end-cap.

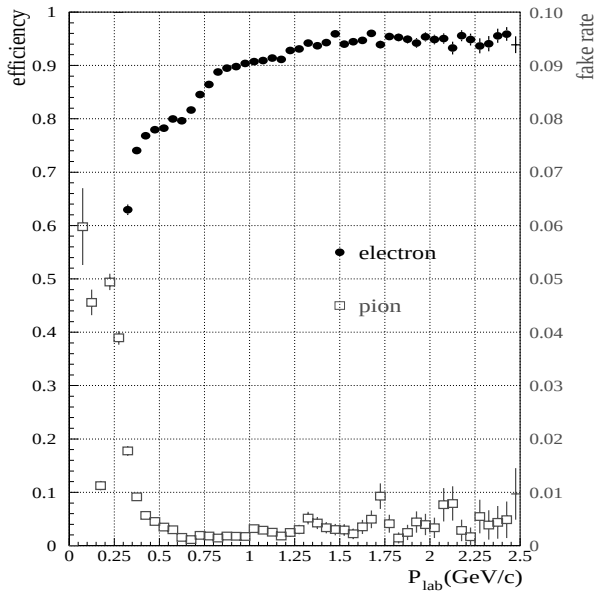


Figure 3.21: Electron identification efficiency and fake rate as a function of momentum.

number of readout channel of resistive plate counters (RPCs) is 21856 in barrel and 16128 in end-cap.

The detection of charged particles is provided by glass-electrode RPCs. RPCs have two parallel plate electrodes with high bulk resistivity ($\geq 10^{10}$ Ωcm) separated by a gas (Argon : Butane : Freon = 30 : 8 : 62)-filled gap. 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.

A K_L that interacts in the iron or ECL produces a shower of ionizing particles. The location of this shower determines the direction of the K_L, but fluctuations in the size of the shower do not allow a useful measurement of the K_L energy. The multiple layers of charged particle detectors and iron allow the discrimination between muons and charged hadrons ($\pi^{\pm}$ or $K^{\pm}$) based upon their range and transverse scattering. Muons travel much farther with smaller deflections on average than strongly interacting hadrons.

Figure 3.23 shows a histogram of the difference between the direction of the K_L cluster candidates and the missing momentum direction which is obtained from the hadronic events. We can see a clear peak where the direction of the neutral cluster measured in the KLM is consistent with the missing momentum in the event. The angular resolution for K_L is also estimated to be 0.03 mrad [38] from the Monte Carlo simulation.

Figure 3.24 shows the muon detection efficiency versus momentum for an appropriate likelihood cut in the cosmic ray events. The likelihood being a muon is obtained from a comparison of the measured range of a particle with the predicted range for a muon. We have a muon identification efficiency of better than 90% with a fake rate of less than 2%

for $p > 1.5$ GeV/c.

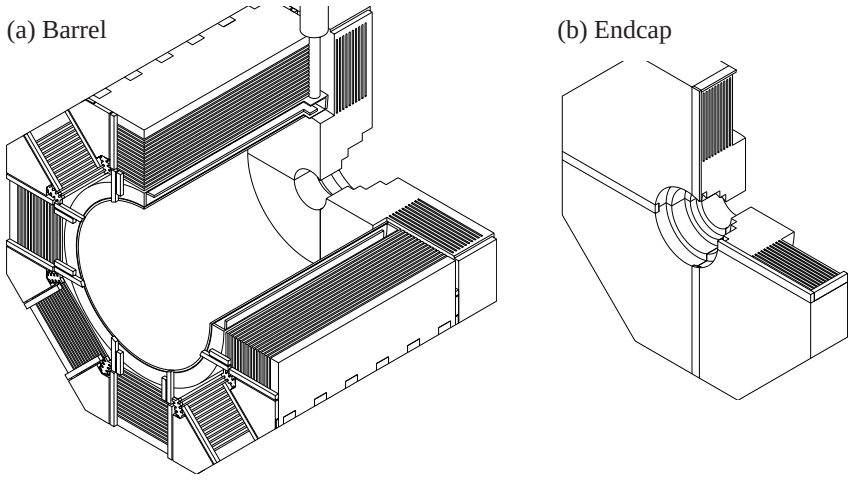


Figure 3.22: Schematic view of (a) barrel and (b) endcap KLM.

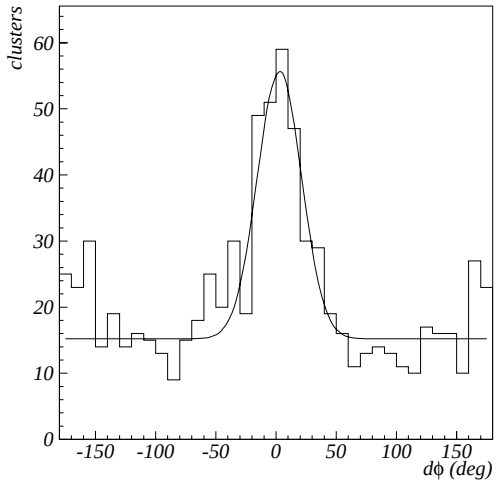


Figure 3.23: Difference between the neutral cluster and the direction of missing momentum in the KLM.

3.2.7 EFC – Extreme Forward Calorimeter

The purpose of the Extreme Forward Calorimeter (EFC) [31] is as a beam monitor for the KEKB control, as a luminosity monitor for the Belle experiment and for the improvement of the experimental sensitivity to some physics process such as $B \to \tau\nu$.

Since the EFC is placed in the very high radiation-level area around the beampipe near the interaction point, it consists of 320 radiation-hard BGO (Bismuth Germanate, $Bi_4Ge_3O_{12}$) crystals.

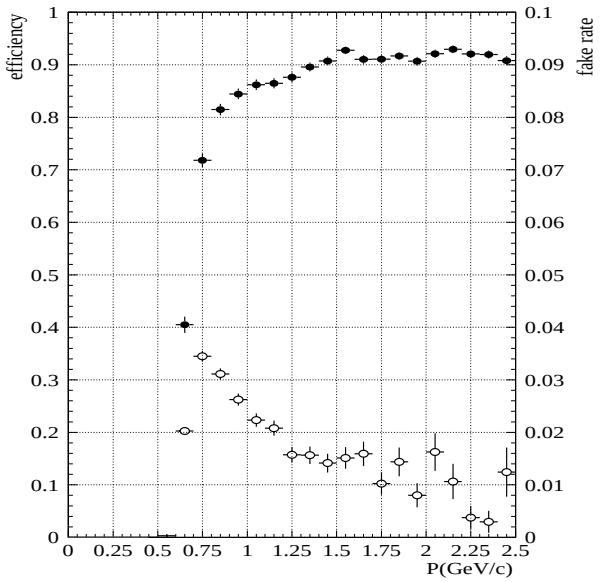


Figure 3.24: Muon identification efficiency and fake rate as a function of momentum.

The EFC covers the angular range from 6.4° to 11.5° in the forward direction and 163.3° to 171.2° in the backward direction.

The energy sum spectra for Bhabha events show a correlation between the forward and backward EFC. A clear peak at 8 GeV with an root-mean-square (rms) resolution of 7.3% is seen for the forward EFC, while a clear peak at 3.5 GeV with an rms resolution of 5.8% is seen in the backward EFC. These results are compatible with the beam test results.

3.2.8 Superconducting Solenoid

The superconducting solenoid provides a magnetic field of 1.5 T in a cylindrical volume of 3.4 m in diameter and 4.4 m in length. The coil is surrounded by a multi layer structure consisting of iron plates and calorimeters, which is integrated into a magnetic return circuit.

The two superconducting magnet complexes, QCS-R and QCS-L, are inserted in the holes along the axis of the end yokes. Each magnet complex consists of a solenoidal coil for compensation of the Belle solenoidal field, a quadrupole for focusing of the beams onto the interaction point, and several correction coils all located in a single cryostat.

The magnetic field [39] is measured as shown in Figure 3.25. The field strength is shown in Figure 3.26 as a function of z for various radii. The non-uniformity of the magnetic field can be clearly seen. The effect of the non-uniformity is taken into account in the reconstruction of charged particles by the Kalman Filter technique [40].

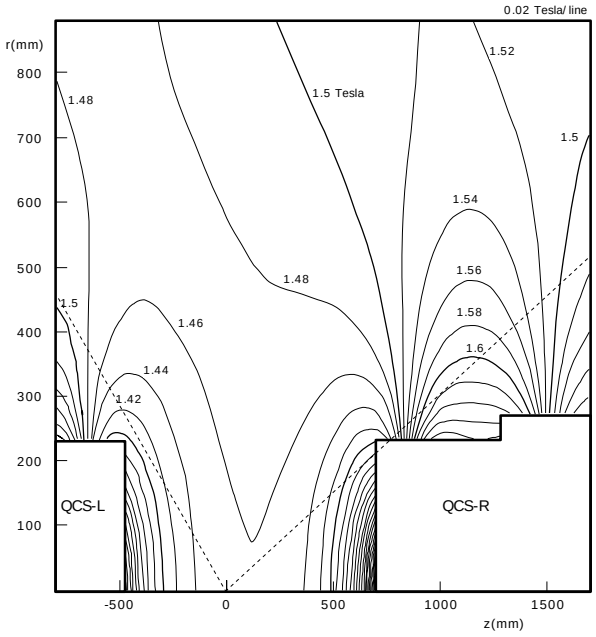


Figure 3.25: Contour plot of the magnetic field.

3.2.9 Trigger and Data Acquisition

The total cross sections and trigger rates at the design luminosity of 10^{34} cm^{-2}s^{-1} for various physical processes of interest are listed in Table 3.4. Since rates of Bhabha and $\gamma\gamma$ events are very large, the trigger rates must be prescaled by a factor of $\sim$1/100. In addition, high beam backgrounds are expected because of the high beam current. The trigger system is required to be robust against unexpectedly high beam background rates. The trigger conditions should be flexible so that background rates are kept within the tolerance of the data acquisition system, while the efficiency for physics events of interest is kept high.

Figure 3.27 shows the schematic view of the Belle trigger system. The trigger system consists of the subdetector trigger systems and the central trigger system called the Global Decision Logic (GDL). The subdetectors process event signals in parallel and provide trigger information to the GDL. The GDL combines the trigger signal from each subdetector and makes a final decision to initiate a Belle data acquisition within 2.2 μs from the event occurrence.

The global scheme of the Belle data acquisition system is shown in Figure 3.28. The distributed-parallel system has been devised in order to satisfy the data acquisition requirements of a 500 Hz trigger rate with a deadtime fraction of less than 10%. A typical data size of a hadronic event by $B\overline{B}$ or continuum production is measured to be $\sim$30 kB, which corresponds to the maximum data transfer rate of $\sim$15 MB/s.

The entire system is segmented into 12 subsystems running in parallel, each handling the data from a subdetector. Data from each subsystem are combined into a single event record by an event builder, which sorts the detector-wise parallel data to the event-wise

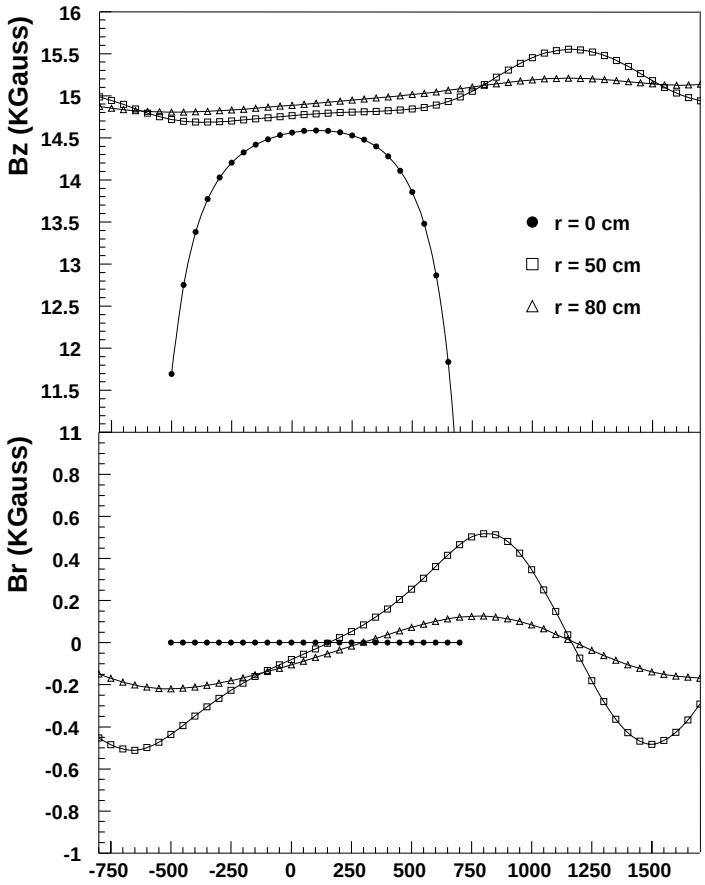


Figure 3.26: Magnetic field strength as a function of z (mm) for r =0, 50, and 80 cm.

Table 3.4: Total cross section and trigger rates with the luminosity of 10^{34} cm^{-2}s^{-1} for various physics processes at $\Upsilon(4S)$. Superscript$^{(A)}$ indicates the values prescaled by a factor of 1/100 and superscript$^{(B)}$ indicates the restricted condition of $p_{\rm t}$ $\geq$0.3 GeV/c.

Physics process	Cross section (nb)	Rate (Hz)
$\Upsilon(4S) \to B\overline{B}$	1.2	12
Hadron production from continuum	2.8	28
$\mu^+\mu^- + \tau^+\tau^-$	1.6	16
Bhabha ($\theta_{\rm lab} \geq 17^\circ$)	44	$4.4^{(A)}$
$\gamma\gamma$ ($\theta_{\rm lab} \geq 17^\circ$)	2.4	$0.24^{(A)}$
2γ processes ($\theta_{\rm lab} \geq 17^\circ$, $p_{\rm t} \geq 0.1$ GeV/c)	$\sim$15	$\sim 35^{(B)}$
Total	$\sim$67	$\sim$96

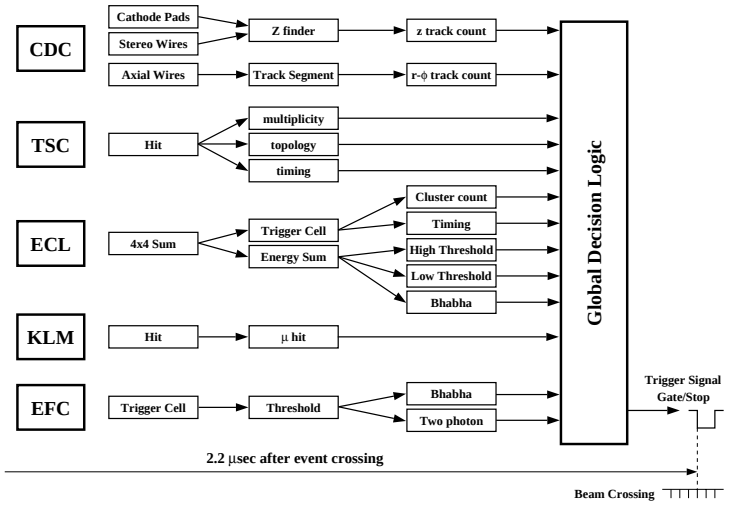


Figure 3.27: Level-1 trigger system for the Belle detector.

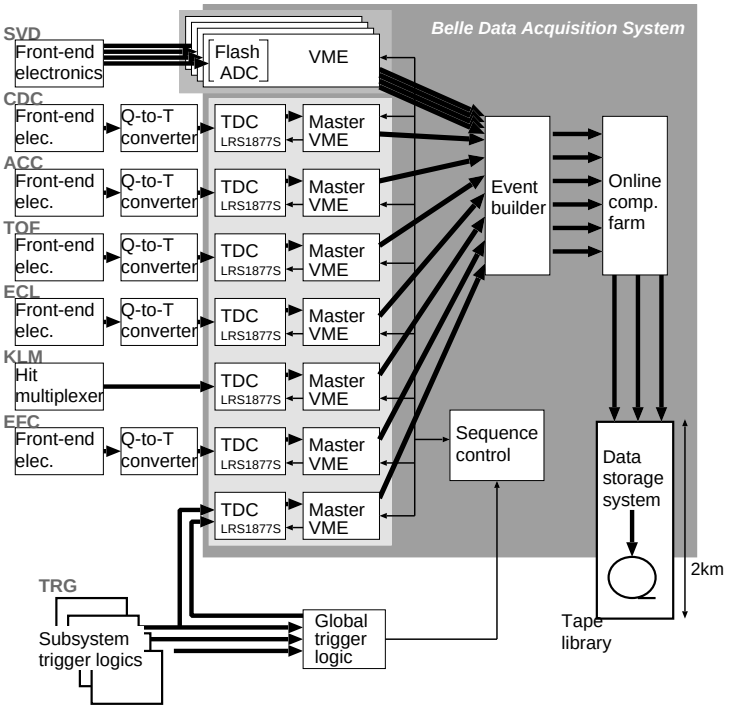


Figure 3.28: Overview of Belle DAQ system.

parallel data. The event builder output is transferred to an online computer farm. The data are then sent to a mass storage system located at the computer center via optical fibers.

3.3 Analysis Tools

3.3.1 Offline Computer System

The collected data by the Belle detector are analyzed at the offline computer system. Monte Carlo simulation is also an important task of the offline computer system since we need larger Monte Carlo samples than the collected data for the precise and reliable measurements of the physics. The parallel processing by multi-CPUs is necessary to achieve the required computational powers. A Framework for the Parallel Data Analysis (FPDA [41]) has been developed to support parallel event processing. The offline computer system consists of Symmetric Multi Processor (SMP) servers, tape libraries and disk servers. We have a total of ∼100 GHz equivalent Ultra Sparc II and ∼500 GHz equivalent Pentium III CPUs at KEK.

3.3.2 Software

The flow of the data analysis and Monte Carlo simulation is shown in Figure 3.29. The raw data obtained by the Belle detector are processed by the reconstruction tools, which perform the tracking of the charged particles in the CDC, clustering in the ECL to find photons, and particle identification. The output of the reconstruction is called a Data Summary Tape (DST) which is converted to a Mini Data Summary Tape (MDST [42]) for the analysis. In Monte Carlo simulation, we use the event generator and the full detector simulator to produce the raw data. The other parts of the analysis chain are the same as the case of the real data.

Based on the FPDA, a BASF (Belle AnalysiS Framework [43]) has been developed for the real data analysis and Monte Carlo simulation. The BASF is the main generic structure for the Belle analysis software and links different "module"s dynamically at run time. We provide an analysis code with a specific purpose as a module that is written as an object of a class of C++. Data are managed by Panther [44] which is a bank system based on an entity relationship model.

3.3.3 Monte Carlo Simulator

The event generator simulates the physical process of the particle decay chain. The initial state is $\Upsilon(4S)$ for $B\overline{B}$ or continuum ($q\overline{q}$) and the final states consists of stable particles. We use the "QQ98" generator which has been developed by CLEO [45]. The QQ98 has been modified [46] for the Belle experiment. The decay of $\Upsilon(4S)$ is performed by referring to the decay table that contains decay modes and branching ratios mainly measured by CLEO. We can control decays by changing this decay table. The continuum generation

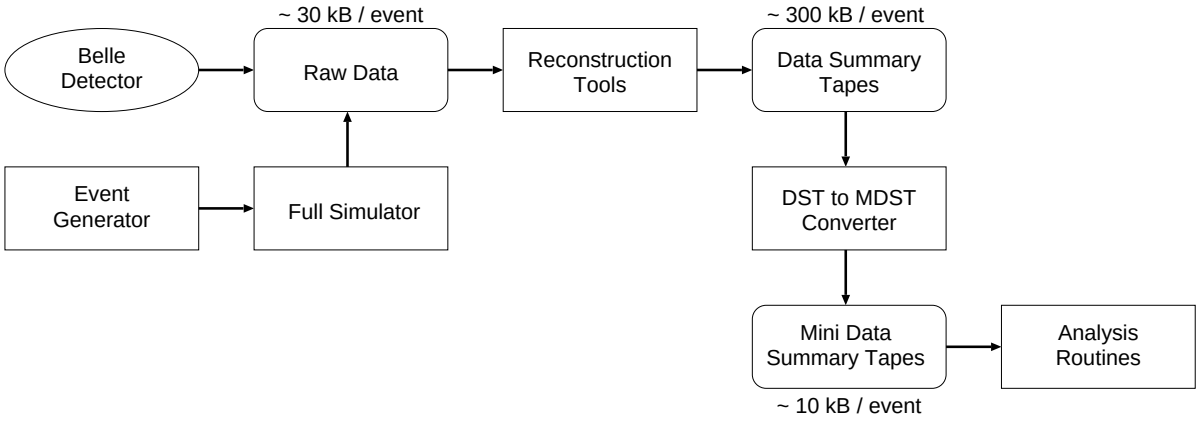


Figure 3.29: Flow of data analysis and Monte Carlo simulation. Rectangular and rounded boxes show programs and data, respectively.

uses the LUND (JETSET 7.3) program [47], in which the subsequent hadronization process is based on the Lund string fragmentation model [48].

The full detector simulator is based on the GEANT [49], which is a large library program developed at CERN to simulate reactions between particles and matters. This simulator takes the data from the QQ98 as an input and then traces the behavior of each particle in the detector, and simulates detector response[c].

[c]It takes (3–4) sec/event with dual 1GHz Pentium III CPUs.

Chapter 4

Selection and Reconstruction

We describe the Belle data set and the criteria used to select charged tracks and photons, and the ways in which they are combined to form the charm mesons used in this analysis. Then we explain the flight length reconstruction and the proper-time calculation to measure D lifetimes.

4.1 Data Set

The analysis is performed on data from the Belle experiment from Oct. 1999 to May 2001. We assign a specific index to our data set according to the period of the experiment. We call the data set which was recorded from Oct. 1999 to Dec. 1999, from Jan. 2000 to Jul. 2000, from Sep. 2000 to Dec. 2000, from Jan. 2001 to Apr. 2001 and from Apr. 2001 to Jul. 2001 as Experiment 5, 7, 9, 11 and 13, respectively as shown in Figure 4.1. For the lifetime measurements we use Experiment 5, 7 and 9 data set whose total integrated luminosity is 11.1 fb^{-1}[a], which corresponds to 1.4×10^7 $c\overline{c}$ events assuming the cross section $\sigma_{c\overline{c}} = 1.3$ nb^{-1}. For the y_{CP} measurement we use Experiment 7, 9, 11 and 13 (up to run number 531) whose total integrated luminosity is 23.4 fb^{-1}[b], which corresponds to 3.0×10^7 $c\overline{c}$ events. The difference between data sets for the lifetime and y_{CP} measurements is the version of the Belle software library which is used in the event reconstruction of the raw data. For y_{CP} measurement we have updated the Belle software library and then reprocessed our data from Experiment 7 through 13.

[a] 10.4 fb^{-1} for $\Upsilon(4S)$ resonance, 0.6 fb^{-1} for continuum, 0.1 fb^{-1} for energy scan. 0.2 fb^{-1} for Exp. 5, 6.5 fb^{-1} for Exp. 7, 4.4 fb^{-1} for Exp. 9.

[b] 20.9 fb^{-1} for $\Upsilon(4S)$ resonance, 2.3 fb^{-1} for continuum, 0.2 fb^{-1} for energy scan. 0.2 fb^{-1} for Exp. 5, 6.6 fb^{-1} for Exp. 7, 4.4 fb^{-1} for Exp. 9, 9.6 fb^{-1} for Exp. 11, 2.6 fb^{-1} for Exp. 13.

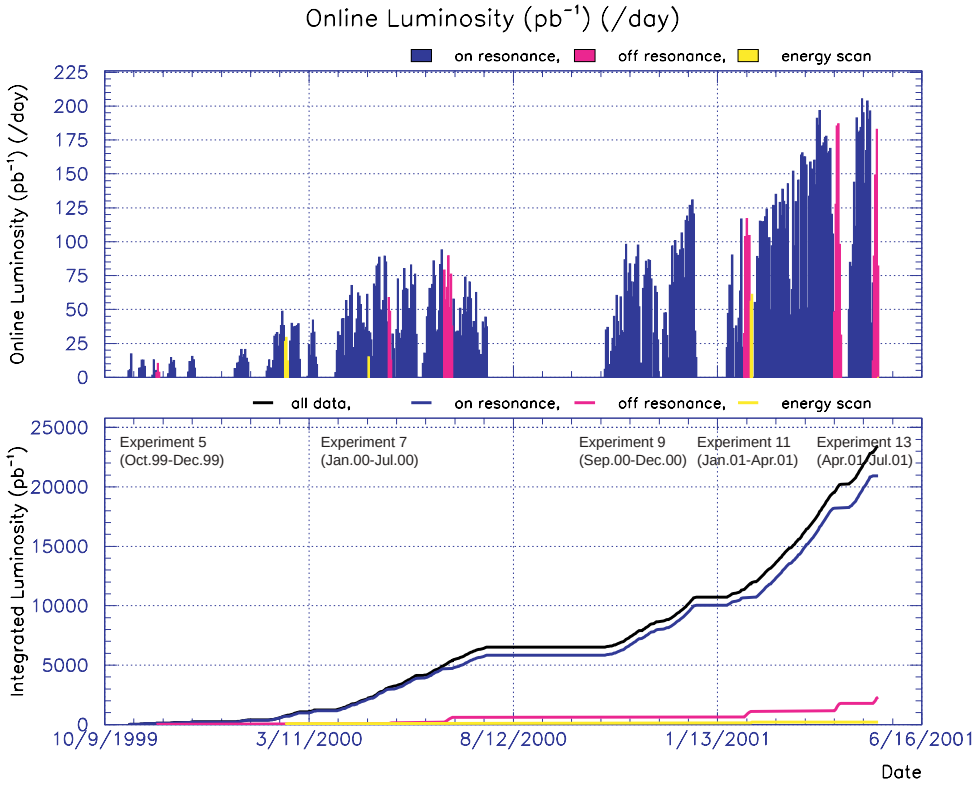


Figure 4.1: Integrated luminosity: per day (top) and as a function of day (bottom). The "on resonance" data are taken at the $\Upsilon(4S)$ resonance. The "off resonance" data are taken at 60 MeV/c^2 below the $\Upsilon(4S)$ resonance. The "energy scan" data are taken to find the $\Upsilon(4S)$ resonance.

4.2 Event Selection

For the lifetime measurements D^0, D^+ and D_s^+ mesons are reconstructed via the decay chains:

$$
\begin{aligned}
&D^0 \to K^-\pi^+ \\
&D^+ \to K^-\pi^+\pi^+ \text{ (with } D^{*+} \to D^+\pi^0 \text{ requirement)}, \\
&D^+ \to \phi\pi^+, \quad \phi \to K^+K^-, \\
&D_s^+ \to \phi\pi^+, \quad \phi \to K^+K^-, \\
&D_s^+ \to \overline{K}^{*0}K^+, \quad \overline{K}^{*0} \to K^-\pi^+.
\end{aligned}
$$

For the D^0 lifetime measurement, the following decay mode is used for the consistency check:

$$D^0 \to K^-\pi^+ \text{ (with } D^{*+} \to D^0\pi^+ \text{ requirement)}.$$

In the case of $D^+ \to K^-\pi^+\pi^+$ in which only a K^- exists, a $D^{*+} \to D^0\pi^+$ requirement is applied to decrease background events.

For the y_{CP} measurement D^0 mesons are reconstructed via the decay modes:

$$
\begin{aligned}
&D^0 \to K^-\pi^+ \\
&D^0 \to K^-K^+.
\end{aligned}
$$

4.2.1 Hadronic Event Selection

We record several event types, $e^+e^- \to u\overline{u}$, $d\overline{d}$, $s\overline{s}$, $c\overline{c}$, $b\overline{b}$, e^+e^-, $\mu^+\mu^-$, $\tau^+\tau^-$ and 2 photons. We also record cosmic rays and beam backgrounds. We call $e^+e^- \to u\overline{u}, d\overline{d}, s\overline{s}, c\overline{c}$ events continuum events. The beam backgrounds occur from two main sources; one is when e^- (or e^+) collides with beam gas and the other is spent electrons, which go out the beam line due to the bremsstrahlung and the collision with the beam gas and then come in the Belle detector.

The hadronic events are selected, which consists of continuum and $B\overline{B}$ events. The following criteria [50] are used to eliminate events from other sources, such as Bhabha (e^+e^-), $\mu^+\mu^-$, $\tau^+\tau^-$, 2 photons, cosmic rays, beam backgrounds.

The requirements are:

- Number of "good" charged tracks $N_{\rm trk}$:

  $$N_{\rm trk} \geq 3,$$

  where a good charged track is defined as a track with $p_{\rm t}$ >0.1 GeV/c, $|d_r|$ <2.0 cm and $|d_z|$ <4.0 cm, where d_r and d_z are the closest distance of a track to the z-axis and the z-position of the interaction point, respectively.

- Distance between a primary event vertex and the origin of the detector coordinate:

$$r_{\rm vtx} < 1.5 \text{ cm and } |z_{\rm vtx}| < 3.5 \text{ cm},$$

where a primary event vertex is obtained from all good charged tracks, $r_{\rm vtx}$ is a distance between a primary event vertex and the origin of the detector coordinate in the $r\phi$ plane, and $z_{\rm vtx}$ is z-position of the primary event vertex.

- Total "good" cluster energy $E_{\rm sum} = \sum_{\substack{\text{all good} \\ \text{clusters}}} E$:

$$0.18\sqrt{s} \leq E_{\rm sum} \leq 0.80\sqrt{s},$$

where a good cluster is defined as a cluster in the ECL with E >0.1 GeV and $\sqrt{s}$ represents the center of mass energy.

- Total visible energy detected by the CDC and ECL $E_{\rm vis} = \sum_{\substack{\text{all good} \\ \text{charged} \\ \text{tracks}}} p + \sum_{\substack{\text{all good} \\ \text{clusters}}} E$:

$$E_{\rm vis} \geq 0.20\sqrt{s}.$$

- Z-component of total visible energy:

$$\left| \sum_{\substack{\text{all good} \\ \text{charged} \\ \text{tracks}}} p_z \right| + \sum_{\substack{\text{all good} \\ \text{clusters}}} E_z \leq 0.50\sqrt{s}.$$

- The number of ECL clusters $n_{\rm ECL}$:

$$n_{\rm ECL} \geq 2.$$

- The average cluster energy:

$$E_{\rm sum}/n_{\rm ECL} < 1 \text{ GeV}.$$

- Heavy jet mass $M_{\rm jet}$:

$$M_{\rm jet} > 0.25E_{\rm vis} \text{ or } M_{\rm jet} > 1.8 \text{ GeV}/c^2.$$

Here the event is split into two hemispheres by a plane perpendicular to the event thrust axis. The invariant mass of tracks in each hemisphere is calculated assuming a pion mass. The larger invariant mass is called "heavy jet mass".

- Interaction point profile, described in Appendix C, can be obtained.

Background events are mostly rejected by requirement of the number of good charged tracks ($N_{\rm trk}$). The beam backgrounds are rejected by the requirement of a primary event vertex ($r_{\rm vtx}$ and $z_{\rm vtx}$) and energy criteria ($E_{\rm sum}$ and $E_{\rm vis}$). To further decrease background events, we impose requirements on the number of ECL clusters ($n_{\rm ECL}$) for e^+e^-, $\mu^+\mu^-$, 2 photons and beam background events, the average cluster energy for e^+e^- and $\mu^+\mu^-$ events, and heavy jet mass for $\tau^+\tau^-$ events. The requirement of the interaction point profile is necessary to obtain a flight length of the charm mesons. This criteria is $\sim$99% efficient for $B\overline{B}$ events and $\sim$80% efficient for the continuum events [c].

4.2.2 Track Selection

The charged tracks that are used to find a vertex must be associated with at least two SVD hits in both the $r\phi$ and the rz planes.

The χ^2/N.D.F (χ^2/number of the degrees of freedom) of track fit must be less than 5 to reduce poorly reconstructed tracks due to mis-association of SVD hits, incorrect SVD-hit clustering, large angle multiple scattering of tracks and so on. About 96% of tracks remain after this selection.

4.2.3 Particle Identification

For a charged track to be a pion or kaon candidate, the ionization energy in the CDC, time-of-flight information from the TOF and the hits in the ACC should be consistent with an expected hypothesis. We use a probability $Prob(X:Y)$ for the particle identification whose definition is described in Appendix B:

$$Prob(X:Y) = \frac{P_X}{P_X + P_Y},$$

where P_X is the PID likelihood calculated for the signal particle species and P_Y for the background particle species.

We require

- $Prob(K:\pi) > 0.6$ for K
- $Prob(\pi:K) > 0.1$ for π.

The measured K efficiency exceeds 80%, while the π fake rate is kept below 10%. No particle identification is required for the low momentum pion from D^{*+}.

Proton Veto

We reject proton candidates from kaon candidates by the next criterion.

- $Prob(p:K) > 0.95$ for K

The proton efficiency and K fake rate are estimated to be 84% and 1.5% from the Monte Carlo simulation by this criterion.

[c] $N_{\rm trk}$ and heavy jet mass cuts decrease the continuum events.

4.2.4 Photon and π^0 Selection

Photon candidates are associated with isolated CsI calorimeter clusters that are not matched to any charged track. A minimum energy cut of 20 MeV is applied. Lateral shower shapes must be consistent with those of photons.

π^0 candidates are reconstructed from two photons. Figure 4.2 shows a π^0 candidate mass distribution which is fitted with a single Gaussian plus a linear function: $\sigma = 4.6$ MeV/c^2. A mass-constrained fit is performed to improve the π^0 momentum resolution, where the mass-constrained fit means that momenta of the photons are varied within these errors to make the mass of the reconstructed π^0 to be the world average π^0 mass. A π^0 candidate must satisfy the following conditions.

- It must have the invariant mass before the mass-constraint fit within ± 14 MeV/c^2 of the world average value of π^0 mass [6].
- It must contain at least one photon with $-0.63 < \cos\theta_\gamma < 0.84$ (barrel region), where θ_γ is the polar angle with respect to the electron beam axis. This is because the energy resolution of the photon in the end-cap regions are worse than that in the barrel region and the number of the converted photon from the materials of the inner detectors is large in the end-cap region.
- It must also satisfy the decay angle cut $|\cos\theta_{\gamma,\pi^0}| < 0.9$, where θ_{γ,π^0} is the angle between the π^0 candidate and one of the daughter photons, evaluated in the π^0 rest frame. Here the $\cos\theta_{\gamma,\pi^0}$ distribution is expected to be flat for signal as described in Section 4.2.7.

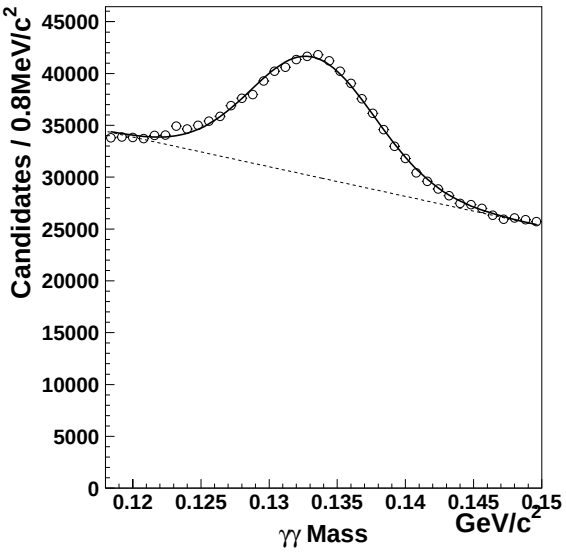


Figure 4.2: Mass distribution of π^0 candidates. The fit yields the width of $\sigma = 4.6$ MeV/c^2.

4.2.5 K_S Veto

Daughters of K_S candidates are removed from kaon and pion candidates. K_S candidates are reconstructed from $\pi^+\pi^-$ combination. A K_S candidate must satisfy the criteria listed in the Table 4.1 [51]. Here "*dr*" is the closest approach from the interaction point to one of the two daughter tracks in the $r\phi$ plane. The "$d\phi$" is the azimuthal angle between the momentum and the decay vertex vectors of a K_S candidate. The decay vertex vector is a vector pointing to the decay vertex from the interaction point. The "*z.dist*" is the z-distance between the two daughter tracks at their interception in the $r\phi$ plane. The "*fl*" is the flight length of a K_S candidate in the $r\phi$ plane. The information of the distance from the interaction point (*dr* and *fl*) can be used to identify a K_S since a K_S has a long lifetime ($c\tau$=2.7 cm). We require that the daughter tracks of K_S originate from the common vertex (*z.dist*). The $d\phi$ is peaked toward 0 since the momentum and the decay vertex vectors of K_S have the same direction. The width of the K_S mass distribution is $\sigma = 3.7$ MeV/c^2 [51].

Table 4.1: K_S selection criteria. A K_S candidate with the invariant mass within ± 12 MeV/c^2 of the world average [6] is vetoed. The definition of these parameters is described in the text.

Momentum (GeV/c)	*dr* (cm)	$d\phi$ (rad.)	*z.dist* (cm)	*fl* (cm)
< 0.5	> 0.05	< 0.3	< 0.8	-
0.5–1.5	> 0.03	< 0.1	< 1.8	> 0.08
> 1.5	> 0.02	< 0.03	< 2.4	> 0.22

4.2.6 Light Meson Selection: ϕ, $\overline{K}^{*0}$

An invariant mass cut is imposed to select ϕ and $\overline{K}^{*0}$ candidates. Table 4.2 lists the cuts on the invariant mass with the decay modes. We use natural widths [6] to select $\overline{K}^{*0}$. Figure 4.3 shows mass distributions of ϕ candidate with two Gaussians [52] and a linear function. We use $\sim 2.5\sigma_{\rm av}$ to select ϕ, where $\sigma_{\rm av}$ is the weighted average of the standard deviations of two Gaussians.

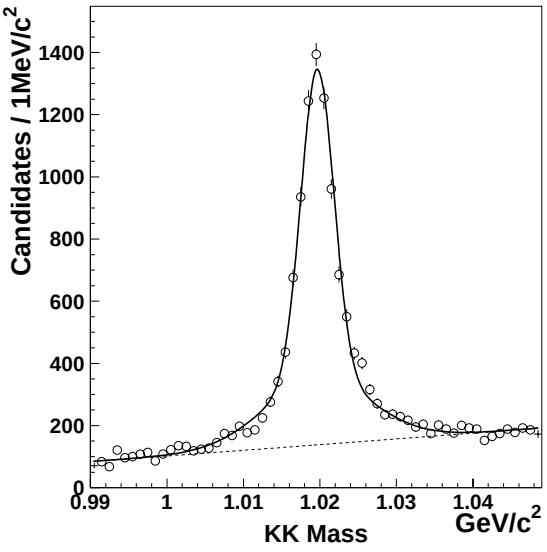


Figure 4.3: Mass distribution of ϕ candidates. The narrow and wide widths are 2.1 MeV/c^2 and 6.9 MeV/c^2, respectively. The area occupied by the narrow Gaussian is 60% of the total Gaussian (=narrow+wide).

Table 4.2: Light meson selection: decay modes, mass cuts used to select $\phi, \overline{K}^{*0}$ and natural width [6].

Mode	Mass cut (MeV/c^2)	Natural width (MeV/c^2)
$\phi \to K^+K^-$	$\pm$12	4.458$\pm$0.032
$\overline{K}^{*0} \to K^-\pi^+$	$\pm$50	50.7$\pm$0.6

4.2.7 D Meson Selection

Charm mesons are reconstructed in $D^0 \to K^-\pi^+$, $D^0 \to K^-K^+$, $D^+ \to K^-\pi^+\pi^+$, $D^+ \to \phi\pi^+$, $D_s^+ \to \phi\pi^+$ and $D_s^+ \to \overline{K}^{*0}K^+$ modes.

In order to select "prompt" charm mesons, we need to reject secondary charm mesons produced via B meson decays. The prompt charm meson decays at the interaction point, that is, comes from the continuum not $B\overline{B}$ events.

The D momentum in the cms is required to be greater than 2.5 GeV/c to eliminate secondary charm mesons from B mesons. A momentum of the D from $B\overline{B}$ events is mostly less than 2.5 GeV/c as shown in Figure 4.4.

In addition we require an event shape variable R_2 to be larger than 0.2 in order to reject combinatorial backgrounds from $B\overline{B}$ events. In general, a topology of $B\overline{B}$ events is spherical while that of continuum events is jet-like as shown in Figure 4.5. The R_2

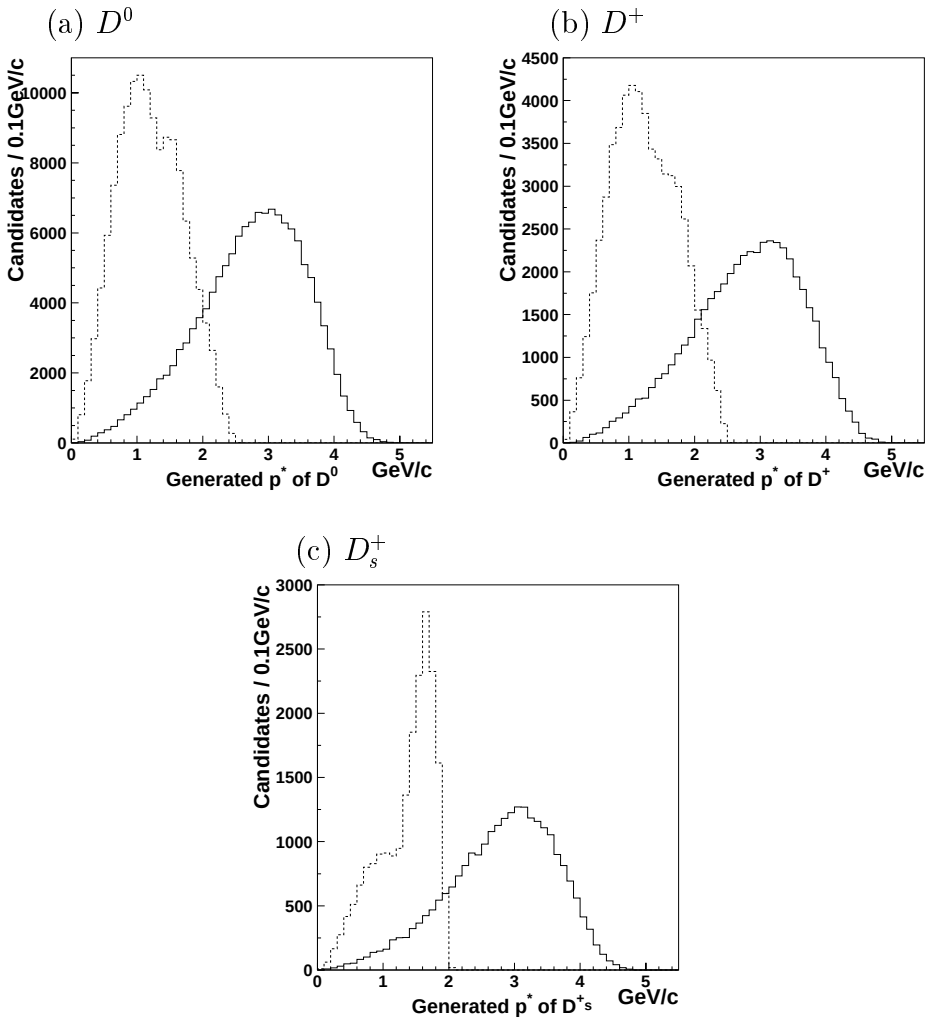


Figure 4.4: Generated D momentum in the cms: (a) D^0, (b) D^+ and (c) D_s^+. A solid (dotted) line is a shape of the continuum ($B\overline{B}$) event.

variable is an event shape variable that represents event topologies. R_2 is close to 1 for continuum events, while R_2 is close to 0 for $B\overline{B}$ events as shown in Figure 4.6.

(a) Continuum event (b) $B\overline{B}$ event

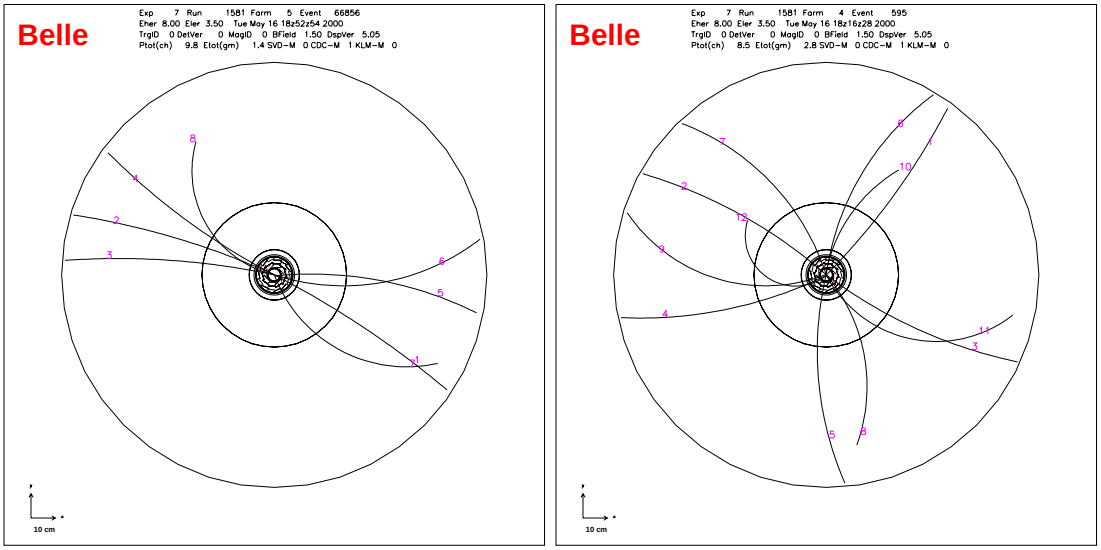


Figure 4.5: Event shape: (a) continuum and (b) $B\overline{B}$ with R_2=0.53 and 0.06, respectively.

The R_2 is defined as a ratio of 2^{nd} to 0^{th} Fox-Wolfram moments [53]:

$$R_2 \equiv H_2/H_0,$$

where the i^{th} Fox-Wolfram moment H_i is defined as

$$H_i \equiv \sum_{j,k} \frac{|\boldsymbol{p}_j^*||\boldsymbol{p}_k^*|}{E^2} P_i(\cos\phi_{jk}),$$

where the indices j, k run over all tracks in the event, $\boldsymbol{p}_j^*$ is the momentum of the track j in the cms, E is the total energy of the event in the cms, ϕ_{jk} is the angle between the track j and k, and $P_i(x)$ is the i^{th} Legendre polynomial[d].

[d] $P_0(x) = 1, P_1(x) = x, P_2(x) = (3x^2 - 1)/2$

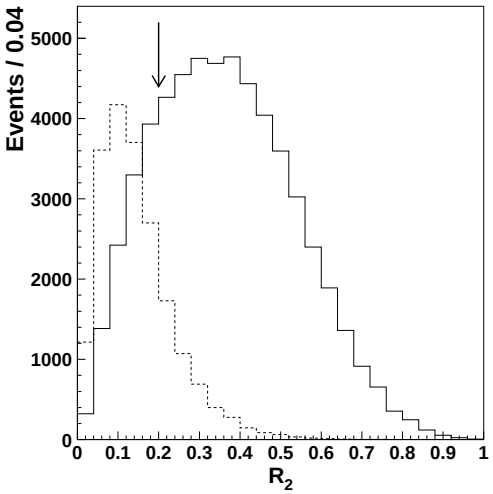


Figure 4.6: R_2 distribution: a solid (dotted) line is a shape of the continuum ($B\overline{B}$) event. By the $R_2 > 0.2$ cut, 77% of tracks in continuum events remain while 23% of tracks in $B\overline{B}$ events.

Using MC samples we check the fraction of secondary D mesons from $B\overline{B}$ events which remain after all selection criteria are applied. The result is shown in Table 4.3. The "fraction of D candidates" is the percentage of D candidates from $B\overline{B}$ events in all D candidates. The "fraction of true D" is the percentage of true D mesons from $B\overline{B}$ events in all true D mesons. The remaining secondary D mesons are negligible.

Table 4.3: Fraction of secondary D mesons after cuts.

Mode	Fraction of D candidates (%)	Fraction of true D (%)
$D^0 \to K^-\pi^+$ w/o D^*-tag	0.8	0.01
$D^0 \to K^-\pi^+$ with D^*-tag	0.08	0.0
$D^0 \to K^-K^+$	0.6	0.01
$D^+ \to K^-\pi^+\pi^+$	0.7	0.005
$D^+ \to \phi\pi^+$	3.2	0.0
$D_s^+ \to \phi\pi^+$	1.3	0.0
$D_s^+ \to \overline{K}^{*0}K^+$	1.1	0.0

In two-body D decay modes, a decay angle cut is imposed. A decay angle $\theta_{\rm Dec}$ is the angle between the momentum of one daughter in the D rest frame and that of D meson in the laboratory frame. For signal events the $\cos\theta_{\rm Dec}$ distribution is flat, while the background from random pions is peaked toward $\cos\theta_{\rm Dec} = -1$ (or $\cos\theta_{\rm Dec} = 1$,

depending on which daughter we choose) because of the low momentum of the pions as shown in Figure 4.7(a).

(a) Decay angle (b) Helicity angle

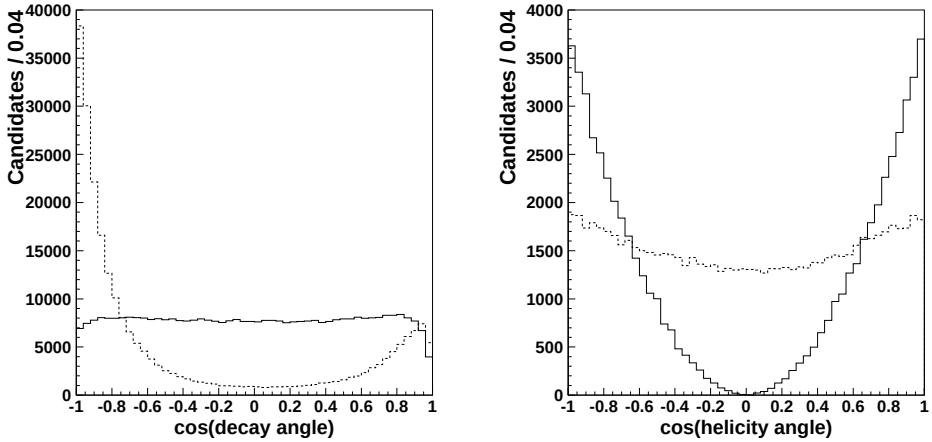


Figure 4.7: Cosine distribution of the decay angle in $D^0 \to K^-\pi^+$ and helicity angle in $D_s^+ \to \phi\pi^+$: a solid line is signal while a dotted is background.

In a $P \to PV$ decay[e], the vector meson is polarized, decaying to two pseudo-scalars $V \to PP$ according to a $\cos^2\theta_{\rm Hel}$ distribution, where $\theta_{\rm Hel}$ is the helicity angle between one of the pseudo-scalar daughters and the D direction evaluated in the rest frame of the vector particle. Figure 4.7(b) shows the $\cos\theta_{\rm Hel}$ distributions for signal and background. Table 4.4 lists the cuts imposed on the decay angle, while Table 4.5 on the helicity angle.

Table 4.4: Decay angle cuts

Mode	Decay angle cut
$D^0 \to K^-\pi^+$	$\cos\theta_{\pi,D^0} > -0.85$
$D^0 \to K^-K^+$	$\lvert\cos\theta_{K,D^0}\rvert < 0.90$
$D^+ \to \phi\pi^+$	$\cos\theta_{\pi,D^+} > -0.60$
$D_s^+ \to \phi\pi^+$	$\cos\theta_{\pi,D_s^+} > -0.70$
$D_s^+ \to \overline{K}^{*0}K^+$	$\lvert\cos\theta_{K,D_s^+}\rvert < 0.90$

[e] P = Pseudo-scalar particle, V = Vector particle.

Table 4.5: Helicity angle cuts

Mode	Helicity angle cut
$D^+ \to \phi\pi^+$	$\lvert\cos\theta_{K,D^+}\rvert > 0.4$
$D_s^+ \to \phi\pi^+$	$\lvert\cos\theta_{K,D_s^+}\rvert > 0.4$
$D_s^+ \to \overline{K}^{*0}K^+$	$\lvert\cos\theta_{K,D_s^+}\rvert > 0.4$

When we have multiple D candidates that share the same pion tracks, D candidates with the best kaon particle-identification are selected. We allow multiple D candidates with the same kaon track and different pion tracks.

To select final D candidates we apply the cuts related to the vertex fits which will be explained in Section 4.3. Figures 4.8(a)–4.8(f) show D^0, D^+ and D_s^+ mass distributions after all selection criteria are applied. Figures 4.8(a)–4.8(d) are for the lifetime measurement with 11.1 fb^{-1} and Figures 4.8(e) and 4.8(f) for the y_{CP} measurement with 23.4 fb^{-1}. The points are data while the solid curve is the result of a fit. Signal is represented by a single Gaussian or a sum of two Gaussians with the same mean while background is represented by a linear function. The background function is indicated by the dotted line in the figures. The number of the signal candidates determined from the fit results is summarized in Tables 4.6 and 4.7. Tables 4.8 and 4.9 show the fitted parameters of the mass distributions. Here $\sigma_{\rm n}$ and $\sigma_{\rm w}$ are a value of the standard deviation of narrow and wide Gaussians, respectively. The "ratio" is the ratio of the area occupied by the narrow Gaussian to all (narrow+wide). The $\sigma_{\rm av}$ is the weighted average of the standard deviations of the two Gaussians. The "purity" is calculated from the number of signal candidates divided by the number of signal and background candidates in the mass region within $3\sigma_{\rm av}$ of the peak of the measured D mass distribution. We call this mass region a signal region.

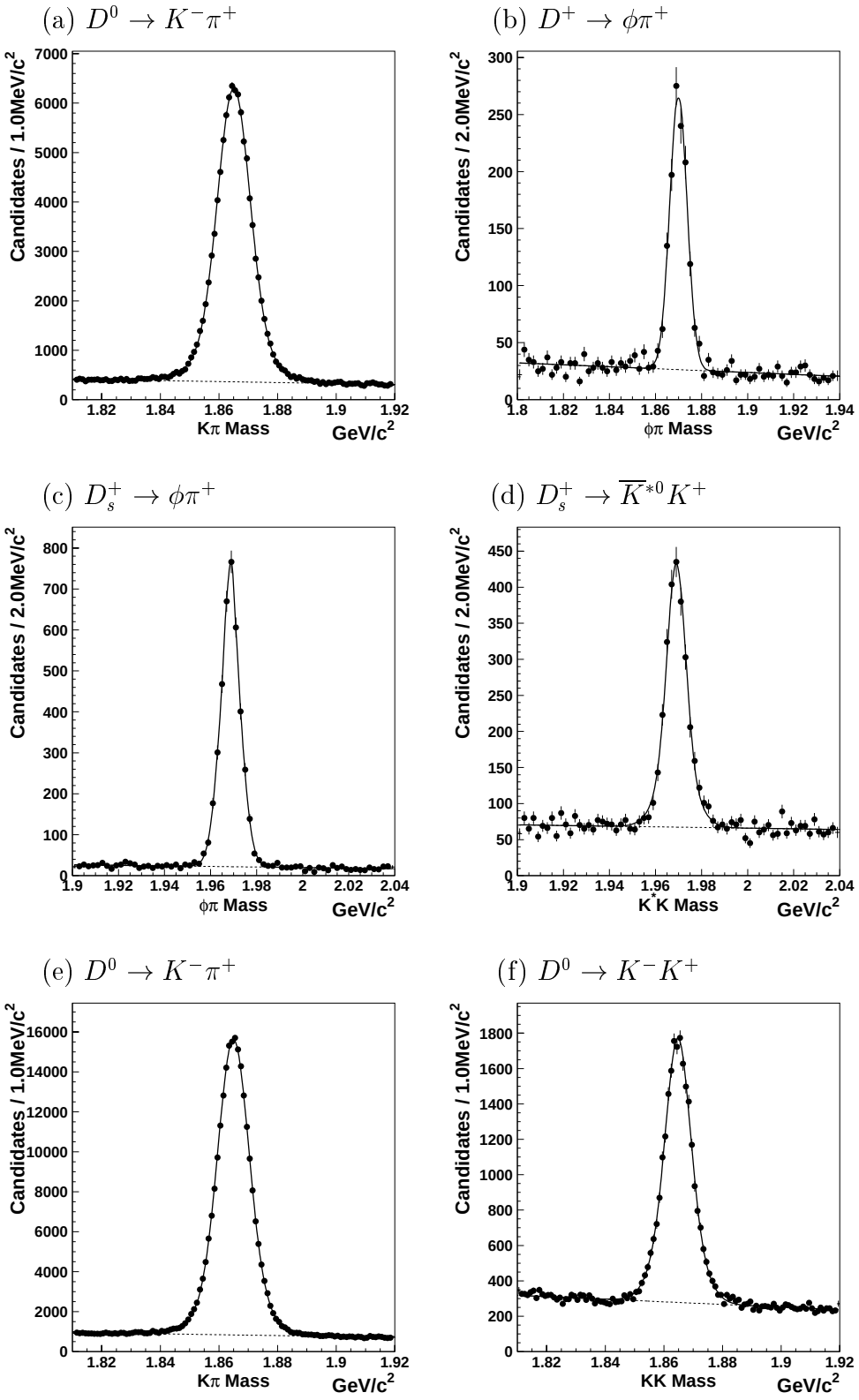


Figure 4.8: Mass distributions for (a) $D^0 \to K^-\pi^+$, (b) $D^+ \to \phi\pi^+$, (c) $D_s^+ \to \phi\pi^+$, (d) $D_s^+ \to \overline{K}^{*0}K^+$ (e) $D^0 \to K^-\pi^+$ and (f) $D^0 \to K^-K^+$ candidates. The distributions shown here are obtained from 11.1 fb^{-1} of data except for (e) and (f), which are obtained from 23.4 fb^{-1}of data.

Table 4.6: Number of the signal candidates and the purity determined from the fit results for $D^0 \to K^-\pi^+$, $D^+ \to \phi\pi^+$, $D_s^+ \to \phi\pi^+$ and $D_s^+ \to \overline{K}^{*0}K^+$ candidates. The definition of the purity is described in the text. They are obtained from 11.1 fb^{-1} data.

Mode	Number of signal	Purity
$D^0 \to K^-\pi^+$	90806 ± 380	0.85
$D^+ \to \phi\pi^+$	1130 ± 37	0.79
$D_s^+ \to \phi\pi^+$	3747 ± 64	0.93
$D_s^+ \to \overline{K}^{*0}K^+$	2179 ± 65	0.66

Table 4.7: Number of the signal candidates and the purity from the fit results for $D^0 \to K^-\pi^+$ and $D^0 \to K^-K^+$ candidates. They are obtained from 23.4 fb^{-1} data.

Mode	Number of signal	Purity
$D^0 \to K^-\pi^+$	214260 ± 562	0.87
$D^0 \to K^-K^+$	18306 ± 189	0.67

Table 4.8: Peak and widths obtained from the fit of the mass distributions for $D^0 \to K^-\pi^+$, $D^+ \to \phi\pi^+$, $D_s^+ \to \phi\pi^+$ and $D_s^+ \to \overline{K}^{*0}K^+$. The definitions of these parameters are described in the text. They are obtained from 11.1 fb^{-1} data.

Mode	Peak (MeV/c^2)	σ_n (MeV/c^2)	σ_w (MeV/c^2)	Ratio (%)	σ_av (MeV/c^2)
$D^0 \to K^-\pi^+$	1865.06 ± 0.02	5.4 ± 0.1	11.4 ± 0.5	79.4 ± 1.8	7.1
$D^+ \to \phi\pi^+$	1870.0 ± 0.1	3.7 ± 0.1	-	-	3.7
$D_s^+ \to \phi\pi^+$	1968.6 ± 0.1	1.7 ± 0.4	4.6 ± 0.1	8.6 ± 3.5	4.4
$D_s^+ \to \overline{K}^{*0}K^+$	1969.0 ± 0.1	3.9 ± 0.4	7.5 ± 1.8	63.8 ± 19.3	5.5

Table 4.9: Peak and widths obtained from the fit of the mass distributions for $D^0 \to K^-\pi^+$ and $D^0 \to K^-K^+$. They are obtained from 23.4 fb^{-1} data.

Mode	Peak (MeV/c^2)	σ_n (MeV/c^2)	σ_w (MeV/c^2)	Ratio (%)	σ_av (MeV/c^2)
$D^0 \to K^-\pi^+$	1864.94 ± 0.01	5.2 ± 0.5	9.8 ± 0.3	77.5 ± 1.8	6.5
$D^0 \to K^-K^+$	1864.80 ± 0.05	3.9 ± 0.3	6.7 ± 0.6	54.1 ± 13.7	5.4

4.2.8 D^* Meson Selection

We require the D^* tagging to obtain high-purity sub-samples of D mesons. D^{*+} candidates are formed by combining D^0 or D^+ candidates with low momentum pions π^+ or π^0, respectively. The mass differences $\Delta M = M_{D^0\pi^+} - M_{D^0}$ or $\Delta M = M_{D^+\pi^0} - M_{D^+}$ become narrow due to the small Q value, where $Q \equiv \Delta M - M_\pi$. When we have multiple pion candidates which share the same D^0 or D^+, the candidate with the ΔM closer to the world average value is retained. In the case of D^0, the ΔM resolution is greatly improved by recalculating the momentum of the pion with a constraint that it originates from the D^0 production point.

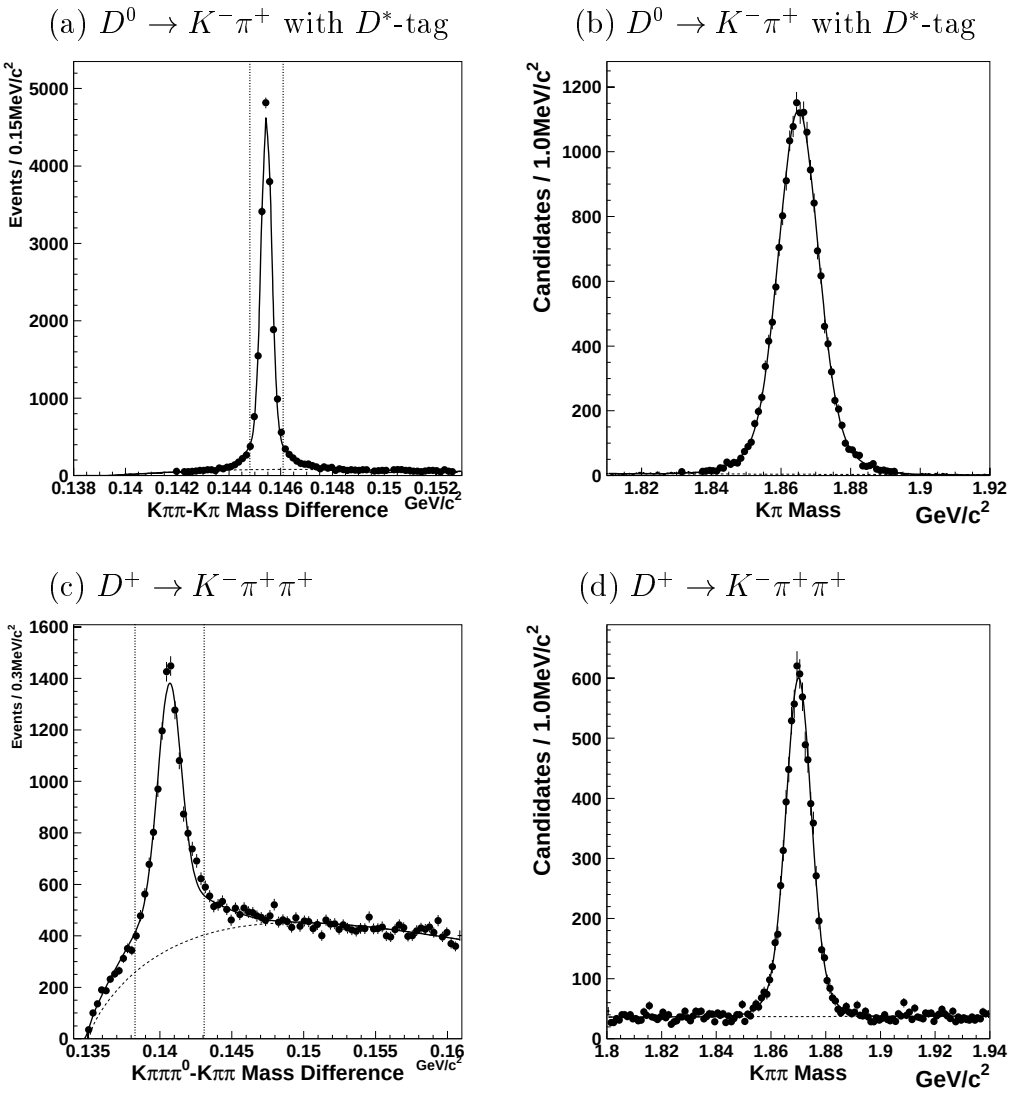


Figure 4.9: Mass difference distributions for (a) $D^0 \to K^-\pi^+$ and (c) $D^+ \to K^-\pi^+\pi^+$ candidates and mass distributions for (b) $D^0 \to K^-\pi^+$ and (d) $D^+ \to K^-\pi^+\pi^+$ candidates. They are obtained from 11.1 fb^{-1} data. The two dotted lines in (a) and (c) represent $\pm 3\sigma$ cuts on the ΔM distributions.

Figures 4.9(a)–4.9(d) shows the ΔM distribution for D^{*+} candidates and D candidate

mass distribution after the ΔM cut. The standard deviations of ΔM are approximately 0.2 MeV/c^2 for D^0, 0.8 MeV/c^2 for D^+ modes. Signal is represented by a sum of two Gaussians with the same mean while background is represented by a phase-space function, $(x - x_0)^p e^{a-b(x-x_0)}$ for the ΔM distribution. For the D mass distribution, the same parameterization is employed as described in Section 4.2.7. We require that the mass differences be within the regions listed in Table 4.10. The number of the signal events and values of the peak and widths determined from the fit are summarized in Tables 4.11 and 4.12.

Table 4.10: Values of ΔM cut for $D^0 \to K^-\pi^+$ and $D^+ \to K^-\pi^+\pi^+$.

Mode	ΔM Cut (MeV/c^2)
$D^0 \to K^-\pi^+$	± 0.64
$D^+ \to K^-\pi^+\pi^+$	± 2.4

Table 4.11: Number of the signal candidates and the purity for $D^0 \to K^-\pi^+$ and $D^+ \to K^-\pi^+\pi^+$. They are obtained from 11.1 fb^{-1} data.

Mode	Number of signal	Purity
$D^0 \to K^-\pi^+$	17285 ± 138	0.99
$D^+ \to K^-\pi^+\pi^+$	6950 ± 99	0.85

Table 4.12: Peak and widths obtained from the fit of mass distributions of $D^0 \to K^-\pi^+$ and $D^+ \to K^-\pi^+\pi^+$. They are obtained from 11.1 fb^{-1} data.

Mode	Peak (MeV/c^2)	$\sigma_{\rm n}$ (MeV/c^2)	$\sigma_{\rm w}$ (MeV/c^2)	Ratio (%)	$\sigma_{\rm av}$ (MeV/c^2)
$D^0 \to K^-\pi^+$	1865.0 ± 0.1	5.5 ± 0.1	11.5 ± 0.6	79.5 ± 2.7	7.1
$D^+ \to K^-\pi^+\pi^+$	1870.1 ± 0.1	4.4 ± 0.2	8.7 ± 1.3	79.2 ± 7.4	5.6

4.3 Decay Length Reconstruction

4.3.1 Flight Length Reconstruction

The charm meson lifetime is measured using the three-dimensional distance between the decay and production vertices of the reconstructed charm mesons. Figure 4.10 illustrates the reconstruction of the decay and production vertices of the charm meson.

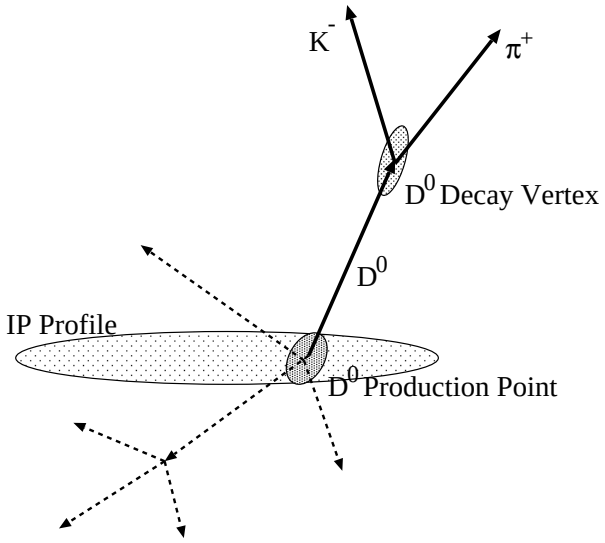


Figure 4.10: Event topology in an $e^+e^- \to c\overline{c}$ decay.

The decay vertex of the reconstructed charm meson is obtained using tracks that form the charm meson candidate. We use all charged tracks that form the charm meson candidate assuming ϕ and $\overline{K}^{*0}$ mesons decay at the decay vertex of the charm meson. In the fit for finding a vertex, all charged tracks are required to originate from a common vertex as described in Appendix D. The $\chi^2/N.D.F.$ of the fit is required to be less than 3 to remove poorly reconstructed candidates. This corresponds to $83-88\%$ efficiency.

The production vertex is obtained by extrapolating the D flight path to the interaction region. The extrapolation is performed by a fit taking into account the error of momentum and the spread of the interaction region. The information on the interaction region is represented by the IP profile as described in Appendix C. The $\chi^2/N.D.F.$ of the fit is required to be less than 20 to remove very poorly reconstructed candidates. This corresponds to $\sim 100\%$ efficiency since most poorly reconstructed events are rejected by the $\chi^2/N.D.F.$ cut of the decay vertex. The low momentum pion track is not used to find the production vertex in the $D^{*+} \to D^0\pi^+$ analysis.

The flight length ℓ is calculated from

$$\ell \equiv (\boldsymbol{v}_{\rm dec} - \boldsymbol{v}_{\rm prod}) \cdot \frac{\boldsymbol{p}_D}{|\boldsymbol{p}_D|},$$

where $\boldsymbol{v}_{\rm dec}, \boldsymbol{v}_{\rm prod}$, and $\boldsymbol{p}_D$ are decay vertex, production vertex and momentum of D in the laboratory frame, respectively. Figure 4.11 shows the relation between the sign of ℓ and the measurements: $\boldsymbol{v}_{\rm dec}, \boldsymbol{v}_{\rm prod}$, and $\boldsymbol{p}_D$. When the direction of $\boldsymbol{p}_D$ is opposite to that of $(\boldsymbol{v}_{\rm dec} - \boldsymbol{v}_{\rm prod})$ due to finite detector resolutions, the flight length ℓ has a negative sign. For D^+ and D_s^+ candidates, the helical flight path is not taken into account since the effect is found to be negligible.

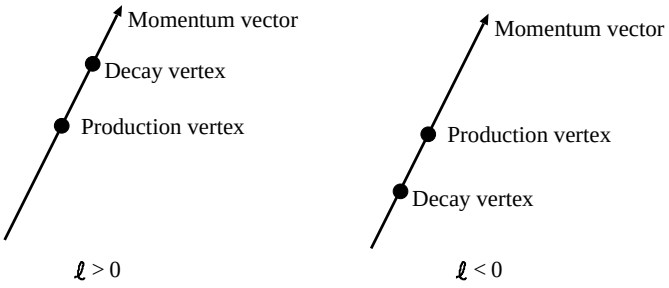


Figure 4.11: Sign of the flight length.

Figures 4.12 shows the calculated error of the flight length (σ_ℓ) for some decay modes. The averages of the calculated error of the flight length are listed in Table 4.13. We require σ_ℓ for a candidate to be less than 300 μm to reject poorly reconstructed candidates. This corresponds to $\sim$98% efficiency.

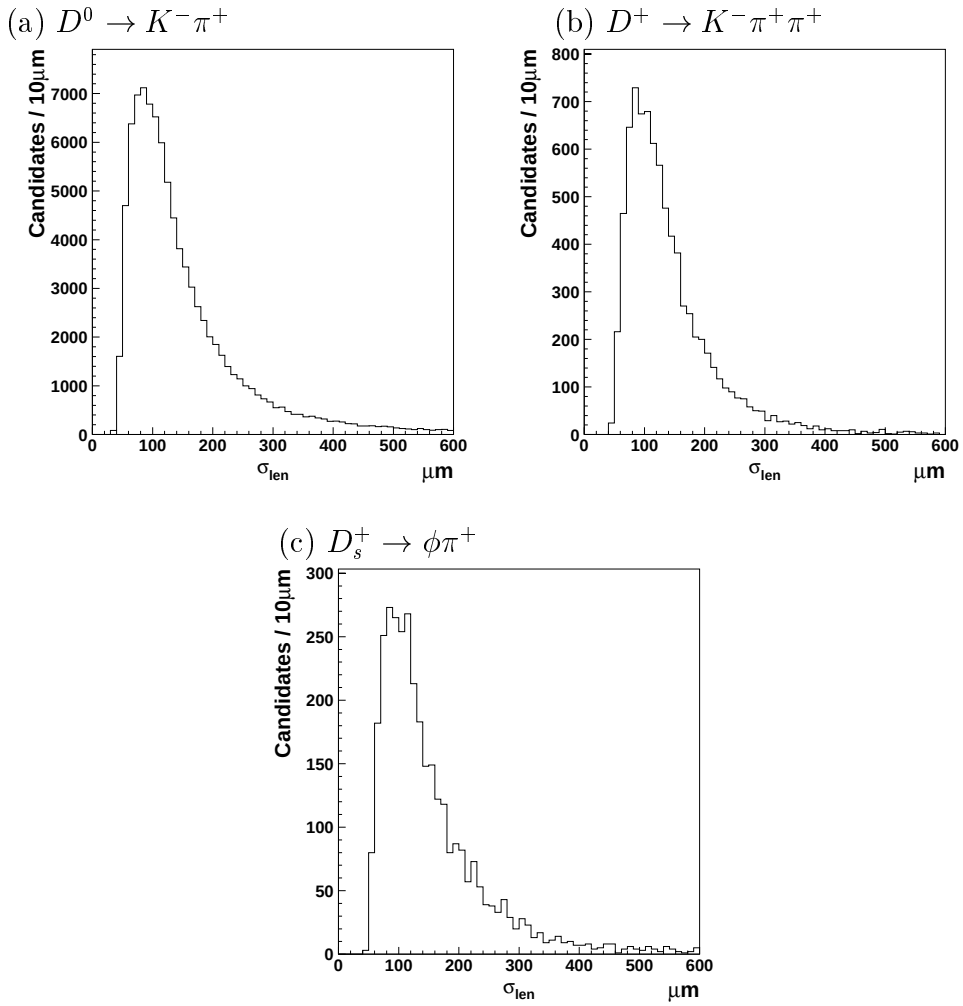


Figure 4.12: σ_ℓ distributions for (a) $D^0 \to K^-\pi^+$, (b) $D^+ \to K^-\pi^+\pi^+$ and (c) $D_s^+ \to \phi\pi^+$ candidates.

Table 4.13: Average of σ_ℓ of the selected events in the signal region before the σ_ℓ cut.

Mode	$\langle\sigma_\ell\rangle$ (μm)
$D^0 \to K^-\pi^+$	141
$D^0 \to K^-K^+$	141
$D^+ \to K^-\pi^+\pi^+$	124
$D^+ \to \phi\pi^+$	153
$D_s^+ \to \phi\pi^+$	140
$D_s^+ \to \overline{K}^{*0}K^+$	128

4.3.2 Proper-time Calculation

For each D candidate, the proper-time t between the production and decay is calculated as

$$t = \frac{\ell}{\beta\gamma c} = \frac{\ell m_D}{|\boldsymbol{p}_D|} \tag{4.1}$$

where ℓ, $\boldsymbol{p}_D$ and m_D are the flight length of the D meson, the momentum vector of the reconstructed D meson and the world average value of the D mass [6].

A proper-time t is a function of the decay vertex, production vertex and D momentum vector, that is,

$$t = t(\boldsymbol{x} \equiv (\boldsymbol{v}_{\rm dec}, \boldsymbol{v}_{\rm prod}, \boldsymbol{p}_D)).$$

Therefore the uncertainty σ_t on the proper-time is calculated from

$$\sigma_t^2 = \left(\frac{\partial t}{\partial \boldsymbol{x}}\right)^T E_{\boldsymbol{x}} \left(\frac{\partial t}{\partial \boldsymbol{x}}\right).$$

$E_{\boldsymbol{x}}$ consists of error matrices as follows.

$$E_{\boldsymbol{x}} = \begin{pmatrix} E_{\rm dec} & cov({\rm dec}, {\rm prod}) & cov({\rm dec}, p) \\ cov({\rm prod}, {\rm dec}) & E_{\rm prod} & cov({\rm prod}, p) \\ cov(p, {\rm dec}) & cov(p, {\rm prod}) & E_p \end{pmatrix}$$

$E_{\rm dec}$, $E_{\rm prod}$ and E_p are the error matrices of the D decay vertex, the production vertex and the D momentum, respectively. A $cov(A, B)$ is a correlation matrix between A and B. These correlation matrices are obtained from a vertex fit with the IP profile, that is, finding the production vertex.

Table 4.14 shows averages of the decay length ℓ, proper-time t and its error σ_t after applying all selection cuts. $\langle\sigma_t\rangle$ is $\sim$ 220 fs in nearly all decay modes while $\langle t\rangle$ depends on the amount of the background events in each mode.

Table 4.14: Average of p, ℓ, t and σ_t of the selected events in the signal region.

Mode	$\langle p \rangle$ (GeV/c)	$\langle \ell \rangle$ (μm)	$\langle t \rangle$ (fs)	$\langle \sigma_t \rangle$ (fs)
$D^0 \to K^-\pi^+$	3.6	213	366	213
$D^0 \to K^-K^+$	3.6	182	314	225
$D^+ \to K^-\pi^+\pi^+$	3.5	501	897	219
$D^+ \to \phi\pi^+$	3.5	505	886	244
$D_s^+ \to \phi\pi^+$	3.7	255	455	230
$D_s^+ \to \overline{K}^{*0}K^+$	3.6	222	390	228

Chapter 5

Lifetime Fit Method

We use an unbinned maximum likelihood fit to extract lifetimes. In this chapter, we describe the fitting method including the definition of the probability density functions.

5.1 Unbinned Maximum Likelihood Fit Method

The general description of an unbinned maximum likelihood method is given here. We assume that the experimental result x obeys a function $\mathcal{P}(\boldsymbol{\alpha}, x)$, that is, when we obtain many experimental results x, the distribution of x becomes close to a function $\mathcal{P}(\boldsymbol{\alpha}, x)$, where $\boldsymbol{\alpha}$ is a set of the parameters to vary the shape of $\mathcal{P}(\boldsymbol{\alpha}, x)$. A function $\mathcal{P}(\boldsymbol{\alpha}, x)$ is a normalized distribution:

$$\int \mathcal{P}(\boldsymbol{\alpha}, x) dx = 1,$$

where $\mathcal{P}(\boldsymbol{\alpha}, x)$ is called a probability density function.

The likelihood function $\mathcal{L}(\boldsymbol{\alpha})$ is the joint probability density of the experimental results x_i $(i = 1, ..., N)$:

$$\mathcal{L}(\boldsymbol{\alpha}) = \prod_{i=1}^{N} \mathcal{P}(\boldsymbol{\alpha}, x_i),$$

where $\boldsymbol{\alpha}$ is called fitting parameters. We try to find the most probable value of $\boldsymbol{\alpha}$ as the solution which gives the maximum value of $\mathcal{L}(\boldsymbol{\alpha})$ when varying the value of $\boldsymbol{\alpha}$. This procedure is called an unbinned maximum likelihood fit [55]. We use the MINUIT library [56] to execute maximization of $\mathcal{L}(\boldsymbol{\alpha})$.

It is important to find a functional form of $\mathcal{P}(\boldsymbol{\alpha}, x)$ which should represent a distribution of our measurement.

5.2 Definition of Probability Density Function

We determine the functional form of the probability density function to extract the lifetimes. We assume that the probability density function $\mathcal{P}$ is given by the linear combi-

nation of signal and background functions as follows,

$$\mathcal{L} = \prod_{i=\text{all candidates}} \mathcal{P}_i = \prod_{i=\text{all candidates}} \mathcal{P}_i(t^i, \sigma_t^i, f_{\rm SIG}^i)$$
$$\equiv \prod_{i=\text{all candidates}} f_{\rm SIG}^i \mathcal{F}_{\rm SIG}^i(t^i, \sigma_t^i) + (1 - f_{\rm SIG}^i)\mathcal{F}_{\rm BG}^i(t^i, \sigma_t^i).$$

where $\mathcal{P}_i$ is a function of the measured values (t^i, σ_t^i and $f_{\rm SIG}^i$) for each candidate as described below. In the following sections, we discuss in detail each form of the functions, $\mathcal{F}_{\rm SIG}^i$ (signal function), $\mathcal{F}_{\rm BG}^i$ (background function), and $f_{\rm SIG}^i$ (signal probability). Finally we summarize the functional form of the likelihood function to measure the lifetimes.

5.2.1 Signal Function

We represent a signal function $\mathcal{F}_{\rm SIG}^i$ by a true decay distribution $\frac{1}{\tau}e^{\frac{-t}{\tau}}$ convoluted with a resolution function $R_{\rm SIG}$;

$$\mathcal{F}_{\rm SIG}^i(t^i, \sigma_t^i) \equiv \int_0^\infty dt' \frac{1}{\tau_{\rm SIG}} e^{\frac{-t'}{\tau_{\rm SIG}}} R_{\rm SIG}(t^i - t', \sigma_t^i),$$

where t^i and σ_t^i are the measured proper-time and its error for each candidate. This resolution $R_{\rm SIG}$ depends on t^i and σ_t^i.

Resolution Function

The parameterization of the resolution function is described in this section. Figure 5.1 shows the distribution of $t_{\rm rec} - t_{\rm gen}$ obtained from MC data, where $t_{\rm rec}$ is the reconstructed proper-time of the candidate and $t_{\rm gen}$ is the true proper-time. A fit to a sum of two Gaussians is also shown. However the fit is not a good representation of the distribution. The vertex resolution depends on track momenta, angles and other factors, and therefore it varies candidate by candidate. The distribution actually consists of many Gaussians. We therefore calculate the error on the proper-time candidate by candidate taking into account the resolution of each hit, multiple scattering and energy loss. Figure 5.2 shows the distribution of $(t_{\rm rec} - t_{\rm gen})/\sigma_t$ where σ_t is the calculated proper-time error for each candidate. The distribution is well represented by a fit with a sum of two Gaussians, shown in the figure by the overlaid dashed curve. The standard deviation of the main Gaussian will be 1 if the error estimation is correct. The fit yields 1.1. This difference is considered to be due to underestimation of the tracking error. The remaining candidates, represented by the broader Gaussian, are considered to be those which are poorly measured due to mis-association of SVD hits, incorrect SVD-hit clustering, large angle multiple scattering of tracks and so on.

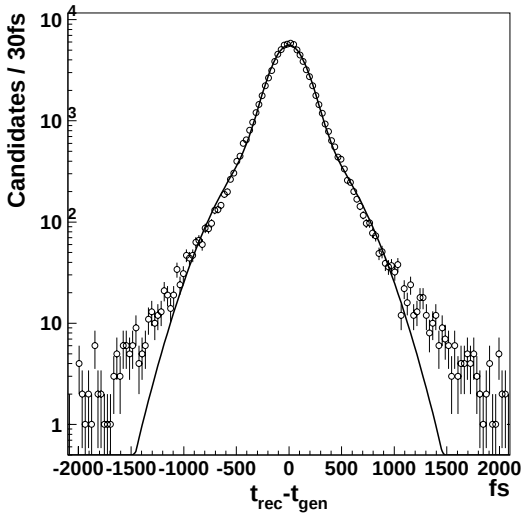


Figure 5.1: Resolution of proper-time for $D^0 \to K^-\pi^+$.

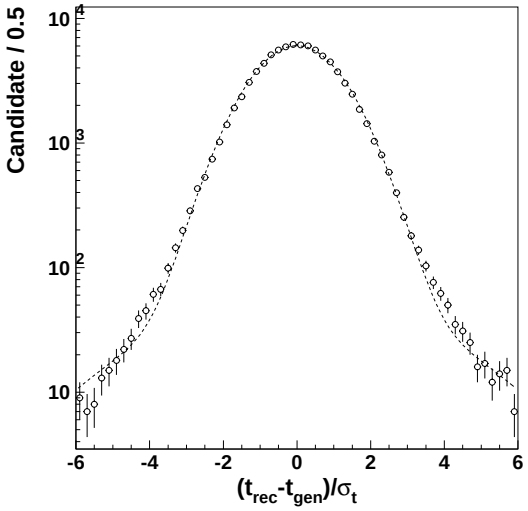


Figure 5.2: Pull distribution of proper-time for $D^0 \to K^-\pi^+$.

Based on the success of this fit, we represent the proper-time resolution using the

function:

$$\begin{aligned} R_{\rm SIG}(t,\sigma_t) &\equiv R_{\rm SIG}(t,\sigma_t,S,S_{\rm tail},f_{\rm tail}) \\ &\equiv (1-f_{\rm tail})\cdot G(t;S\sigma_t) + f_{\rm tail}\cdot G(t;S_{\rm tail}\sigma_t), \\ G(x;\sigma) &\equiv \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, \end{aligned}$$

where S and $S_{\rm tail}$ are the global scaling factors for the estimated error σ_t for the main and tail Gaussian distributions, and $f_{\rm tail}$ is the fraction of the tail part. The S, $S_{\rm tail}$, and $f_{\rm tail}$ are output parameters of the likelihood fit. Figure 5.3 shows the σ_t distribution for signal candidates in the Monte Carlo (MC) simulation. The distribution has a long tail. We evaluate the effect of those candidates with large σ_t on the resolution function by dividing the signal candidates into three samples according to σ_t: one sample with $\sigma_t < 120$ fs (lower side of the peak), another sample with $120 < \sigma_t < 350$ fs (higher side of the peak) and the third sample with $\sigma_t > 350$ fs (the tail). We fit the $t_{\rm rec} - t_{\rm gen}$ distribution with a resolution function for each of these three samples as shown in Figures 5.4. Table 5.1 lists the fit results: the parameters determined for the three σ_t regions are comparable.

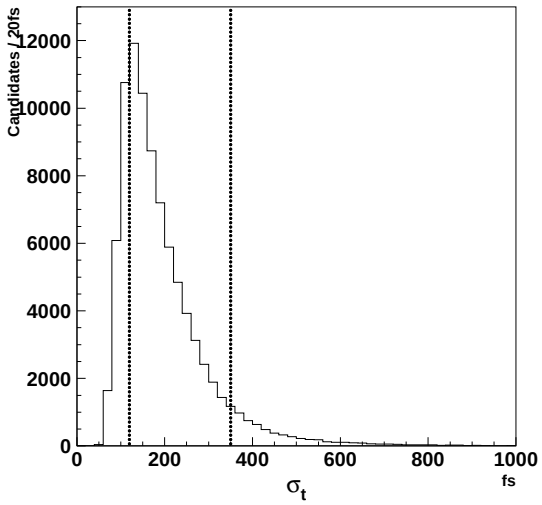


Figure 5.3: σ_t distribution for $D^0 \to K^-\pi^+$.

Table 5.1: Fit results of pull distributions for $D^0 \to K^-\pi^+$.

Region (fs)	S	$S_{\rm tail}$	$f_{\rm tail}$
$\sigma_t < 120$	1.14 ± 0.02	4.01 ± 0.31	0.018 ± 0.005
$120 < \sigma_t < 350$	1.14 ± 0.01	3.19 ± 0.20	0.027 ± 0.005
$\sigma_t > 350$	1.09 ± 0.02	3.01 ± 0.32	0.040 ± 0.013
all	1.13 ± 0.01	3.27 ± 0.16	0.026 ± 0.004

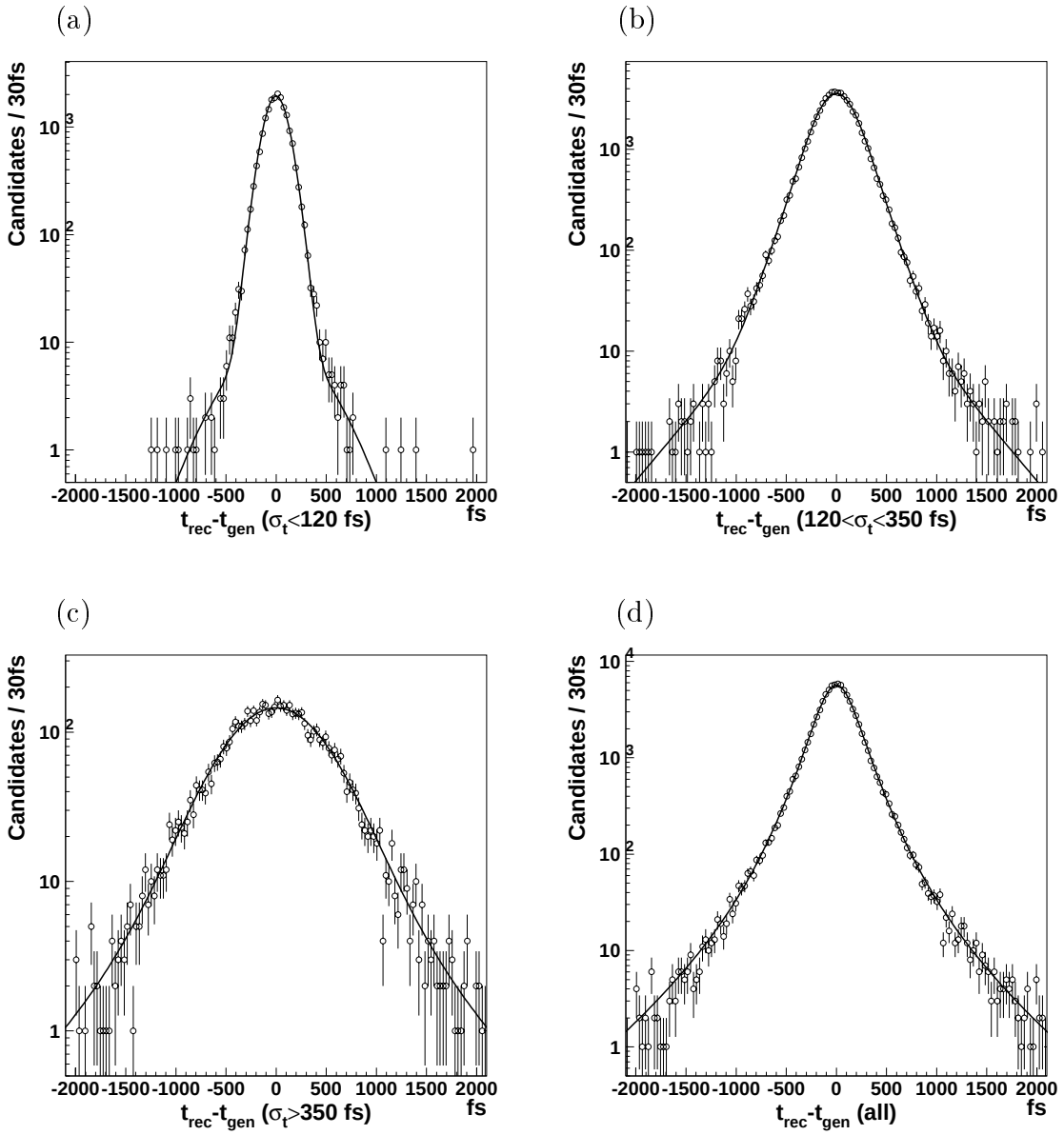


Figure 5.4: Resolutions of proper-time for $D^0 \to K^-\pi^+$ at each region: (a) $\sigma_t < 120$, (b) $120 < \sigma_t < 350$, (c) $350 < \sigma_t$ and (d) all σ_t.

In the analysis presented here we use the same S, $S_{\rm tail}$, and $f_{\rm tail}$ for the resolution function, regardless of the σ_t value. The effect of ignoring the small σ_t dependence of the resolution function parameters has been assessed by performing a lifetime fit for signal candidates in the Monte Carlo simulation. The fit to MC data yields a lifetime 410.3 ± 0.7 fs, which is consistent with the true lifetime of the Monte Carlo sample, 410.0 ± 0.7 fs. We conclude that the σ_t dependence of the resolution function parameters does not significantly affect the lifetime fit.

5.2.2 Background Function

We express a background function $\mathcal{F}^i_{\rm BG}$ as an appropriate background proper-time distribution $\mathcal{T}_{\rm BG}$ convoluted with a resolution function $R_{\rm BG}$;

$$\mathcal{F}^i_{\rm BG}(t^i, \sigma^i_t) \equiv \int_0^\infty dt' \mathcal{T}_{\rm BG}(t') R_{\rm BG}(t^i - t', \sigma^i_t).$$

As described in Section 4.2.7, our background events consist of the continuum events. The combinatorial backgrounds from the $c\overline{c}$ events are caused by daughter particles of D mesons. The background proper-time distribution $\mathcal{T}_{\rm BG}(t)$ is therefore modeled by a fraction $f_{\tau_{\rm BG}}$ with effective lifetime $\tau_{\rm BG}$ and a fraction $(1 - f_{\tau_{\rm BG}})$ with zero lifetime represented by the Dirac delta function $\delta(t)$;

$$\mathcal{T}_{\rm BG}(t) \equiv f_{\tau_{\rm BG}} \frac{1}{\tau_{\rm BG}} e^{\frac{-t}{\tau_{\rm BG}}} + (1 - f_{\tau_{\rm BG}})\delta(t).$$

We test the modeling of background using a Monte Carlo simulation. Background contributions can be categorized into several groups according to the origin of charged tracks that are used in the decay vertex reconstruction. In the case of $D^0 \to K^-\pi^+$ with D^*-tag, we do not study background compositions since the purity is very high ($\simeq 0.99$). Tables 5.2 and 5.3 summarize the definition of the groups for the case of two daughter charged tracks (D^0) and three daughter charged tracks (D^+, D^+_s), respectively.

Tables 5.4 and 5.6 show fractions of the groups. Tables 5.5 and 5.7 list average proper-times of the background components. In the case of two daughter charged tracks, a zero lifetime component "c0p0" is dominant and other components can be represented by one lifetime term. Our model of the background terms is reasonable. Also in the case of three daughter charged tracks, a zero lifetime component "c0p0" is dominant except for $D^+ \to K^-\pi^+\pi^+$. Two non-zero lifetime components "c2p1" and "c1p0" have large contributions. Since they have different lifetimes, we would need two lifetimes in the background terms for $D^+ \to K^-\pi^+\pi^+$ and $D^+_s \to \phi\pi^+$. However we employ a model which has one lifetime term in the background. This model difference does not make a change in the fitted D^+ and D^+_s lifetimes for the Monte Carlo sample: 1058.9 ± 4.5 (one lifetime in background) $\to 1057.0 \pm 4.5$ fs (two lifetimes in background) in D^+ and 473.2 ± 2.8 (one lifetime in background) $\to 473.4 \pm 2.8$ fs (two lifetimes in background) in D^+_s. We also check it with the real data for $D^+ \to K^-\pi^+\pi^+$ and $D^+_s \to \phi\pi^+$ and the fitted lifetimes change only within $0.5\sigma_{\rm stat.}$. Therefore we use a model which has one lifetime term in the background in all decay modes.

Table 5.2: Groups for the case of two daughter charged tracks.

Type	Description
c2p1	tracks from the same charm meson.
c2p0	tracks from different charm mesons.
c1p0	a track from a charm meson and a track from the primary vertex.
c0p0	tracks from the primary vertex.

Table 5.3: Groups for the case of three daughter charged tracks.

Type	Description
c3p2	three tracks from the same charm meson.
c3p1	two tracks from the same charm meson and a track from a different charm meson.
c3p0	tracks from three different charm mesons.
c2p1	two tracks from the same charm meson and a track from the primary vertex.
c2p0	two tracks from different charm mesons and a track from the primary vertex.
c1p0	a track from a charm meson and two tracks from the primary vertex.
c0p0	three tracks from the primary vertex.

Table 5.4: Fractions (%) of background components of D^0 decay modes.

Mode	c2p1	c2p0	c1p0	c0p0
$D^0 \to K^-\pi^+$ w/o D^*-tag	0.01	0.9	23.1	76.0
$D^0 \to K^-K^+$	0.2	1.3	27.0	71.6

Table 5.5: $\langle t \rangle$ of the background components of D^0 decay modes.

Mode	c2p1 (fs)	c2p0 (fs)	c1p0 (fs)	c0p0 (fs)
$D^0 \to K^-\pi^+$ w/o D^*-tag	1002 ± 448	-207 ± 10	28.7 ± 0.3	4.59 ± 0.02
$D^0 \to K^-K^+$	655 ± 122	-116 ± 7	249 ± 3	7.74 ± 0.07

Table 5.6: Fractions (%) of background components of D^+ and D_s^+ decay modes.

Mode	c3p2	c3p1	c3p0	c2p1	c2p0	c1p0	c0p0
$D^+ \to K^-\pi^+\pi^+$	0.0	5.4	0.1	49.1	3.4	20.5	21.6
$D^+ \to \phi\pi^+$	2.2	1.6	0.0	35.0	0.5	7.0	53.6
$D_s^+ \to \phi\pi^+$	0.0	1.4	0.05	17.1	1.6	11.9	68.0
$D_s^+ \to \overline{K}^{*0}K^+$	0.0	1.8	0.0	29.0	1.1	7.5	60.6

Table 5.7: $\langle t \rangle$ of the background components of D^+ and D_s^+ decay modes.

Mode	c3p2 (fs)	c3p1 (fs)	c3p0 (fs)	c2p1 (fs)	c2p0 (fs)	c1p0 (fs)	c0p0 (fs)
$D^+ \to K^-\pi^+\pi^+$	–	357 ± 20	41 ± 15	330 ± 6	93 ± 7	182 ± 5	0.96 ± 0.03
$D^+ \to \phi\pi^+$	858 ± 117	-116 ± 19	–	199 ± 7	-241 ± 67	78 ± 6	1.75 ± 0.05
$D_s^+ \to \phi\pi^+$	–	142 ± 15	190 ± 110	408 ± 13	94 ± 10	158 ± 6	3.50 ± 0.05
$D_s^+ \to \overline{K}^{*0}K^+$	–	11 ± 2	–	92 ± 4	27 ± 6	128 ± 11	8.0 ± 0.2

We assume that the functional form of the background resolution $R_{\rm BG}$ is the same as that used for the signal, but a different set of fitting parameters are used, as a combination of tracks coming from different vertices gives a resolution worse than the signal.

$$\begin{aligned} R_{\rm BG}(t,\sigma_t) &\equiv R_{\rm BG}(t,\sigma_t,S_{\rm BG},S_{\rm tail}^{\rm BG},f_{\rm tail}^{\rm BG}) \\ &\equiv (1-f_{\rm tail}^{\rm BG})\cdot G(t;S_{\rm BG}\sigma_t) + f_{\rm tail}^{\rm BG}\cdot G(t;S_{\rm tail}^{\rm BG}\sigma_t). \end{aligned}$$

Table 5.8 compares fit parameters in the four D-mass regions. The results in the four regions are similar within 3σ, which supports the D-mass independence of the proper-time. In addition the correlations between them and the signal lifetime $\tau_{\rm SIG}$ are small in the result of the lifetime fit, e.g. $\pm 3\sigma$ variations of $f_{\tau_{\rm BG}}$, $\tau_{\rm BG}$ or $S_{\rm tail}^{\rm BG}$ affect less than $\pm 0.5\sigma$ of $\tau_{\rm SIG}$ in $D^0 \to K^-\pi^+$.

Table 5.8: Comparison of fit parameters of backgrounds in each region. The Δm_D (in units of MeV/c^2) is defined as the difference from the mass peak of the mass distribution.

	$-40 < \Delta m_D < -20$	$-20 < \Delta m_D < 0$	$0 < \Delta m_D < 20$	$20 < \Delta m_D < 40$	$-40 < \Delta m_D < 40$
$f_{\tau_{\rm BG}}$	0.059 ± 0.007	0.033 ± 0.006	0.026 ± 0.007	0.052 ± 0.007	0.045 ± 0.003
$\tau_{\rm BG}$	327 ± 25	445 ± 46	576 ± 84	326 ± 29	381 ± 18
$S_{\rm BG}$	1.177 ± 0.008	1.178 ± 0.008	1.175 ± 0.010	1.153 ± 0.009	1.174 ± 0.004
$S_{\rm tail}^{\rm BG}$	5.47 ± 0.24	4.69 ± 0.25	4.15 ± 0.36	4.65 ± 0.21	4.98 ± 0.18
$f_{\rm tail}^{\rm BG}$	0.022 ± 0.003	0.028 ± 0.004	0.033 ± 0.007	0.026 ± 0.003	0.024 ± 0.002

5.2.3 Signal Probability

The $f^i_{\rm SIG}$ is the probability that the candidate is a signal. The $f^i_{\rm SIG}$ is calculated as a function of the invariant mass of each candidate: $f^i_{\rm SIG} = f_{\rm SIG}(\Delta m^i_D)$.

The invariant mass distribution is fitted by a Gaussian or a sum of two Gaussians with the same mean for the D signal and a linear function for the background. Then $f_{\rm SIG}(\Delta m^i_D)$ is defined as

$$f_{\rm SIG}(\Delta m^i_D) \equiv \frac{F_{\rm SIG}(\Delta m^i_D)}{F_{\rm BG}(\Delta m^i_D) + F_{\rm SIG}(\Delta m^i_D)},$$

where $F_{\rm SIG}$ and $F_{\rm BG}$ are the invariant mass functions for the D signal and background, respectively.

Figure 5.5 shows the mass and signal probability distributions. The value of the signal probability becomes larger than 0.90 around the peak of the mass distribution while it becomes almost zero in the tail region of the mass distribution.

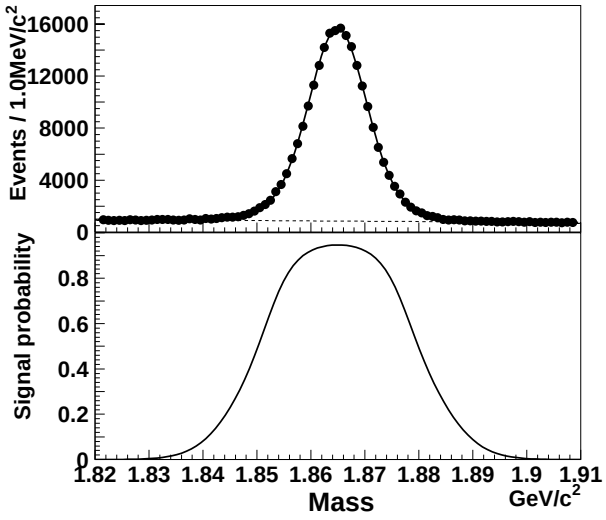


Figure 5.5: Relation between the mass distribution and the signal probability calculated for $D^0 \to K^-\pi^+$.

5.2.4 Form of the Likelihood

We summarize the functional form of the likelihood described in the previous sections. The probability density function $\mathcal{P}$ for each candidate consists of a signal term and the two background terms representing components of the background with non-zero lifetime

and zero lifetime respectively. The complete form of the likelihood function $\mathcal{L}$ is given by

$$
\begin{aligned}
\mathcal{L} &= \mathcal{L}(\tau_{\rm SIG}, S, S_{\rm tail}, f_{\rm tail}, \tau_{\rm BG}, f_{\tau_{\rm BG}}, S_{\rm BG}, S^{\rm BG}_{\rm tail}, f^{\rm BG}_{\rm tail}) \\
&= \prod_{i=\text{all candidates}} \mathcal{P}_i = \prod_i \mathcal{P}(t^i, \sigma_t^i, f^i_{\rm SIG}) \\
&= \prod_i [\underbrace{f^i_{\rm SIG} \int_0^\infty dt' \frac{1}{\tau_{\rm SIG}} e^{\frac{-t'}{\tau_{\rm SIG}}} R_{\rm SIG}(t^i - t', \sigma_t^i, S, S_{\rm tail}, f_{\rm tail})}_{\textbf{Signal term}} \\
&+ \underbrace{(1 - f^i_{\rm SIG}) f_{\tau_{\rm BG}} \int_0^\infty dt' \frac{1}{\tau_{\rm BG}} e^{\frac{-t'}{\tau_{\rm BG}}} R_{\rm BG}(t^i - t', \sigma_t^i, S_{\rm BG}, S^{\rm BG}_{\rm tail}, f^{\rm BG}_{\rm tail})}_{\textbf{BG term with lifetime}} \\
&+ \underbrace{(1 - f^i_{\rm SIG})(1 - f_{\tau_{\rm BG}}) \int_0^\infty dt' \delta(t') R_{\rm BG}(t^i - t', \sigma_t^i, S_{\rm BG}, S^{\rm BG}_{\rm tail}, f^{\rm BG}_{\rm tail})}_{\textbf{BG term without lifetime}}], \qquad (5.1)
\end{aligned}
$$

where t^i and σ_t^i are the measured proper-time and its error for each candidate. The $f^i_{\rm SIG}$ is the probability that the candidate is a signal. The $\tau_{\rm SIG}$ is the signal lifetime and the functions $R_{\rm SIG}$ and $R_{\rm BG}$ represent the resolution of the proper-time t^i. The background proper-time distribution is modeled by a fraction $f_{\tau_{\rm BG}}$ with effective lifetime $\tau_{\rm BG}$ and a fraction $(1 - f_{\tau_{\rm BG}})$ with zero lifetime represented by the Dirac delta function $\delta(t')$. The fraction $f^i_{\rm SIG}$ is calculated based on the D mass for each candidate, using a functional form $f_{\rm SIG}(\Delta m^i_D)$ derived from the data.

We summarize input and output parameters in Table 5.9. We have 9 fit parameters in the likelihood function.

Table 5.9: Input and output parameters of likelihood for the lifetime measurement.

Input parameter	Description
t^i	proper-time
σ_t^i	proper-time uncertainty
$f^i_{\rm SIG}$	signal probability (a function of the D mass Δm^i_D)
Output parameter	**Description**
$\tau_{\rm SIG}$	signal lifetime
S	scaling factor for σ_t (signal)
$S_{\rm tail}$	scaling factor for the tail part (signal)
$f_{\rm tail}$	fraction of the tail part (signal)
$\tau_{\rm BG}$	background lifetime
$f_{\tau_{\rm BG}}$	fraction of background with lifetime
$S_{\rm BG}$	scaling factor for σ_t (background)
$S^{\rm BG}_{\rm tail}$	scaling factor for the tail part (background)
$f^{\rm BG}_{\rm tail}$	fraction of the tail part (background)

Chapter 6

Lifetime and D^0-$\overline{D}^0$ Mixing Measurements

In this chapter, we show the proper-time distributions for the real data and the results of the likelihood fit explained in the previous chapter. First we present the results of the fit for each decay mode. We then describe the method of combining the likelihood functions for two decay modes. To obtain the final results, we apply a small correction which is discussed in the last section.

6.1 Lifetime Measurements of Each Decay Mode

We perform an unbinned likelihood fit described in the previous chapter on the D meson candidates given in Chapter 4. Only the candidates within ± 40 MeV/c^2 of the mass peak are used in the fit. We use different data samples for lifetime and y_{CP} measurements as described in Section 4.1. Figures 6.1(a) and 6.1(b) show the result of the lifetime fit of $D^0 \to K^-\pi^+$ without D^*-tag. For demonstration we prepare two plots of proper-time distributions, one is for the signal-dominated region and the other is for the background-dominated region. The signal-dominated region is defined as the mass region within $\pm 3\sigma_{\rm av}$ of the peak, where $\sigma_{\rm av}$ is the weighted average of the standard deviations of the two Gaussians in the mass distribution. The background-dominated region is defined as the outside of the signal-dominated region. Figures 6.2(a) and 6.2(b) show the average of resolution functions of the signal and background terms, respectively, which are defined as

$$\sum_{i=\text{signal region}} f^i_{\rm SIG} R_{\rm SIG}(t, \sigma^i_t, S, S_{\rm tail}, f_{\rm tail}) \ \text{(signal term)},$$

$$\sum_{i=\text{signal region}} (1 - f^i_{\rm SIG}) R_{\rm BG}(t, \sigma^i_t, S_{\rm BG}, S^{\rm BG}_{\rm tail}, f^{\rm BG}_{\rm tail}) \ \text{(background term)}.$$

Table 6.1 is the results of the fit to the curves in Figure 6.2 with two Gaussians with the same mean. Typical resolutions are 210 fs for signal term and 250 fs for background term. Figure 6.3 shows log-likelihood as a function of proper-time. It has a parabolic shape implying a Gaussian error for the D^0 lifetime measurement.

Figures 6.4, 6.5, 6.6, 6.7 and 6.8 show the results of the likelihood fit for $D^+ \to K^-\pi^+\pi^+$, $D^+ \to \phi\pi^+$, $D_s^+ \to \phi\pi^+$, $D_s^+ \to \overline{K}^{*0}K^+$ and $D^0 \to K^-\pi^+$ with D^*-tag, respectively. Table 6.2 summarizes the results of the fit parameters of each lifetime fit. We obtain $\tau(D^0 \to K^-\pi^+) = 417.9 \pm 1.7$ fs, $\tau(D^+ \to K^-\pi^+\pi^+) = 1028.9 \pm 13.1$ fs, $\tau(D^+ \to \phi\pi^+) = 1081.5 \pm 35.0$ fs, $\tau(D_s^+ \to \phi\pi^+) = 486.8 \pm 9.3$ fs, $\tau(D_s^+ \to \overline{K}^{*0}K^+) = 504.1 \pm 13.8$ fs, where the uncertainties are statistical only. The obtained lifetimes from two modes are consistent with each other.

Figures 6.9 and 6.10 show the results of $D^0 \to K^-\pi^+$ and $D^0 \to K^-K^+$ for the y_{CP} measurement. Table 6.3 summarizes the results of the fit parameters of each lifetime fit for the y_{CP} measurement. The size of the scaling factor S is less than 1, which indicates the overestimation of the proper-time error. When we calculate the average scaling factor $(1-f_{\rm tail})S+f_{\rm tail}S_{\rm tail}$, we obtain 0.9–1.1 for all decay modes. We also calculate the average scaling factor for the background term, and we obtain 1.2–1.4, which is larger than that of the signal term. The fraction of background with lifetime, $f_{\tau_{\rm BG}}$, is less than 0.2. This indicates that $u\overline{u}$, $d\overline{d}$ and $s\overline{s}$ events are dominant in the background events.

(a) $D^0 \to K^-\pi^+$ (signal region) (b) $D^0 \to K^-\pi^+$ (background region)

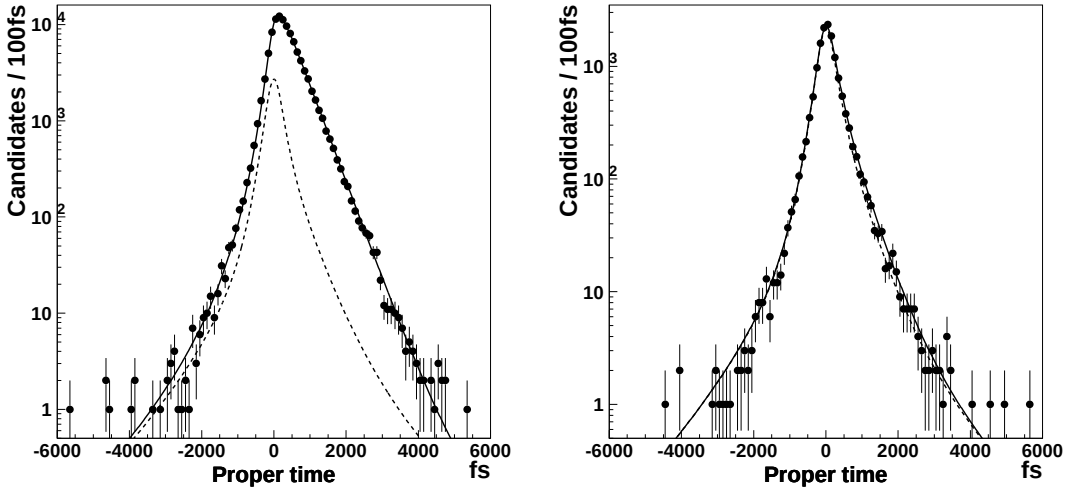
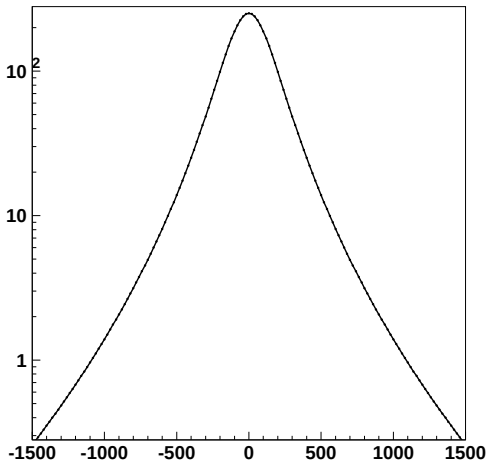


Figure 6.1: Proper-time distributions and fit results for $D^0 \to K^-\pi^+$ candidates in the lifetime measurement: (a) signal- and (b) background-dominated regions.

(a) Average of resolution functions for $D^0 \to K^-\pi^+$ (signal term)



(b) Average of resolution functions for $D^0 \to K^-\pi^+$ (background term)

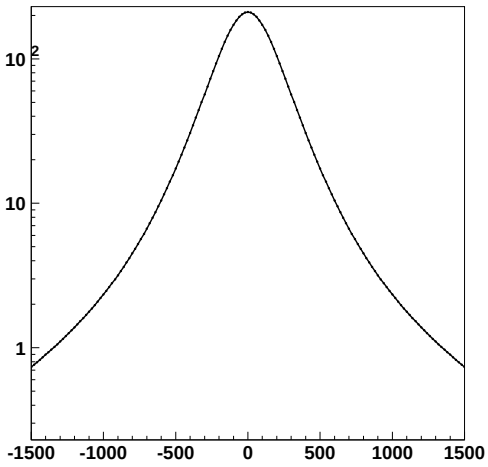


Figure 6.2: Average of the resolution functions obtained from the lifetime fit for $D^0 \to K^-\pi^+$ candidates: (a) signal and (b) background terms. The unit of the horizontal axis is "fs".

Table 6.1: Results of the fit to the curves in Figure 6.2. The $\sigma_{\rm n}$ and $\sigma_{\rm w}$ are the standard deviation of the narrow and wide Gaussians, respectively. The "fraction" is the ratio of the area occupied by the narrow Gaussian to all (=narrow+wide).
(a) Signal term $R_{\rm SIG}$

	$\sigma_{\rm n}$ (fs)	$\sigma_{\rm w}$ (fs)	Fraction	$\sigma_{\rm av}$ (fs)
$D^0 \to K^-\pi^+$ for lifetime	118	282	0.51	215
$D^0 \to K^-\pi^+$ for y_{CP}	113	274	0.50	210
$D^0 \to K^-K^+$	123	286	0.50	220
$D^+ \to K^-\pi^+\pi^+$	127	288	0.50	223
$D^+ \to \phi\pi^+$	167	354	0.52	273
$D_s^+ \to \phi\pi^+$	148	324	0.53	247
$D_s^+ \to \overline{K}^{*0}K^+$	106	257	0.41	209

(b) Background term $R_{\rm BG}$

	$\sigma_{\rm n}$ (fs)	$\sigma_{\rm w}$ (fs)	Fraction	$\sigma_{\rm av}$ (fs)
$D^0 \to K^-\pi^+$ for lifetime	145	339	0.54	253
$D^0 \to K^-\pi^+$ for y_{CP}	139	328	0.53	247
$D^0 \to K^-K^+$	151	345	0.53	261
$D^+ \to K^-\pi^+\pi^+$	151	342	0.53	259
$D^+ \to \phi\pi^+$	172	379	0.52	290
$D_s^+ \to \phi\pi^+$	193	407	0.55	308
$D_s^+ \to \overline{K}^{*0}K^+$	166	359	0.52	276

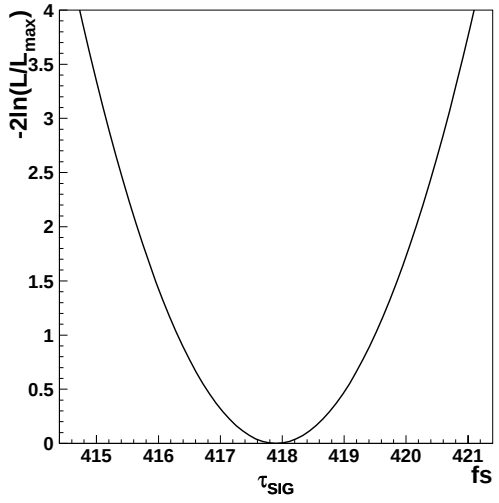


Figure 6.3: Log-likelihood versus proper-time obtained from the $D^0 \to K^-\pi^+$ lifetime fit.

(a) $D^+ \to K^-\pi^+\pi^+$ (signal region) (b) $D^+ \to K^-\pi^+\pi^+$ (background region)

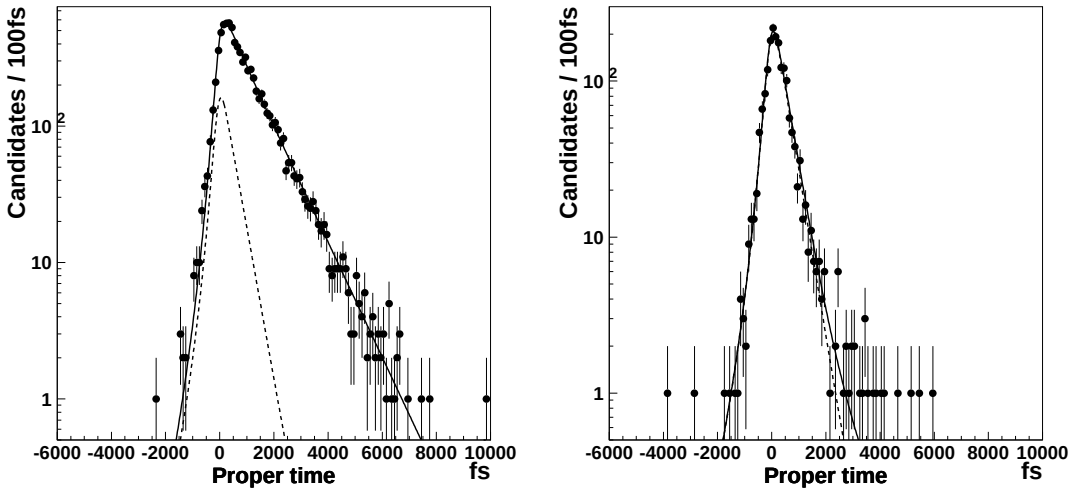


Figure 6.4: Proper-time distributions and fit results for $D^+ \to K^-\pi^+\pi^+$ candidates in the lifetime measurement: (a) signal- and (b) background-dominated regions.

(a) $D^+ \to \phi\pi^+$ (signal region) (b) $D^+ \to \phi\pi^+$ (background region)

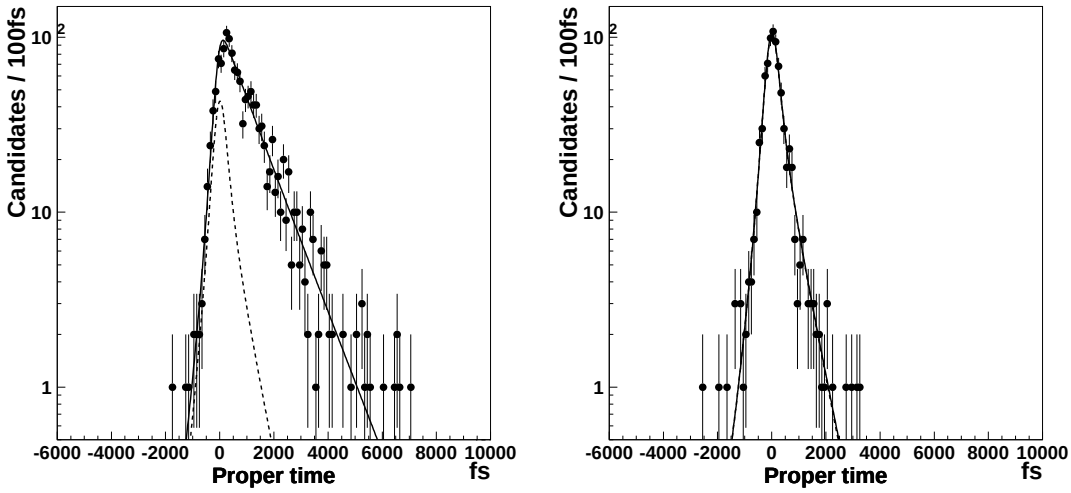


Figure 6.5: Proper-time distributions and fit results for $D^+ \to \phi\pi^+$ candidates in the lifetime measurement: (a) signal- and (b) background-dominated regions.

(a) $D_s^+ \to \phi\pi^+$ (signal region) (b) $D_s^+ \to \phi\pi^+$ (background region)

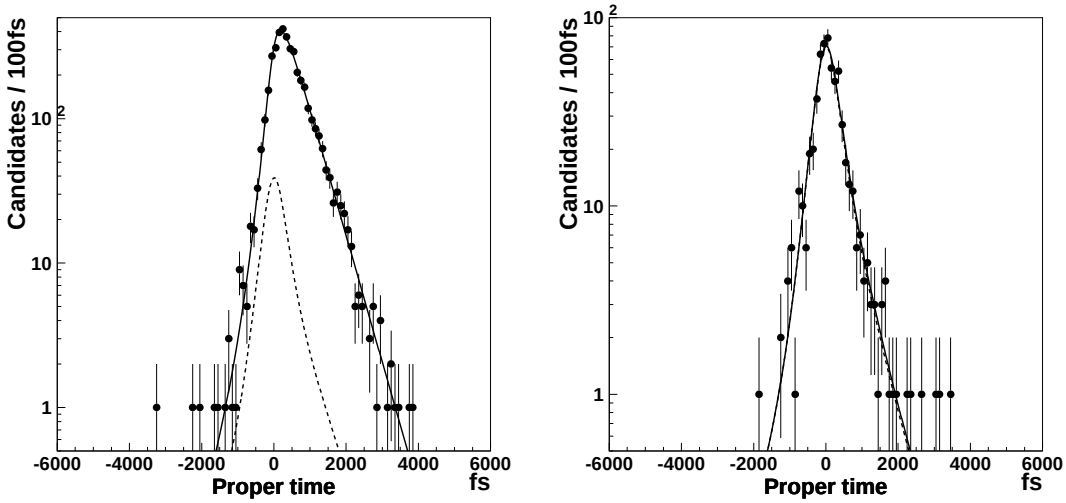


Figure 6.6: Proper-time distributions and fit results for $D_s^+ \to \phi\pi^+$ candidates in the lifetime measurement: (a) signal- and (b) background-dominated regions.

(a) $D_s^+ \to \overline{K}^{*0}K^+$ (signal region) (b) $D_s^+ \to \overline{K}^{*0}K^+$ (background region)

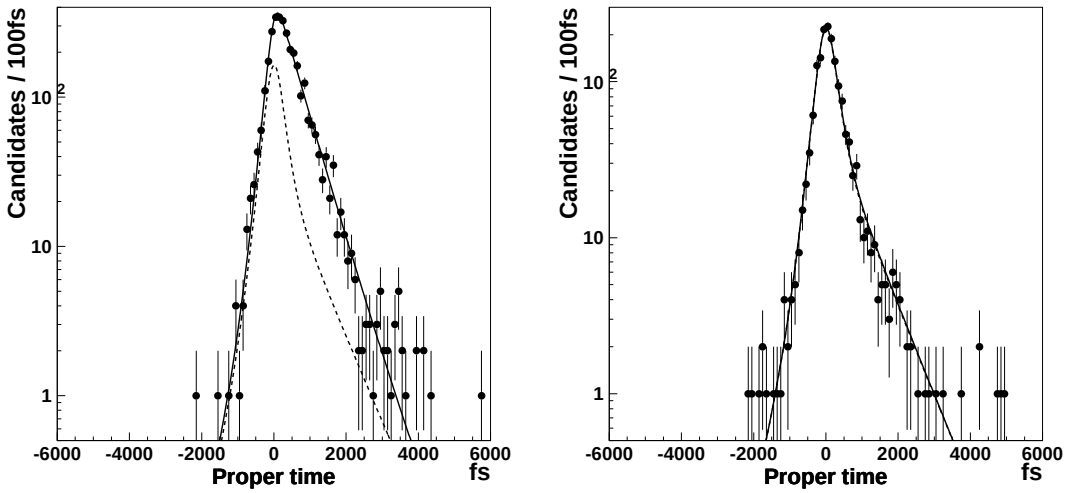


Figure 6.7: Proper-time distributions and fit results for $D_s^+ \to \overline{K}^{*0}K^+$ candidates in the lifetime measurement: (a) signal- and (b) background-dominated regions.

(a) $D^0 \to K^-\pi^+$ with D^*-tag (signal region)

(b) $D^0 \to K^-\pi^+$ with D^*-tag (background region)

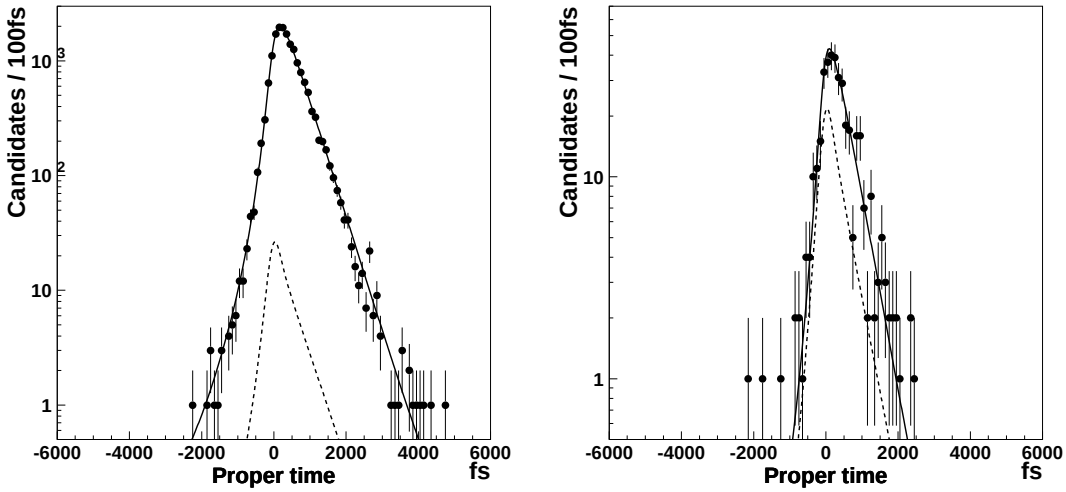


Figure 6.8: Proper-time distributions and fit results for $D^0 \to K^-\pi^+$ candidates with D^*-tag in the lifetime measurement: (a) signal- and (b) background-dominated regions.

(a) $D^0 \to K^-\pi^+$ (signal region) (b) $D^0 \to K^-\pi^+$ (background region)

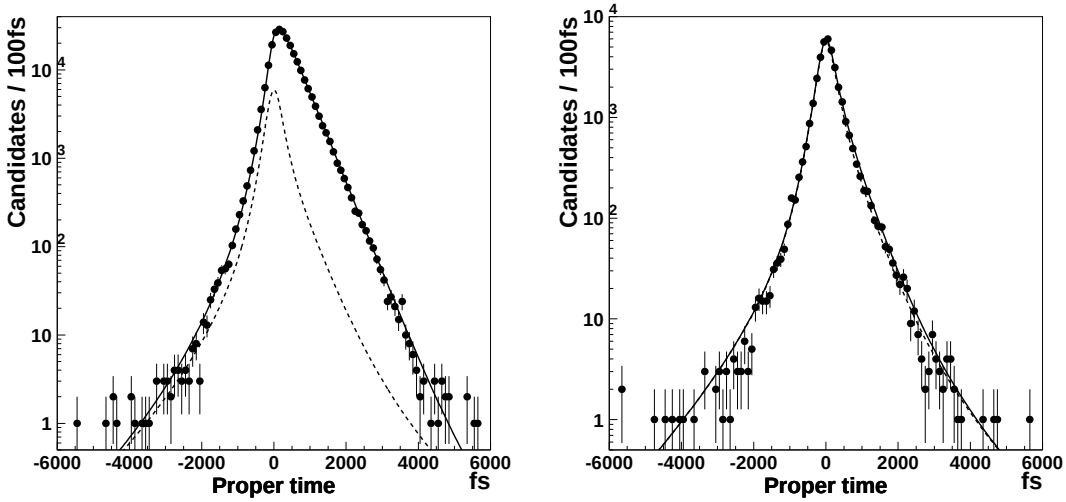


Figure 6.9: Proper-time distributions and fit results for $D^0 \to K^-\pi^+$ candidates in the y_{CP} measurement: (a) signal- and (b) background-dominated regions. We use different data samples for lifetime and y_{CP} measurements.

(a) $D^0 \to K^-K^+$ (signal region) (b) $D^0 \to K^-K^+$ (background region)

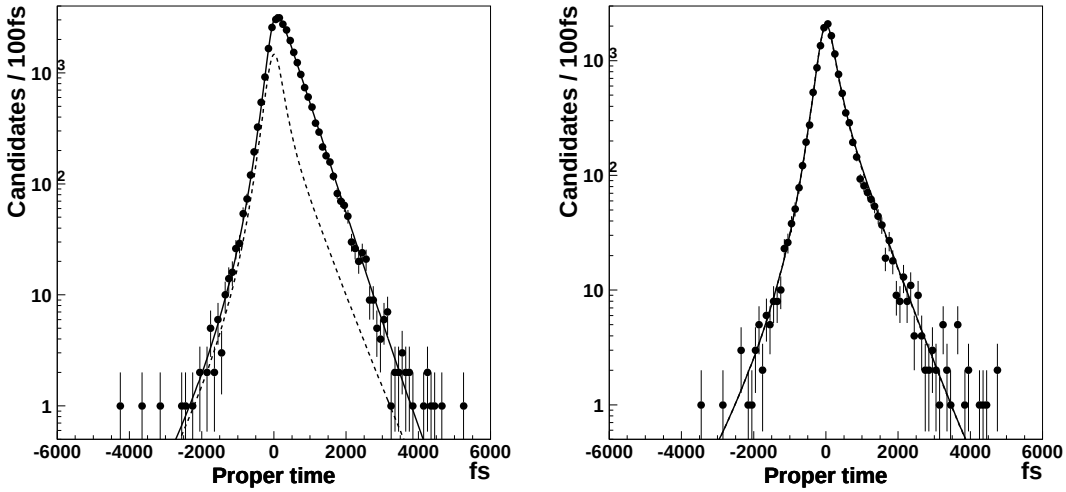


Figure 6.10: Proper-time distributions and fit results for $D^0 \to K^-K^+$ candidates in the y_{CP} measurement: (a) signal- and (b) background-dominated regions. We use different data samples for lifetime and y_{CP} measurements.

Table 6.2: Summary of the fit parameters obtained from the lifetime fit in the lifetime measurement. For the $D^0 \to K^-\pi^+$ with D^*-tag, the parameters in the $R_{\rm BG}$ cannot be determined because of the high purity of the data sample. The parameters are therefore constrained to be the same as those in $R_{\rm SIG}$.

(a) Fit parameters for D^0.

	$D^0 \to K^-\pi^+$	$D^0 \to K^-\pi^+$ with D^*-tag
$\tau_{\rm SIG}$	417.9 ± 1.7 fs	419.2 ± 3.7 fs
S	0.89 ± 0.01	1.01 ± 0.02
$S_{\rm tail}$	1.94 ± 0.08	3.26 ± 0.21
$f_{\rm tail}$	0.15 ± 0.02	0.041 ± 0.008
$\tau_{\rm BG}$	402 ± 25 fs	450 ± 108 fs
$f_{\tau_{\rm BG}}$	0.12 ± 0.01	0.53 ± 0.11
$S_{\rm BG}$	1.10 ± 0.01	-
$S_{\rm tail}^{\rm BG}$	5.00 ± 0.18	-
$f_{\rm tail}^{\rm BG}$	0.055 ± 0.005	-

(b) Fit parameters for D^+.

	$D^+ \to K^-\pi^+\pi^+$ with D^*-tag	$D^+ \to \phi\pi^+$
$\tau_{\rm SIG}$	1028.9 ± 13.1 fs	1081.5 ± 35.0 fs
S	0.83 ± 0.13	1.01 ± 0.60
$S_{\rm tail}$	1.62 ± 0.45	1.86 ± 8.26
$f_{\rm tail}$	0.23 ± 0.23	0.12 ± 0.60
$\tau_{\rm BG}$	364 ± 24 fs	544 ± 70 fs
$f_{\tau_{\rm BG}}$	0.59 ± 0.03	0.23 ± 0.04
$S_{\rm BG}$	1.00 ± 0.06	1.02 ± 0.07
$S_{\rm tail}^{\rm BG}$	2.46 ± 0.36	2.03 ± 0.36
$f_{\rm tail}^{\rm BG}$	0.12 ± 0.06	0.17 ± 0.10

(c) Fit parameters for D_s^+.

	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \overline{K}^{*0}K^+$
$\tau_{\rm SIG}$	486.8 ± 9.3 fs	504.1 ± 13.8 fs
S	0.96 ± 0.06	0.59 ± 0.30
$S_{\rm tail}$	2.16 ± 0.41	1.13 ± 0.20
$f_{\rm tail}$	0.10 ± 0.06	0.60 ± 0.57
$\tau_{\rm BG}$	520 ± 142 fs	768 ± 58 fs
$f_{\tau_{\rm BG}}$	0.17 ± 0.05	0.20 ± 0.02
$S_{\rm BG}$	1.26 ± 0.06	1.10 ± 0.03
$S_{\rm tail}^{\rm BG}$	5.97 ± 1.15	2.75 ± 0.32
$f_{\rm tail}^{\rm BG}$	0.03 ± 0.02	0.07 ± 0.02

Table 6.3: Summary of the fit parameters obtained from the lifetime fit in the y_{CP} measurement.

	$D^0 \to K^-\pi^+$	$D^0 \to K^-K^+$
$\tau_{\rm SIG}$	416.2 ± 1.1 fs	417.0 ± 4.2 fs
S	0.84 ± 0.01	0.84 ± 0.03
$S_{\rm tail}$	1.79 ± 0.06	1.73 ± 0.18
$f_{\rm tail}$	0.17 ± 0.02	0.16 ± 0.05
$\tau_{\rm BG}$	391 ± 16 fs	497 ± 16 fs
$f_{\tau_{\rm BG}}$	0.15 ± 0.01	0.21 ± 0.01
$S_{\rm BG}$	1.05 ± 0.01	1.05 ± 0.01
$S^{\rm BG}_{\rm tail}$	4.15 ± 0.10	3.22 ± 0.16
$f^{\rm BG}_{\rm tail}$	0.062 ± 0.004	0.063 ± 0.008

6.2 Lifetime and D^0-$\overline{D}^0$ Mixing Measurements: D^+, D_s^+ and y_{CP}

The D^0-$\overline{D}^0$ mixing parameter y_{CP} and lifetimes of D^+ and D_s^+ are each obtained from two decay modes. We use a combined likelihood instead of combining lifetime results of each decay mode since this method allows us to properly estimate correlated systematic errors, e.g. on the decay length, as described in Chapter 7.

The combined likelihoods to obtain y_{CP} and lifetimes are defined as

$$\begin{aligned}
\mathcal{L}_{y_{CP}} &= \prod_{i=\text{all candidates}} \mathcal{P}_i \\
&= \prod_{i=K\pi} \mathcal{P}_i^{K^-\pi^+} \cdot \prod_{i=KK} \mathcal{P}_i^{K^-K^+}, \qquad (6.1) \\
\mathcal{L}_{D^+} &= \prod_{i=\text{all candidates}} \mathcal{P}_i \\
&= \prod_{i=K\pi\pi} \mathcal{P}_i^{K^-\pi^+\pi^+} \cdot \prod_{i=\phi\pi} \mathcal{P}_i^{\phi\pi^+}, \qquad (6.2) \\
\mathcal{L}_{D_s^+} &= \prod_{i=\text{all candidates}} \mathcal{P}_i \\
&= \prod_{i=\phi\pi} \mathcal{P}_i^{\phi\pi^+} \cdot \prod_{i=K^{*0}K} \mathcal{P}_i^{\overline{K}^{*0}K^+}, \qquad (6.3)
\end{aligned}$$

where the definition of each $\mathcal{P}$ is the same as that of Eq. (5.1).

The relations between y_{CP} and D^0 lifetime are obtained from Eq. (2.18):

$$\tau^{KK}_{\rm SIG} = \tau^{K\pi}_{\rm SIG}/(1 + y_{CP}) \qquad (6.4)$$

We replace the $\tau^{KK}_{\rm SIG}$ of Eq. (6.1) by that of Eq. (6.4). The y_{CP} and $\tau^{K\pi}_{\rm SIG}$ are fitted simultaneously in the unbinned maximum likelihood fit.

In Eq. (6.2) and Eq. (6.3), we impose the following constraints:

$$\tau_{\rm SIG}^{K\pi\pi} = \tau_{\rm SIG}^{\phi\pi},$$
$$\tau_{\rm SIG}^{\phi\pi} = \tau_{\rm SIG}^{K^{*0}K}.$$

Therefore we have 18 fit parameters for the y_{CP} measurement and 17 for D^+ and D_s^+ lifetime measurements.

Table 6.4 shows the results of the fit and Figures 6.11 and 6.12 show the log-likelihood as a function of proper-time and y_{CP}, respectively. They have a parabolic shape implying a Gaussian error for the lifetime and y_{CP} measurements.

Table 6.4: Summary of main fit parameters obtained from the combined lifetime fit.

	D^0	D^+	D_s^+
$\tau_{\rm SIG}$	416.2 ± 1.1	1036 ± 12	492.0 ± 7.8
y_{CP}	-0.002 ± 0.010	–	–

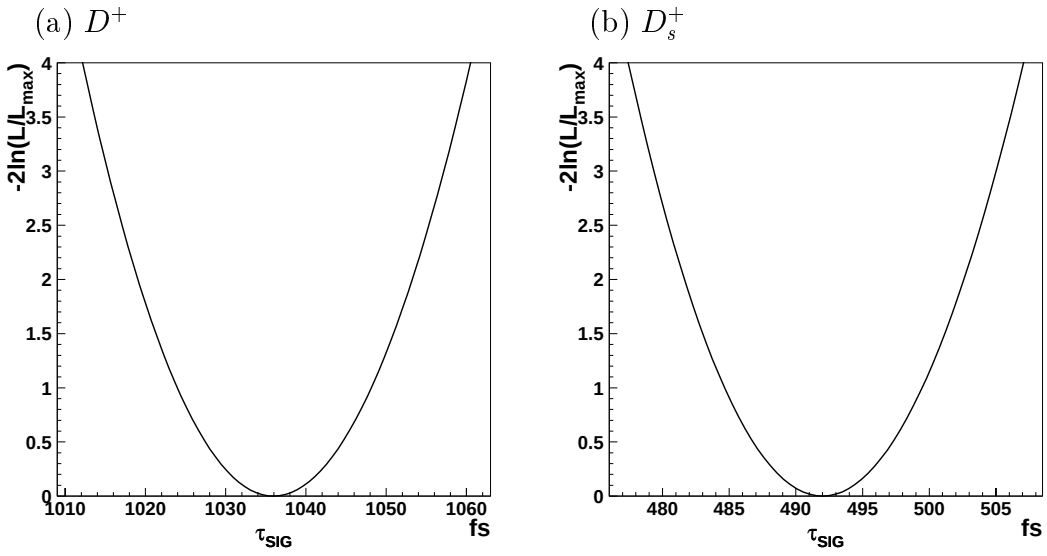


Figure 6.11: Log-likelihood versus proper-time: (a) D^+ and (b) D_s^+.

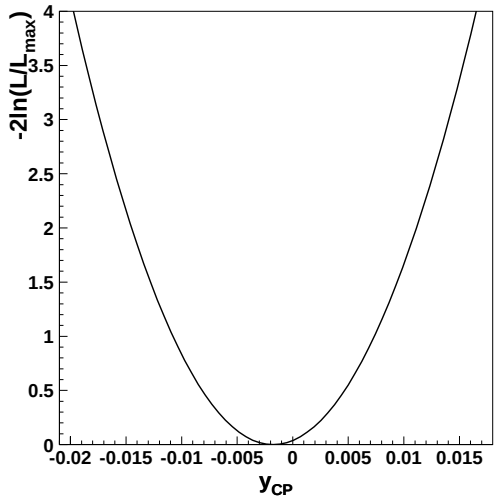


Figure 6.12: Log-likelihood versus y_{CP}.

We compare the results of the combined likelihood fits with the average of the lifetimes of two decay modes. The averaged lifetime $\langle\tau\rangle$ is obtained by the formula,

$$\langle\tau\rangle = \frac{\sum_i \omega_i \tau_i}{\sum_i \omega_i},$$

where $\omega_i = 1/\sigma_i^2$. Table 6.5 shows the results of the two methods and we can see that the two methods give nearly identical results. We use the result of the combined likelihood method for lifetimes and y_{CP} since it allow us to estimate properly correlated systematic uncertainties such as errors in decay length and interaction point.

Small corrections derived from an extensive MC simulation study will be applied to obtain the final results.

Table 6.5: Comparison of the methods for combining results for the lifetime and y_{CP} measurements.

	Combined likelihood method	Average method
D^+	1036 ± 12	1035 ± 12
D_s^+	492.0 ± 7.8	492.2 ± 7.7
y_{CP}	-0.002 ± 0.010	-0.002 ± 0.010

6.3 Corrections

A study based on large Monte Carlo (MC) samples ($\sim$30–50 fb^{-1}) indicates that there remains a small bias in the reconstruction and fit procedures.

6.3.1 Lifetime Measurement

Table 6.6 shows the results of the lifetime fit to MC samples. A small difference is found between the lifetime used in the MC generation ($\tau_{\rm gen}$) and the result of the fit ($\tau_{\rm SIG}$). The origin of the bias can be broken down to three sources, that is,

- Event selection,
- Reconstruction of proper-time, and
- Fitting.

We may have a bias in our event selection criteria. In the reconstruction of proper-time we could have a bias due to multiple scattering and other physics reasons. We also would have a bias in the fitting if the modeling of signal and background distributions is not complete.

The value of ($\tau_{\rm SIG} - \tau_{\rm gen}$) includes these three effects as described later in detail. We correct the lifetimes for real data with values of ($\tau_{\rm SIG} - \tau_{\rm gen}$) in Table 6.6.

6.3.2 y_{CP} Measurement

Table 6.7 shows the results of the lifetime fit to MC samples for the y_{CP} measurement. Since y_{CP} is a ratio of lifetimes, some of biases described above would cancel or be negligible. We study each bias in detail.

Event Selection

We investigate a possible bias in event selection. We use a MC sample of $D^0 \to K^-\pi^+$. Figure 6.13 shows the distribution of the generated proper-times for the selected events. We fit the distribution with a single exponential function $Ae^{-t/\tau_{\rm sel}}$. The fit result $\tau_{\rm sel}$ should be the same as the event-generator input value $\tau_{\rm gen}$ if there is no bias in the event selection. Table 6.8 shows the difference between $\tau_{\rm sel}$ and $\tau_{\rm gen}$. There is no bias within 1σ. We also check the same value with the other modes and find no bias.

Table 6.6: Summary of fit parameters obtained from the lifetime fit in MC simulation for the lifetime measurements.

(a) D^0

	$D^0 \to K^-\pi^+$	$D^0 \to K^-\pi^+$ with D^*-tag
$\tau_{\rm gen}$	414.7	414.7
$\tau_{\rm SIG}$	417.6 ± 0.9	415.8 ± 2.0
S	0.97 ± 0.01	1.10 ± 0.01
$S_{\rm tail}$	1.68 ± 0.07	4.56 ± 0.22
$f_{\rm tail}$	0.14 ± 0.03	0.013 ± 0.002
$\tau_{\rm BG}$	432 ± 20	605 ± 84
$f_{\tau_{\rm BG}}$	0.11 ± 0.01	0.67 ± 0.08
$S_{\rm BG}$	1.15 ± 0.01	-
$S_{\rm tail}^{\rm BG}$	4.97 ± 0.17	-
$f_{\rm tail}^{\rm BG}$	0.044 ± 0.003	-
$\tau_{\rm SIG} - \tau_{\rm gen}$	$+2.9 \pm 0.9$	$+1.1 \pm 2.0$

(b) D^+

	$D^+ \to K^-\pi^+\pi^+$ with D^*-tag	$D^+ \to \phi\pi^+$
$\tau_{\rm gen}$	1056.7	1056.7
$\tau_{\rm SIG}$	1056.5 ± 4.5	1050.4 ± 11.6
S	0.96 ± 0.08	0.89 ± 0.33
$S_{\rm tail}$	1.54 ± 0.64	1.59 ± 1.24
$f_{\rm tail}$	0.12 ± 0.22	0.25 ± 0.42
$\tau_{\rm BG}$	410 ± 11	594 ± 41
$f_{\tau_{\rm BG}}$	0.68 ± 0.02	0.14 ± 0.01
$S_{\rm BG}$	1.12 ± 0.02	1.11 ± 0.03
$S_{\rm tail}^{\rm BG}$	3.39 ± 0.37	2.51 ± 0.44
$f_{\rm tail}^{\rm BG}$	0.04 ± 0.01	0.05 ± 0.03
$\tau_{\rm SIG} - \tau_{\rm gen}$	-0.2 ± 4.5	-6.3 ± 11.6

(c) D_s^+

	$D_s^+ \to \phi\pi^+$	$D_s^+ \to \overline{K}^{*0}K^+$
$\tau_{\rm gen}$	466.7	466.7
$\tau_{\rm SIG}$	473.3 ± 2.7	472.3 ± 2.8
S	1.08 ± 0.03	0.60 ± 0.05
$S_{\rm tail}$	2.46 ± 2.65	1.17 ± 0.02
$f_{\rm tail}$	0.01 ± 0.04	0.78 ± 0.04
$\tau_{\rm BG}$	227 ± 29	443 ± 19
$f_{\tau_{\rm BG}}$	0.21 ± 0.02	0.08 ± 0.01
$S_{\rm BG}$	1.87 ± 0.27	1.11 ± 0.01
$S_{\rm tail}^{\rm BG}$	4.70 ± 0.65	2.73 ± 0.09
$f_{\rm tail}^{\rm BG}$	0.12 ± 0.08	0.05 ± 0.01
$\tau_{\rm SIG} - \tau_{\rm gen}$	$+6.6 \pm 2.7$	$+5.6 \pm 2.8$

Table 6.7: Summary of D^0 fit parameters obtained from the lifetime fit in MC simulation for the y_{CP} measurement.

	$D^0 \to K^-\pi^+$	$D^0 \to K^-K^+$
$\tau_{\rm gen}$	412.6	412.6
$\tau_{\rm SIG}$	413.0 ± 0.7	411.5 ± 3.0
S	1.00 ± 0.01	0.99 ± 0.10
$S_{\rm tail}$	1.92 ± 0.07	1.71 ± 0.78
$f_{\rm tail}$	0.08 ± 0.01	0.12 ± 0.21
$\tau_{\rm BG}$	492 ± 13	582 ± 17
$f_{\tau_{\rm BG}}$	0.093 ± 0.003	0.15 ± 0.01
$S_{\rm BG}$	1.17 ± 0.01	1.14 ± 0.01
$S_{\rm tail}^{\rm BG}$	4.81 ± 0.12	3.15 ± 0.16
$f_{\rm tail}^{\rm BG}$	0.039 ± 0.002	0.055 ± 0.009
$\tau_{\rm SIG} - \tau_{\rm gen}$	$+0.4 \pm 0.7$	-1.1 ± 3.0

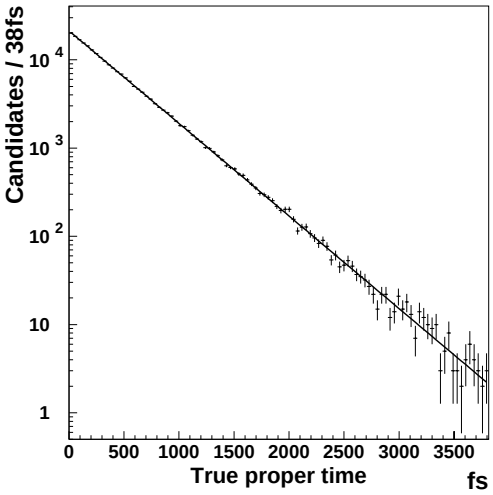


Figure 6.13: Generated proper-time distribution of the selected events in $D^0 \to K^-\pi^+$.

Table 6.8: Bias of event selection in the MC samples.

Mode	$\tau_{\rm sel} - \tau_{\rm gen}$
$D^0 \to K^-\pi^+$ w/o D^*-tag	-0.5 ± 0.6 fs
$D^0 \to K^-K^+$	$+0.4 \pm 1.4$ fs

Reconstruction of Proper-time

We investigate a possible bias in reconstruction of proper-time. We compare the reconstructed proper-time $t_{\rm rec}$ with the generated proper-time $t_{\rm gen}$. Figure 6.14 shows a distribution of the difference between the reconstructed and generated proper-times of the $D^0 \to K^-\pi^+$ events. We fit the distribution with two Gaussians. We take the mean of the distribution as the reconstruction bias $\Delta t_{\rm rec}$. As shown in Table 6.9 we see a small bias, that is, a negative shift in both decay modes. We also check the bias in other modes and see the same negative tendency in almost all decay modes.

We investigate the source of this negative bias. When we turn off the physics effects in the MC simulation: multiple scattering, decay in flight, and hadronic interaction [49], this negative bias becomes small as shown in Table 6.10. We conclude that this bias comes from these physics effects. We correct y_{CP} measurement with the values shown in Table 6.9.

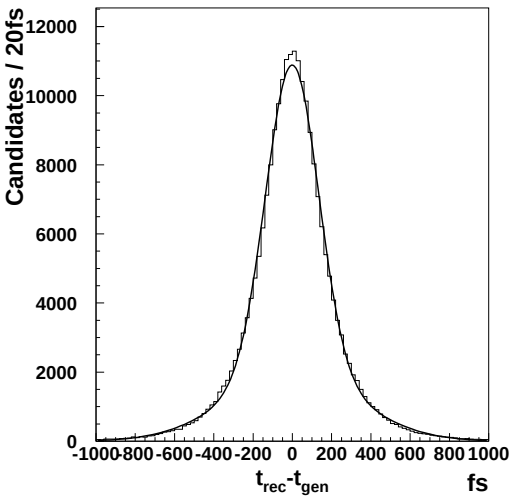


Figure 6.14: Distribution of the difference between the reconstructed and generated proper-times in $D^0 \to K^-\pi^+$.

Table 6.9: Bias of proper-time reconstruction in signal MC samples for the y_{CP} measurements.

Mode	$\Delta t_{\rm rec}$
$D^0 \to K^-\pi^+$ w/o D^*-tag	-1.5 ± 0.3 fs
$D^0 \to K^-K^+$	-2.7 ± 0.6 fs

Table 6.10: Bias of proper-time reconstruction in the special MC environment, where one of the physics effects is turned off.

Off	$D^0 \to K^-\pi^+$ w/o D^*-tag	$D^0 \to K^-K^+$
Nothing	-1.8 ± 0.8 fs	-2.7 ± 0.6 fs
Multiple scattering	-0.4 ± 0.3 fs	-1.2 ± 0.3 fs
Decay in flight	-	-1.7 ± 0.6 fs
Hadronic interaction	-	-0.3 ± 0.6 fs

Fitting

We investigate a possible bias in the fitting procedure. The result of the fit to signal MC events is compared with the result for the MC samples of signal+background. Table 6.11 shows the results of the fit. Here $\langle t_{\rm rec} \rangle$ is an average of reconstructed proper-times of the signal events. The $\tau_{\rm wobg}$ and $\tau_{\rm wbg}$ are fit results for signal events and signal+background events, respectively. D^+ and D_s^+ modes are listed for a reference purpose only. Some of decay modes are also tested in signal only MC simulation to decrease the statistical uncertainties. We see that differences between $\langle t_{\rm rec} \rangle$ and $\tau_{\rm wobg}$ are negligible, which indicates that there is no fitting bias in the case of no background event. On the other hand we find that $\tau_{\rm wbg}$ is larger than $\tau_{\rm wobg}$ except for $D^+ \to K^-\pi^+\pi^+$. The positive bias is due to the existence of the background events. However, since these fitting biases for $D^0 \to K^-\pi^+$ and $D^0 \to K^-K^+$ have the same sign and a similar size ($\sim$+3 fs), they cancel in the measurement of y_{CP} which is a ratio of the lifetimes.

Table 6.11: Fitting results for signal events and signal+background events.

Mode	$\langle t_{\rm rec} \rangle$ (fs)	$\tau_{\rm wobg}$ (fs)	$\tau_{\rm wbg}$ (fs)	$\tau_{\rm wbg} - \tau_{\rm wobg}$
Signal+background				
$D^0 \to K^-\pi^+$ w/o D^*-tag	410.0 ± 0.7	410.3 ± 0.7	413.0 ± 0.7	$+2.7 \pm 0.7$
$D^0 \to K^-K^+$	406.7 ± 2.6	408.7 ± 2.6	411.5 ± 3.0	$+2.8 \pm 3.0$
Signal only (high stat.)				
$D^0 \to K^-\pi^+$ w/o D^*-tag	410.5 ± 0.6	410.9 ± 0.6	-	-
$D^0 \to K^-K^+$	410.1 ± 1.1	410.4 ± 1.1	-	-
Signal+background				
$D^+ \to K^-\pi^+\pi^+$	1056.0 ± 4.6	1056.9 ± 4.6	1056.5 ± 4.5	-0.4 ± 4.6
$D^+ \to \phi\pi^+$	1038.6 ± 11.0	1042.6 ± 11.0	1050.4 ± 11.6	$+7.8 \pm 11.6$
$D_s^+ \to \phi\pi^+$	470.6 ± 2.8	470.3 ± 2.8	473.3 ± 2.7	$+3.0 \pm 2.8$
$D_s^+ \to \overline{K}^{*0}K^+$	465.2 ± 2.7	465.7 ± 2.7	472.3 ± 2.8	$+6.6 \pm 2.8$

Check

We check whether the breakdown of the bias is reasonable using the signal+background MC samples where we can obtain each bias of three sources and "$\tau_{\rm SIG} - \tau_{\rm gen}$" from the same sample. Table 6.12[a] shows each bias, sum of the three biases and the difference between the fit result and the input value of the event-generator ($\tau_{\rm SIG} - \tau_{\rm gen}$). We see good agreements between "Total" and "$\tau_{\rm SIG} - \tau_{\rm gen}$", which indicates that the breakdown is reasonable and that other unknown bias sources are small, if any.

Table 6.12: Breakdown of the three sources of the bias in the signal+background MC samples. Since we use the different Belle software libraries in the lifetime and y_{CP} measurements, we show the MC results for each measurement.

Mode (in units of fs)	Event selection	Reconst-ruction	Fitting	Total	$\tau_{\rm SIG} - \tau_{\rm gen}$
Lifetime					
$D^0 \to K^-\pi^+$ w/o D^*-tag	$+1.4 \pm 0.8$	-2.1 ± 0.5	$+3.1 \pm 0.9$	$+2.4$	$+2.9 \pm 0.9$
$D^+ \to K^-\pi^+\pi^+$	$+3.4 \pm 5.5$	-1.9 ± 1.3	-0.4 ± 4.6	$+1.1$	-0.2 ± 4.5
$D^+ \to \phi\pi^+$	-9.5 ± 11.0	-8.5 ± 3.0	$+7.8 \pm 11.6$	-10.2	-6.3 ± 11.6
$D_s^+ \to \phi\pi^+$	$+2.3 \pm 2.5$	$+0.8 \pm 2.7$	$+3.0 \pm 2.8$	$+6.1$	$+6.6 \pm 2.7$
$D_s^+ \to \overline{K}^{*0}K^+$	-0.4 ± 2.5	-1.9 ± 1.5	$+6.6 \pm 2.8$	$+4.3$	$+5.6 \pm 2.8$
y_{CP}					
$D^0 \to K^-\pi^+$ w/o D^*-tag	-0.5 ± 0.6	-1.9 ± 0.4	$+2.7 \pm 0.7$	$+0.3$	$+0.4 \pm 0.7$
$D^0 \to K^-K^+$	-3.0 ± 2.4	-2.7 ± 1.4	$+2.8 \pm 1.4$	-2.9	-1.1 ± 3.0

Summary

We correct y_{CP} with a bias of proper-time reconstruction listed in Table 6.9 since the effects of the other two biases are negligible, that is, there is no event selection bias and the fitting biases cancel in y_{CP}.

6.3.3 Final Results

We refit D^+ and D_s^+ lifetimes and y_{CP} parameter including the corrections (Table 6.6, Table 6.9) in the probability density functions and obtain $\tau(D^+) = 1037.2^{+12.4}_{-12.2}$ fs, $\tau(D_s^+) = 485.7^{+7.9}_{-7.8}$ fs and $y_{CP} = -0.005 \pm 0.010$, where the uncertainties are statistical only.

For D^0 lifetime measurement, we use the $D^0 \to K^-\pi^+$ w/o D^*-tag decay mode. We correct the fitted lifetime 417.9 ± 1.7 fs (Table 6.2) by 2.9 fs (Table 6.6) and we obtain 415.0 ± 1.7(stat.) fs. We compare this D^0 result with that of $D^0 \to K^-\pi^+$ with D^*-tag

[a]Note that in Tables 6.8 and 6.9 we use the large signal only MC samples to decrease the statistical uncertainties.

which is 418.1 ± 3.7(stat.) fs after the correction of 1.1 fs[b]. They are consistent within one standard deviation.

We summarize the final results of the lifetimes and y_{CP} in Table 6.13.

Table 6.13: Results of the lifetime and y_{CP} measurements. The uncertainties are statistical only.

Measurement	Result	Data size
D^0 lifetime	415.0 ± 1.7 fs	11.1 fb^{-1}
D^+ lifetime	$1037.2^{+12.4}_{-12.2}$ fs	11.1 fb^{-1}
D_s^+ lifetime	$485.7^{+7.9}_{-7.8}$ fs	11.1 fb^{-1}
y_{CP}	-0.005 ± 0.010	23.4 fb^{-1}

[b]Using the data for the y_{CP} measurement we obtain 415.8 ± 1.1(stat.) fs with the correction 0.4 fs (Table 6.7).

Chapter 7

Systematic Uncertainties

We estimate the size of possible systematic effects from reconstruction of the charm mesons and procedure of the fit. In the following sections we describe each systematic error in detail.

7.1 IP Size and Position Dependence

We consider systematic error associated with the size and position of the interaction point (IP). Size and position of the IP is measured run-by-run as described in Appendix C. Figure 7.1 shows measured size and position of the IP profile as a function of run numbers in Experiment 7. Measurement error of the IP position is typically (2–5) μm in x, (2–3) μm in y and $\sim$90 μm in z directions. We vary the IP position by $\pm$5 μm in x, $\pm$3 μm in y and $\pm$90 μm in z and repeat the analysis. The IP size has uncertainty of (1–3) μm in x, (1–2) μm in y and (20–30) μm in z directions. We vary the IP size by $\pm$5 μm in x, $\pm$2 μm in y[a] and $\pm$30 μm in z and repeat the analysis. We observe very small effects on the lifetimes and y_{CP}. In particular, the effect of the changes in z direction is 10 times smaller than that of x and y since the size of z direction of IP profile is (3–4) mm.

7.2 Vertexing Cut Dependence

We evaluate systematic uncertainty due to quality of the vertex fit for the D decay vertex. When the fit quality criterion is loosened, the fraction of the poorly measured events increases and modeling of resolution becomes difficult. It is important to make sure that the criterion is tight enough. The $\chi^2/N.D.F.$ cut for vertex fit is varied from 3 to 1, 1.5, ..., 4.5 and 5 in order to estimate the systematic uncertainty. Figure 7.2 shows the results of lifetimes and y_{CP} as a function of the $\chi^2/N.D.F.$ cut value.

As described in Section 4.3.1 we use a loose cut for the D production point since we use a tight cut for the D decay vertex. When we vary the cut on $\chi^2/N.D.F.$ of the production point by $\pm$1, the number of events changes only by $\pm$ $\sim$0.05% while $\pm$ $\sim$8%

[a] When the size becomes less than 1 μm, the size is set to 1 μm.

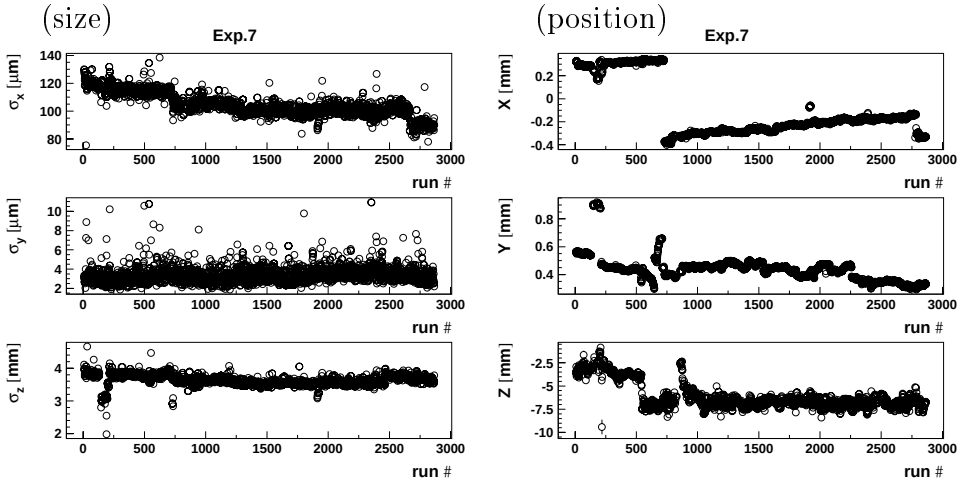


Figure 7.1: Size and position of IP profile versus run number in Experiment 7.

for the decay vertex. In fact, even if we vary the $\chi^2/N.D.F.$ cut on the production point by ± 2, our lifetimes change by < 0.01 fs in case of $D^0 \to K^-\pi^+$, < 0.1 fs in case of $D^0 \to K^-K^+$ which are much smaller than the effect of $\chi^2/N.D.F.$ cut on decay vertex. We ignore this systematic uncertainty.

In the estimation of the systematic uncertainty on the decay vertex, we first average fitted lifetimes for $\chi^2/N.D.F. < 3$ $(1, ..., 2.5)$ and those for $\chi^2/N.D.F. > 3$ $(3.5, ..., 5)$. We take the difference between the average and the value for $\chi^2/N.D.F. = 3$ as the systematic uncertainty.

7.3 World Average of Charm Meson Mass

We use the appropriate world average [6] of the charm meson mass in Eq. (4.1). We estimate the systematic errors by varying the mass by $\pm 1\sigma$ where σ is the error in the world average mass shown in Table 7.1.

We also consider the difference between the measured mass peak and the world average of the charm meson mass. It is due to the uncertainty in the momentum scale which comes primarily from the uncertainty of the magnetic field. We estimate the systematic errors by using the measured mass in Eq. (4.1).

Table 7.1: World average of the charm meson mass [6].

	PDG mass (MeV/c^2)
D^0	1864.5 ± 0.5
D^+	1869.3 ± 0.5
D_s^+	1968.6 ± 0.6

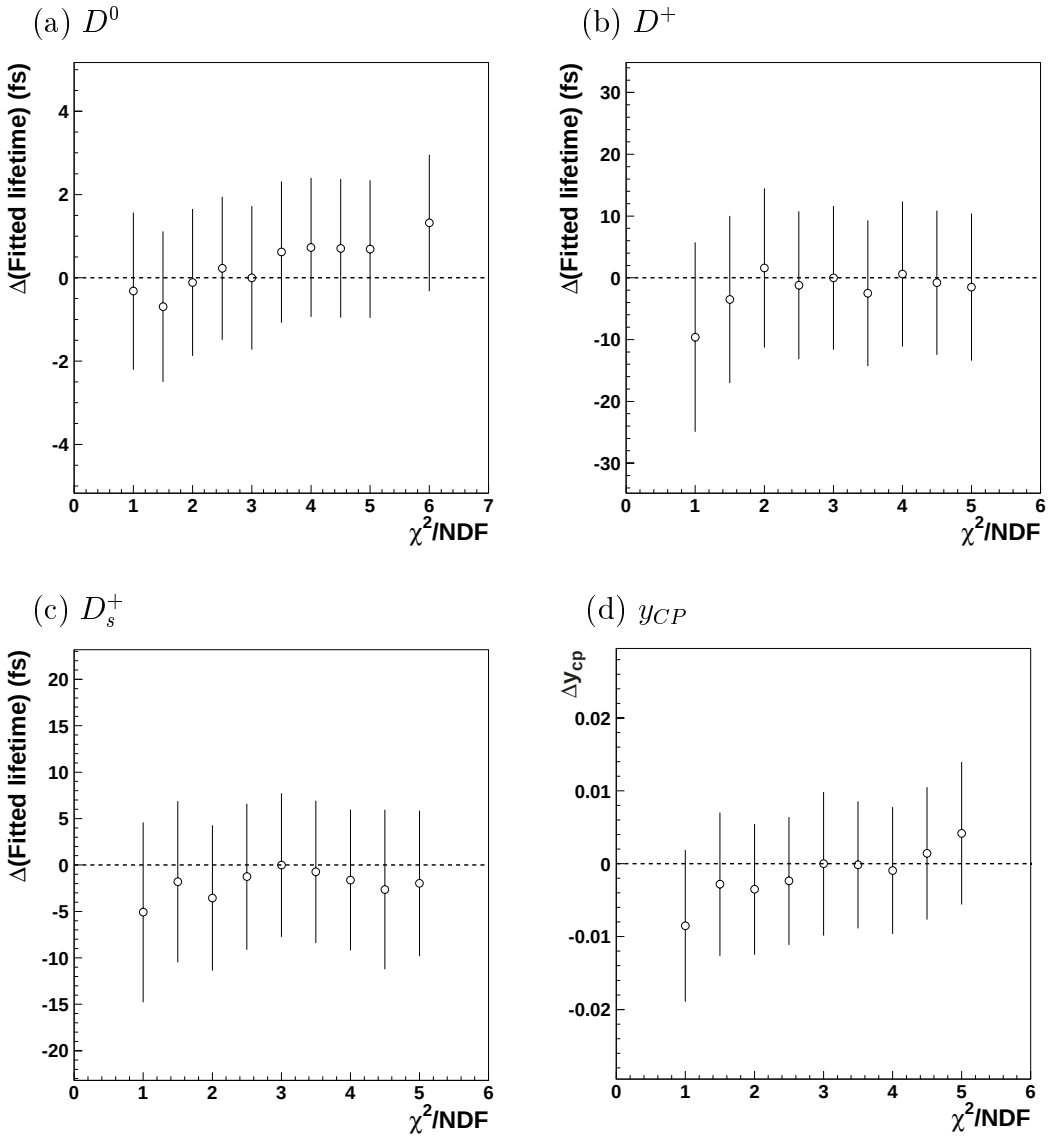


Figure 7.2: Fitted lifetimes and y_{CP} as a function of $\chi^2/N.D.F.$ cut on the decay vertex: (a) D^0, (b) D^+, (c) D_s^+ and (d) y_{CP}. The vertical axis is the difference between the fitted value and the value for $\chi^2/N.D.F. = 3$.

7.4 Decay Length

We evaluate the possibility of a bias in the reconstruction of the decay length by using a "zero-lifetime" sample, i.e., $\gamma\gamma \to \rho^0\rho^0 \to \pi^+\pi^-\pi^+\pi^-$ events. For the estimation of the bias in two-body decays of D mesons, two pairs of $\pi^+\pi^-$ are chosen whose mass is close to the world average value of the ρ mass [6]. For three-body decays of D mesons, a combination that gives maximum momentum is chosen for each event. Figure 7.3 shows the decay length distribution. A fit with the sum of two Gaussians yields a mean value as shown in Table 7.2 and 7.3.

The observed mean values are consistent with zero within two standard deviations. We expect two mechanisms that cause a mean value shift of the decay length. One is the effect of multiple scattering and the other is mis-alignment of the SVD.

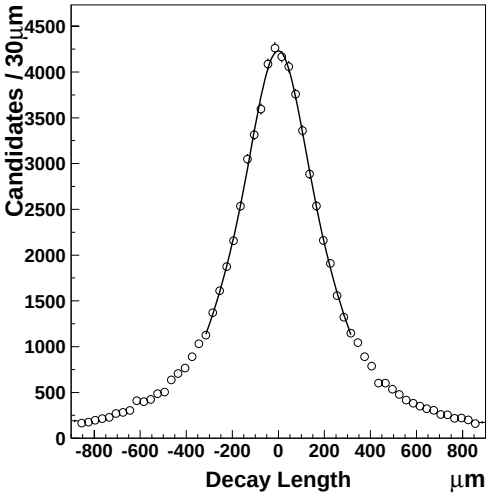
A decay length has a negative bias due to multiple scattering since the effect of the multiple scattering is asymmetric as shown in Figure 7.4, that is, $A > B$. We confirm the effect with MC simulation. For $D^0 \to K^-\pi^+$ we observe a mean shift of -0.48 ± 0.43 (μm) for the MC sample with the effect of the multiple scattering (MS) switched on and -0.09 ± 0.19 (μm) for the sample without MS. We also obtain -2.15 ± 0.52 (μm) with MS and -0.19 ± 0.25 (μm) without MS for a MC sample of $\gamma\gamma \to \pi^+\pi^-\pi^+\pi^-$.

If the measured mean shift is mainly due to the multiple scattering, we need to take into account momentum since the multiple scattering angle is inversely proportional to the momentum. The average momentum of K and π tracks from D^0 decays and pion tracks from the 2γ interaction is 2.0 GeV/c and 0.45 GeV/c, respectively. The measured mean shifts from $\gamma\gamma \to \pi^+\pi^-\pi^+\pi^-$ should be scaled by a factor of about 1/4. The maximum shift is estimated to be $\sim 2.10/4 = \sim 0.525$ μm for two-body decays in the Experiment 9.

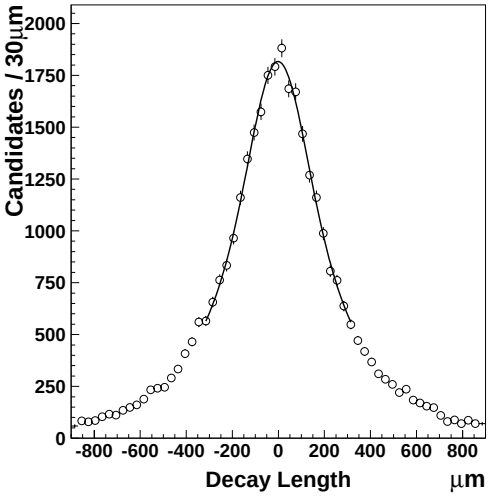
To study possible mis-alignment of the SVD, we check the resolution of impact parameter in the $r\phi$ plane, d_r, using the cosmic-ray particles passing near the interaction point. We use the cosmic-ray events with the transverse momentum of around 1.5 GeV/c which corresponds to the momenta of kaon and pion from the D^0 decay. A cosmic ray is reconstructed as two tracks by our tracking programs. In principle, two tracks must have the same track parameters, therefore the difference between the two track parameters divided by a factor of $\sqrt{2}$ represents the resolution. A impact parameter d_r represents the signed closest distance to the interaction point in the track parameterization [57]. We do not know the explicit relation between the impact parameter and the decay length. However since our lifetime measurement is based on the $r\phi$ plane and not z direction, it is important to check the size of the mis-alignment in the $r\phi$ plane. The mean shift and width are -0.45 ± 0.44 μm and 28.0 ± 1.8 μm for Experiment 7 and -2.2 ± 0.3 μm and 26.8 ± 2.2 μm for Experiment 9 as shown in Figures 7.5. These results indicate that the alignment of Experiment 9 could be worse than that of Experiment 7.

We estimate the systematic error associated with a possible bias in the decay length reconstruction by adding a fixed value to the reconstructed D decay length in the likelihood fit. The fixed value is the statistical error of the mean shift shown in Table 7.2 and 7.3. For the two-body decays in Experiment 9, however, we use the mean shift value

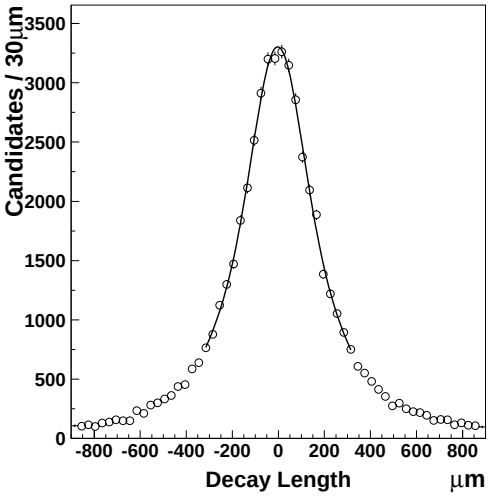
(a) 2 body-decay in Exp. 7



(b) 3 body-decay in Exp. 7



(c) 2 body-decay in Exp. 9



(d) 3 body-decay in Exp. 9

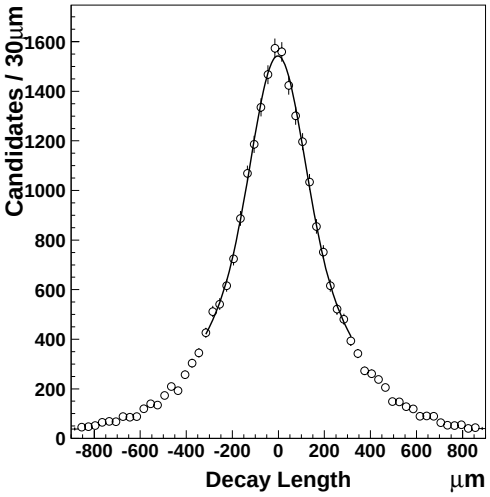


Figure 7.3: Decay length of the "zero-lifetime" sample $2\gamma \to \pi^+\pi^-\pi^+\pi^-$ in each experiment: (a) two body-decay in Experiment 7, (b) three body-decay in Experiment 7, (c) two body-decay in Experiment 9, and (d) three body-decay in Experiment 9.

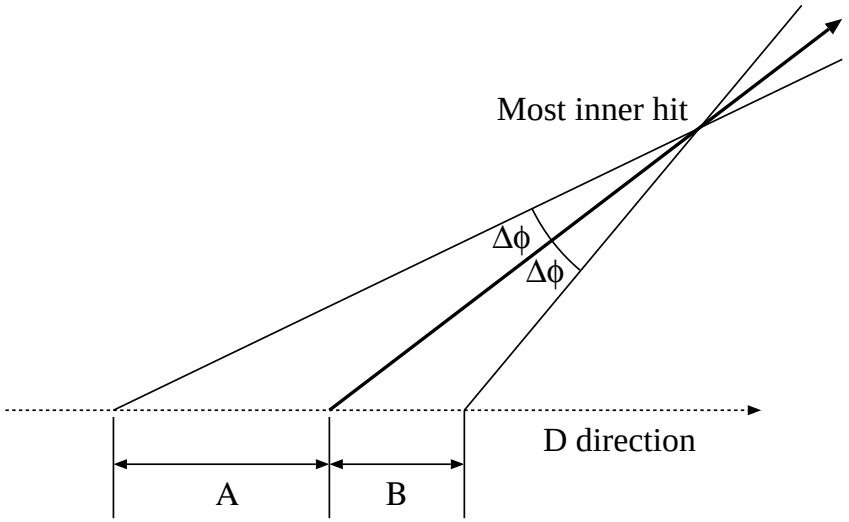


Figure 7.4: Asymmetric decay length due to multiple scattering. "A" is always longer than "B".

itself as the fixed value, since the observed mean shift could be due to the mis-alignment. For Experiment 5, we use the fixed value of Experiment 7 multiplied by a factor 5 which is based on the integrated luminosities. In y_{CP} measurement, this systematic uncertainty is largely canceled since y_{CP} is measured as a ratio of two lifetimes.

Table 7.2: Mean value of decay length of the "zero-lifetime" sample processed by the same library software as used for the lifetime measurement.

Modes	Experiment 7 (μm)	Experiment 9 (μm)
Two-body decay	-0.44 ± 0.99	-2.10 ± 1.06
Three-body decay	-0.29 ± 1.58	-1.26 ± 1.71

Table 7.3: Mean value of decay length of the "zero-lifetime" sample processed by the same library software as used for the y_{CP} measurement.

Modes	Exp. 7 (μm)	Exp. 9 (μm)	Exp. 11 (μm)	Exp. 13 (μm)
Two-body decay	-1.90 ± 1.25	-1.15 ± 1.10	-1.46 ± 0.72	-0.78 ± 1.28

With the "zero-lifetime" sample $2\gamma \to \pi^+\pi^-\pi^+\pi^-$, we also check ϕ dependence of the decay length, where the ϕ is an angle between charm meson momentum and x-axis in the $r\phi$ plane. Figure 7.6 shows the ϕ dependence of the mean of the decay length in each experiment. Even if we include this dependence in the decay length calculation,

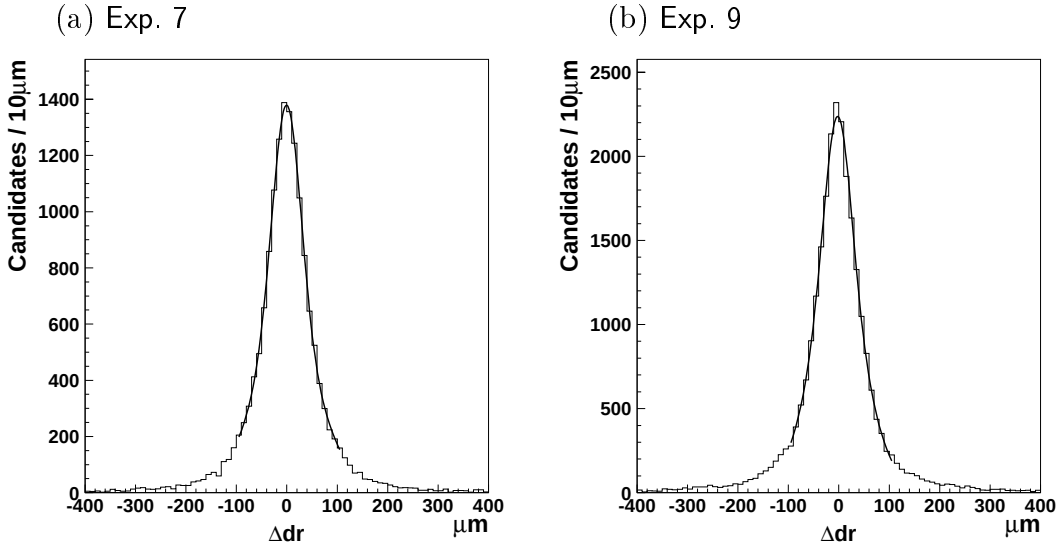


Figure 7.5: Difference between the impact parameters for the cosmic-ray events: (a) Experiment 7 and (b) Experiment 9.

our likelihood fit results are not affected. The systematic error on the ϕ dependence is negligible.

7.5 Detector Scale between the SVD and the CDC

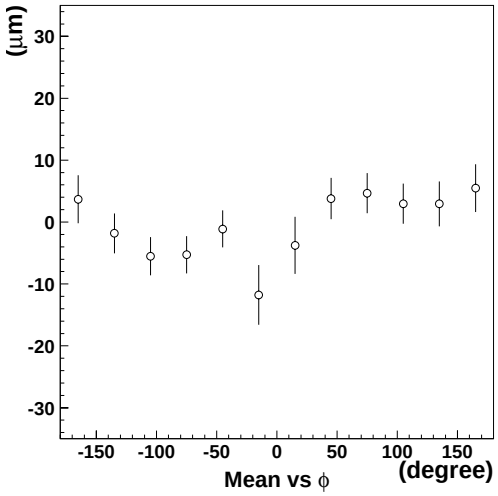
We find that the impact parameter d_r measured by the SVD and that by the CDC have a systematic difference as $d_r^{\rm CDC} \sim d_r^{\rm SVD} \times 0.99$ (a known problem of the CDC). This could be due to a difference of the scale between the SVD and the CDC. When we include this effect in the tracking program, our lifetime is shifted by $\sim$0.05%, which is very small.

7.6 D Mass Dependence of Proper-time

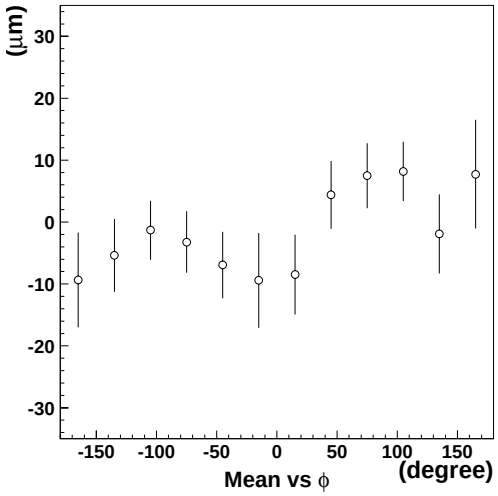
A large correlation between the proper-time and the reconstructed mass exists in all events. When the measured value of the opening angle between K and π in the $D^0 \to K^-\pi^+$ is smaller (larger) than the true value, the measured values of the D^0 mass and the decay length become smaller (larger). Figure 7.7 shows the correlations between Δt and the reconstructed D mass in MC samples, where Δt is the difference between the reconstructed proper-time and the true proper-time.

To evaluate the systematic uncertainties, we repeat the likelihood fit by adding a correction to proper-times. For the correction, we fit the Δt with a linear function of the reconstructed mass as shown in Figure 7.7. Table 7.4 shows the slopes of the fitted lines. We set this correction to zero at the measured mass peak ($= m_{\rm peak}$), that is, $\Delta t_{\rm corr} = a \cdot (m - m_{\rm peak})$ where $\Delta t_{\rm corr}$ is the correction of the proper-time, m is the measured mass and a is the slope shown in Table 7.4. Since we know that the background

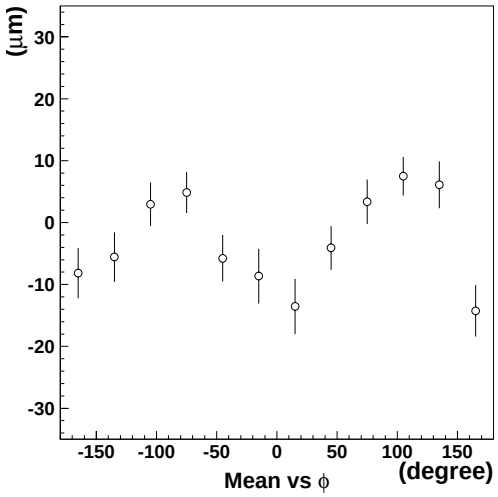
(a) 2 body-decay in Exp. 7



(b) 3 body-decay in Exp. 7



(c) 2 body-decay in Exp. 9



(d) 3 body-decay in Exp. 9

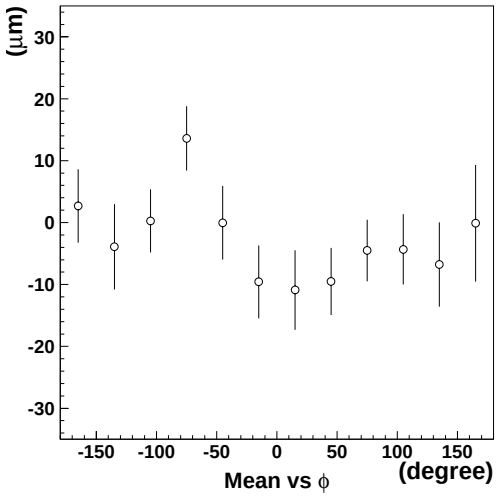


Figure 7.6: ϕ dependence of the decay length of the "zero-lifetime" sample $2\gamma \to \pi^+\pi^-\pi^+\pi^-$ in each experiment for the lifetime measurement: (a) two-body decay in Exp. 7, (b) three-body decay in Exp. 7, (c) two-body decay in Exp. 9 and (d) three-body decay in Exp. 9.

shape does not depend on the mass from Monte Carlo study as described in Section 5.2.2, the proper-time is corrected only in the signal term of the likelihood function. As shown in Tables 7.7 and 7.8, this effect is small since this correlation cancels when averaging over the mass regions above and below the measured mean D mass.

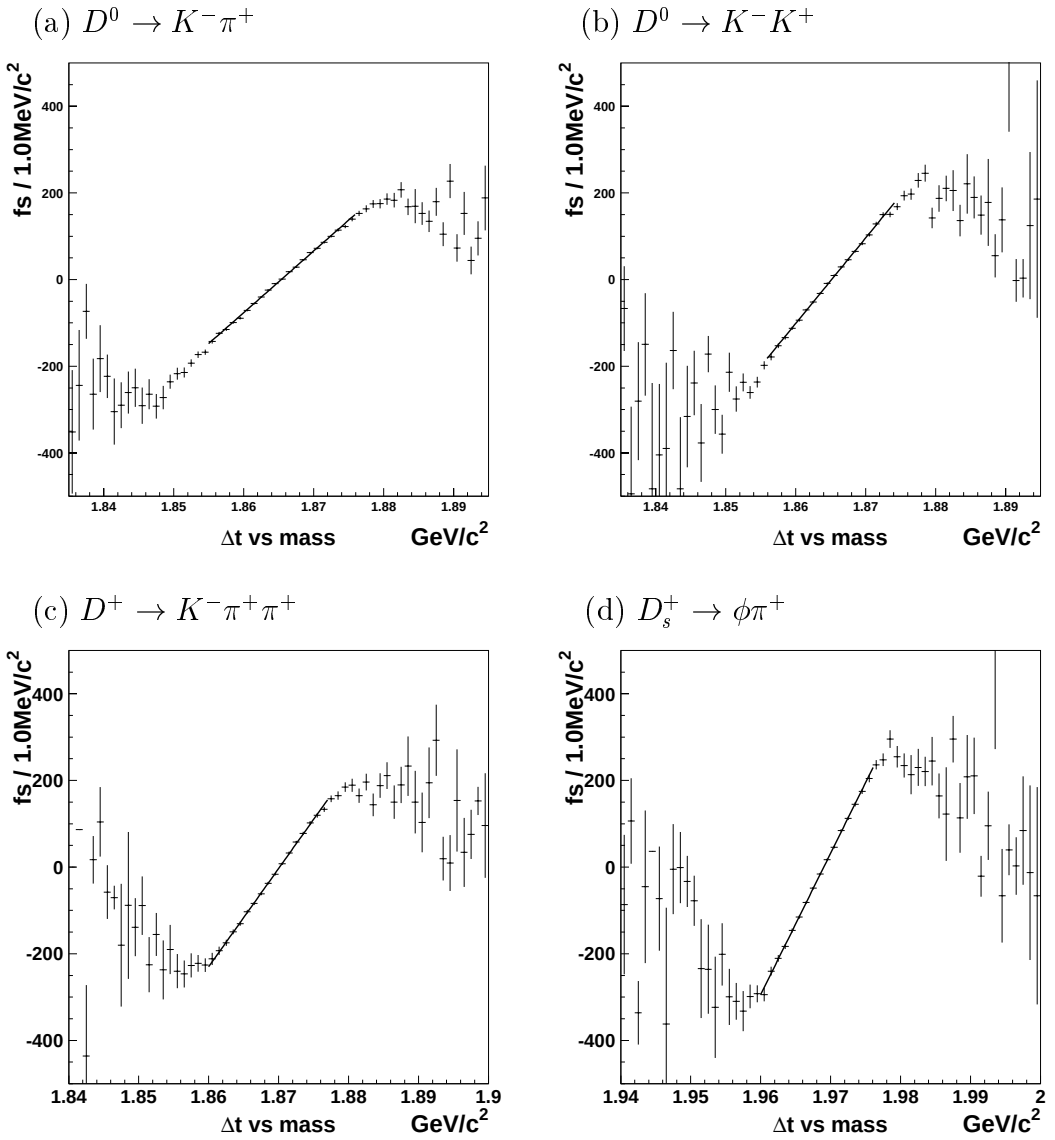


Figure 7.7: Difference between reconstructed proper-time and true proper-time Δt versus reconstructed mass: (a) $D^0 \to K^-\pi^+$, (b) $D^0 \to K^-K^+$, (c) $D^+ \to K^-\pi^+\pi^+$ and (d) $D_s^+ \to \phi\pi^+$.

We also check the dependence of the slopes and the mass peaks on the D momentum, opening-angle and polar angle as shown in Tables 7.5 for $D^0 \to K^-\pi^+$ and 7.6 for $D^0 \to K^-K^+$. An "offset" Δm is defined as $\Delta m \equiv m_{\rm peak} - 1864.5$ MeV/c^2. The $\langle t_{\rm rec}\rangle$ and $\langle t_{\rm gen}\rangle$ are the averages of the reconstructed proper-times and generated proper-times, respectively. The $\langle t_{\rm corr\_rec}\rangle$ is the average of the reconstructed proper-times corrected by $\Delta t_{\rm corr}$ which depends on the slope and the $m_{\rm peak}$. The opening-angle $\theta_{\rm open}$ is defined

Table 7.4: Slope in the Δt versus reconstructed mass plots.

Modes	Slope (fs/(GeV/c^2))
$D^0 \to K^-\pi^+$	14236 ± 4
$D^0 \to K^-K^+$	19521 ± 6
$D^+ \to K^-\pi^+\pi^+$	22195 ± 8
$D^+ \to \phi\pi^+$	36148 ± 10
$D_s^+ \to \phi\pi^+$	27690 ± 10
$D_s^+ \to \overline{K}^{*0}K^+$	31963 ± 8

as the angle between the daughter particles of D meson in the laboratory frame. The polar-angle $\theta_{\rm polar}$ is defined as the angle between the D meson momentum and the z-axis.

We observe no significant difference between $\langle t_{\rm rec}\rangle$ and $\langle t_{\rm gen}\rangle$ in all regions. We find that the difference between $\langle t_{\rm corr\_rec}\rangle$ and $\langle t_{\rm rec}\rangle$ does not depend on the D momentum, opening-angle and polar angle. Therefore, the systematic error estimation based on Table 7.4 is sufficient.

7.7 Signal Probability

The signal probability $f_{\rm SIG}$ has a statistical uncertainty which leads to a systematic uncertainty in the fit. We estimate these systematic uncertainties by varying the signal probability by $\pm 1\sigma_{f_{\rm SIG}}$[b], where the error $\sigma_{f_{\rm SIG}}$ is calculated as $\sigma^2_{f_{\rm SIG}} = \frac{\partial f_{\rm SIG}}{\partial \boldsymbol{\eta}}^T V_{\boldsymbol{\eta}} \frac{\partial f_{\rm SIG}}{\partial \boldsymbol{\eta}}$ where $\boldsymbol{\eta}$ are the fit parameters of the D-mass distribution and $V_{\boldsymbol{\eta}}$ is the covariance matrix of $\boldsymbol{\eta}$. This systematic uncertainty is small as shown in Tables 7.7 and 7.8.

7.8 Background Proper-time Distribution

The result of the likelihood fit may depend on the D mass region used for the fit, since the background parameters are determined effectively by the D mass sideband region. The mass range is varied from ± 40 MeV/c^2 to ± 35 MeV/c^2 and ± 45 MeV/c^2 . This systematic uncertainty is small as shown in Tables 7.7 and 7.8.

7.9 D Mass Peak Shift

We evaluate systematic uncertainties from the shift of the measured mass peak since the correlation between proper-time and mass is large as explained in Section 7.6. We vary it by $\pm 1\sigma_{m_{\rm peak}}$ in the calculation of the signal probability $f_{\rm SIG}$, where $\sigma_{m_{\rm peak}}$ is the statistical

[b] When the signal probability becomes larger (smaller) than 1 (0), we use 1 (0) as the signal probability.

Table 7.5: Dependences of the slope and the mass peak in $D^0 \to K^-\pi^+$.

(a) D momentum: p (GeV/c)

Region	Slope (fs/GeV/c^2)	Δm (MeV/c^2)	$\langle t_{\rm rec}\rangle$ (fs)	$\langle t_{\rm gen}\rangle$ (fs)	$\langle t_{\rm corr\_rec}\rangle$ (fs)
$p < 2.5$	20850	0.46	406.8	407.8	403.9
$2.5 < p < 3.0$	15340	0.56	416.9	416.0	415.0
$3.0 < p < 3.5$	13600	0.65	413.4	411.7	412.1
$3.5 < p < 4.0$	13350	0.85	410.5	410.8	408.9
$4.0 < p$	11930	1.30	414.9	413.0	412.5

(b) Opening-angle: $\theta_{\rm open}$

Region	Slope (fs/GeV/c^2)	Δm (MeV/c^2)	$\langle t_{\rm rec}\rangle$ (fs)	$\langle t_{\rm gen}\rangle$ (fs)	$\langle t_{\rm corr\_rec}\rangle$ (fs)
$\cos\theta_{\rm open} < 0.3$	20020	0.65	414.2	413.1	410.5
$0.3 < \cos\theta_{\rm open} < 0.5$	14260	0.68	415.1	412.5	413.1
$0.5 < \cos\theta_{\rm open} < 0.6$	12630	0.79	410.6	409.9	408.7
$0.6 < \cos\theta_{\rm open} < 0.7$	11860	0.96	411.2	411.5	410.0
$0.7 < \cos\theta_{\rm open}$	11240	1.27	412.7	413.4	410.8

(c) Polar-angle: $\theta_{\rm polar}$

Region	Slope (fs/GeV/c^2)	Δm (MeV/c^2)	$\langle t_{\rm rec}\rangle$ (fs)	$\langle t_{\rm gen}\rangle$ (fs)	$\langle t_{\rm corr\_rec}\rangle$ (fs)
$\cos\theta_{\rm polar} < -0.5$	20580	0.48	408.0	411.0	403.7
$-0.5 < \cos\theta_{\rm polar} < 0.0$	15900	0.54	409.7	409.7	407.6
$0.0 < \cos\theta_{\rm polar} < 0.2$	12980	0.60	411.4	409.1	410.3
$0.2 < \cos\theta_{\rm polar} < 0.4$	12360	0.67	414.7	413.5	413.5
$0.4 < \cos\theta_{\rm polar} < 0.6$	12830	0.78	412.5	412.2	411.9
$0.6 < \cos\theta_{\rm polar}$	13920	1.32	416.4	414.5	413.1

(d) All regions

Slope (fs/GeV/c^2)	Δm (MeV/c^2)	$\langle t_{\rm rec}\rangle$ (fs)	$\langle t_{\rm gen}\rangle$ (fs)	$\langle t_{\rm corr\_rec}\rangle$ (fs)
14236	0.84	413.0	412.2	410.7

Table 7.6: Dependences of the slope and the mass peak in $D^0 \to K^-K^+$.

(a) D momentum: p (GeV/c)

Region	Slope (fs/GeV/c^2)	Δm (MeV/c^2)	$\langle t_{\rm rec} \rangle$ (fs)	$\langle t_{\rm gen} \rangle$ (fs)	$\langle t_{\rm corr\_rec} \rangle$ (fs)
$p < 2.5$	28760	0.16	415.5	419.6	411.6
$2.5 < p < 3.0$	21120	0.31	409.2	408.5	407.7
$3.0 < p < 3.5$	18570	0.43	409.8	408.3	408.4
$3.5 < p < 4.0$	18460	0.71	415.4	416.0	414.4
$4.0 < p$	17030	1.00	409.0	407.4	405.1

(b) Opening-angle: $\theta_{\rm open}$

Region	Slope (fs/GeV/c^2)	Δm (MeV/c^2)	$\langle t_{\rm rec} \rangle$ (fs)	$\langle t_{\rm gen} \rangle$ (fs)	$\langle t_{\rm corr\_rec} \rangle$ (fs)
$\cos\theta_{\rm open} < 0.3$	32680	0.18	417.2	418.7	413.2
$0.3 < \cos\theta_{\rm open} < 0.5$	21660	0.29	416.3	413.7	414.3
$0.5 < \cos\theta_{\rm open} < 0.6$	18470	0.46	403.6	404.6	402.3
$0.6 < \cos\theta_{\rm open} < 0.7$	17440	0.70	412.4	412.8	411.2
$0.7 < \cos\theta_{\rm open}$	15860	0.94	409.1	408.7	405.9

(c) Polar-angle: $\theta_{\rm polar}$

Region	Slope (fs/GeV/c^2)	Δm (MeV/c^2)	$\langle t_{\rm rec} \rangle$ (fs)	$\langle t_{\rm gen} \rangle$ (fs)	$\langle t_{\rm corr\_rec} \rangle$ (fs)
$\cos\theta_{\rm polar} < -0.5$	28590	0.10	421.4	419.2	416.1
$-0.5 < \cos\theta_{\rm polar} < 0.0$	22360	0.26	408.7	411.9	406.8
$0.0 < \cos\theta_{\rm polar} < 0.2$	17040	0.45	412.2	410.4	410.1
$0.2 < \cos\theta_{\rm polar} < 0.4$	17300	0.46	411.4	410.9	410.5
$0.4 < \cos\theta_{\rm polar} < 0.6$	17440	0.58	407.5	407.4	407.2
$0.6 < \cos\theta_{\rm polar}$	19930	1.01	412.9	411.3	408.2

(d) All regions

Slope (fs/GeV/c^2)	Δm (MeV/c^2)	$\langle t_{\rm rec} \rangle$ (fs)	$\langle t_{\rm gen} \rangle$ (fs)	$\langle t_{\rm corr\_rec} \rangle$ (fs)
19521	0.58	411.3	411.1	408.8

error of the measured mean D mass. This systematic uncertainty is very small as shown in Tables 7.7 and 7.8.

7.10 Large Proper-times

In the likelihood fit we impose no cut on the measured proper-time. We observe a small number of D candidates with large proper-times which we call outliers. The effect of these candidates is studied by varying the proper-time range used in the fit as shown in Figures 7.8 and 7.9. Here the indices denote different ranges of the proper-time. The maximum index in each plot corresponds to the case where no limit on the range of the proper-time is imposed. The fraction of the removed events as outliers is less than 0.1%. Since we find small dependence on the choice of the proper-time range, we estimate systematic uncertainties from the slopes[c] obtained by a linear function fit for the lifetime measurements and from the maximum difference for the y_{CP} measurement.

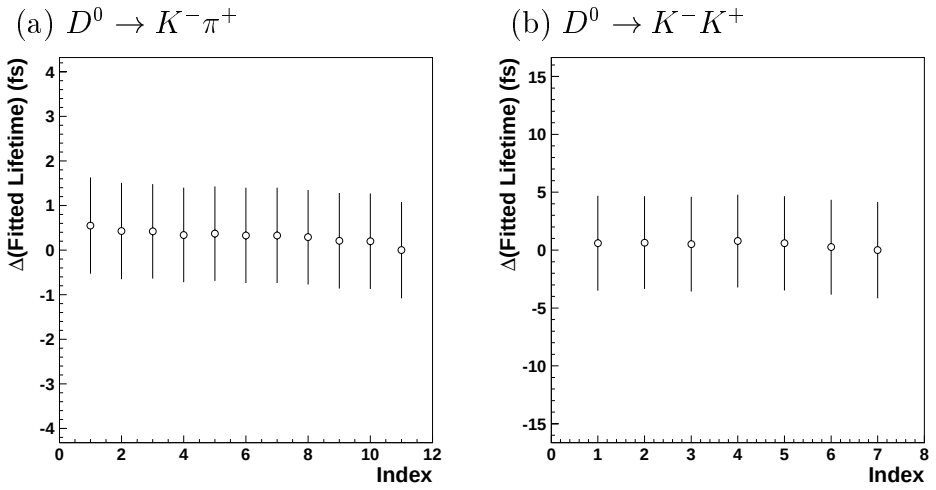


Figure 7.8: Effect of outlier events: difference between the fitted lifetime and lifetime obtained in the range of $-\infty < t < +\infty$ versus the proper-time range in the y_{CP} measurement. (a) $D^0 \to K^-\pi^+$, 1: $-3000 - +4000$, 2: $-3500 - +4500$, 3: $-4000 - +5000$, 4: $-4500 - +5500$, 5: $-5000 - +6000$, 6: $-5500 - +6500$, 7: $-6000 - +7000$, 8: $-6500 - +7500$, 9: $-7000 - +8000$, 10: $-7500 - +8500$, 11: $-\infty - +\infty$, (b) $D^0 \to K^-K^+$, 1: $-2000 - +4000$, 2: $-2500 - +4000$, 3: $-3000 - +4000$, 4: $-3500 - +4500$, 5: $-4000 - +5000$, 6: $-4500 - +5500$, 7: $-\infty - +\infty$ (in units of fs)

7.11 Reconstruction and Fit Biases – MC Statistics

As described in Section 6.3, we correct our results for possible biases in the reconstruction and fit procedures using Monte Carlo samples. We consider the statistical error of the

[c]The size of "obtained slope × (maximum index−1)" is considered as a systematic error.

(a) $D^0 \to K^-\pi^+$ (b) $D^+ \to K^-\pi^+\pi^+$ (c) $D_s^+ \to \phi\pi^+$

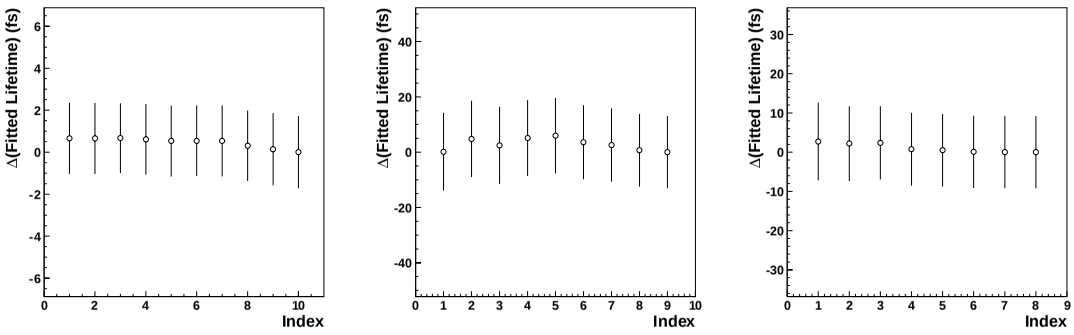


Figure 7.9: Effect of outlier events: difference between the fitted lifetime and lifetime obtained in the range of $-\infty < t < +\infty$ versus the proper-time range in the lifetime measurement. (a) $D^0 \to K^-\pi^+$, 1: $-3000 - +4000$, 2: $-3500 - +4500$, 3: $-4000 - +5000$, 4: $-4500 - +5500$, 5: $-5000 - +6000$, 6: $-5500 - +6500$, 7: $-6000 - +7000$, 8: $-6500 - +7500$, 9: $-7000 - +8000$, 10: $-\infty - +\infty$, (b) $D^+ \to K^-\pi^+\pi^+$, 1: $-1000 - +6000$, 2: $-1500 - +6500$, 3: $-2000 - +6000$, 4: $-2000 - +6500$, 5: $-2000 - +7000$, 6: $-2500 - +7500$, 7: $-3000 - +8000$, 8: $-5000 - +10000$, 9: $-\infty - +\infty$, (c) $D_s^+ \to \phi\pi^+$, 1: $-1000 - +3000$, 2: $-1500 - +3500$, 3: $-2000 - +4000$, 4: $-2500 - +4500$, 5: $-3000 - +5000$, 6: $-3500 - +5500$, 7: $-4000 - +6000$, 8: $-\infty - +\infty$ (in units of fs)

correction as the systematic uncertainties. For the y_{CP} measurement, we take into account 2σ of the statistical error[d].

7.12 PID cuts

As described in Section 6.3 the existence of background events affects our fitting results. We estimate the effect of signal-to-background ratio by varying PID cut.

Figures 7.10(a) and 7.10(b) show the results for $D^0 \to K^-\pi^+$ and $D^0 \to K^-K^+$ in the data sample of the y_{CP} measurement. We see a small dependence in both results. Using these results we obtain y_{CP} as shown in Figures 7.10(c) and 7.10(d) as a function of PID cut and signal-to-background ratio of K^-K^+, respectively.

For the y_{CP} and D^0 lifetime measurements, we use conservatively half of the maximum difference as a systematic error. For the D^+ and D_s^+ lifetime measurements, we fit the dependence with a linear function and take the change for 0.1 PID variation as a systematic error.

[d]The correction for y_{CP} is -0.003 and the systematic error of ± 0.003 is assigned.

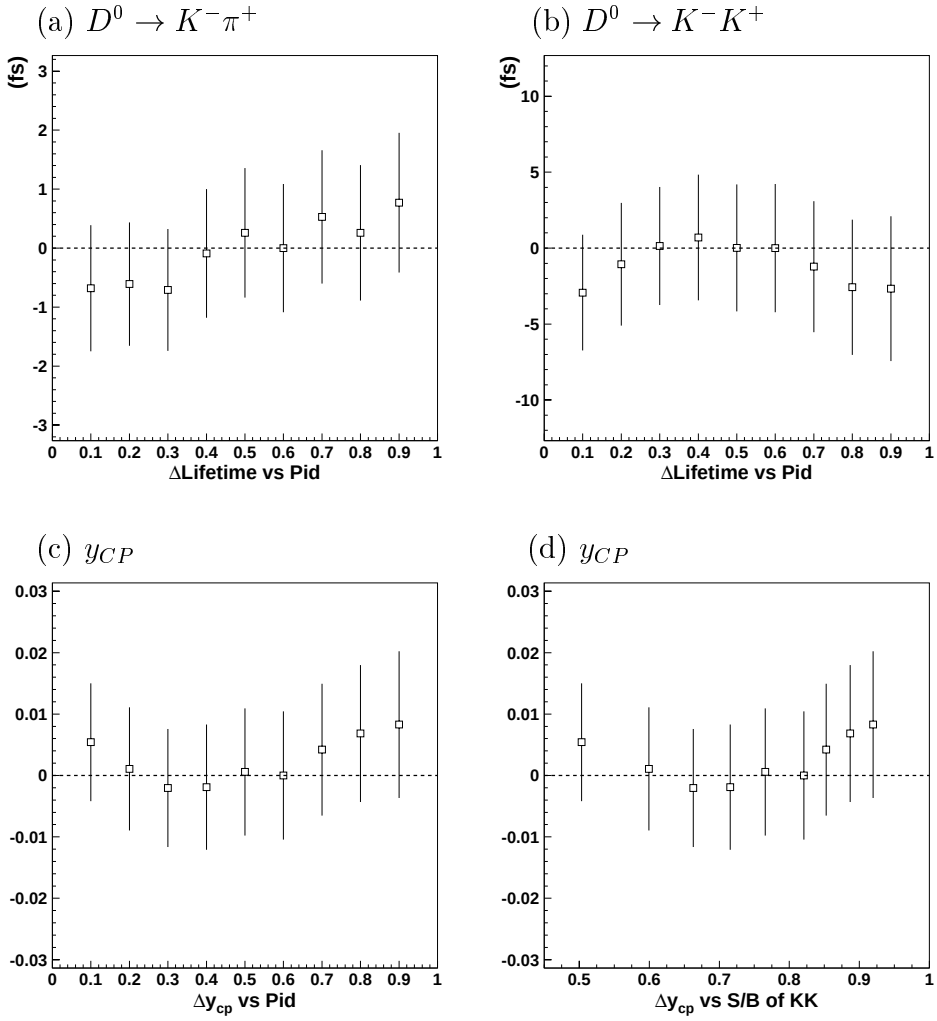


Figure 7.10: PID-cut dependence: (a) lifetime in $D^0 \to K^-\pi^+$ versus PID, (b) lifetime in $D^0 \to K^-K^+$ versus PID, (c) y_{CP} versus PID, and (d) y_{CP} versus signal-to-background ratio of $D^0 \to K^-K^+$.

7.13 Summary

We summarize various systematic uncertainties described in the previous sections. Tables 7.7 and 7.8 show the contributions to the total systematic error. All systematic uncertainties are combined in quadrature to obtain the total systematic error.

Systematic uncertainties of the lifetime ratios, $\tau(D^+)/\tau(D^0)$ and $\tau(D_s^+)/\tau(D^0)$, are calculated by the error propagation formula except for "IP", "Decay length" and "Detector scale" whose systematic uncertainties are correlated. The correlated systematic errors mean that when one value goes to positive (negative), the other value also goes to positive (negative). When a ratio r is defined as $r \equiv \frac{a}{b}$, we can calculate the correlated systematic error Δr as

$$\begin{aligned}
\Delta r &= \frac{a \pm \Delta a}{b \pm \Delta b} - r \\
&= r\left(\frac{1 \pm \Delta a/a}{1 \pm \Delta b/b} - 1\right) \\
&\simeq r\{(1 \pm \Delta a/a)(1 \mp \Delta b/b) - 1\} \quad (\because |\Delta b/b| \ll 1) \\
&\simeq \pm r(\Delta a/a - \Delta b/b). \quad (\because |\Delta a/a| \cdot |\Delta b/b| \ll 1)
\end{aligned}$$

Then we obtain ± 0.02 and ± 0.01 for the systematic uncertainties of the lifetime ratios $\tau(D^+)/\tau(D^0)$ and $\tau(D_s^+)/\tau(D^0)$, respectively.

Table 7.7: Systematic errors for the D^0, D^+ and D_s^+ lifetime measurements.

	Systematic error		
	$D^0 \to K^-\pi^+$	D^+	D_s^+
IP	± 0.2	$^{+0.4}_{-0.3}$	$^{+0.1}_{-0.8}$
Vertexing cuts	$^{+0.5}_{-0.4}$	$^{+0.0}_{-3.2}$	$^{+0.0}_{-2.9}$
Decay length	± 1.0	$^{+1.7}_{-1.6}$	± 1.7
Detector scale	± 0.2	± 0.5	± 0.2
PDG D mass	$^{+0.2}_{-0.1}$	$^{+0.5}_{-0.3}$	$^{+0.2}_{-0.1}$
Mass dependence of t	± 0.8	± 0.3	± 0.9
Signal probability	$^{+0.1}_{-0.2}$	$^{+1.8}_{-2.0}$	± 0.5
Background proper-time distribution	$^{+0.4}_{-0.3}$	$^{+0.0}_{-0.6}$	$^{+0.0}_{-0.2}$
Mass peak	negligible	$^{+0.0}_{-0.2}$	$^{+0.1}_{-0.0}$
Large proper-times	± 0.6	± 1.9	± 2.4
MC statistics	± 0.9	± 4.2	± 2.0
PID cuts	± 0.6	± 2.8	± 1.5
Total	± 1.9	$^{+6.0}_{-6.8}$	$^{+4.0}_{-5.1}$

Table 7.8: Systematic errors for the y_{CP} measurement.

	Systematic error
IP	$^{+0.000}_{-0.001}$
Vertexing cuts	$^{+0.001}_{-0.004}$
Decay length	negligible
Detector scale	negligible
PDG D mass	± 0.001
Mass dependence of t	± 0.003
Signal probability	± 0.001
Background proper-time distribution	$^{+0.002}_{-0.001}$
Mass peak	negligible
Large proper-times	± 0.002
Reconstruction bias	± 0.003
PID cuts	± 0.005
Total	$^{+0.007}_{-0.008}$

Chapter 8

Conclusions

We have measured charm meson lifetimes using 11.1 fb^{-1} of e^+e^- data collected with the Belle detector near the $\Upsilon(4S)$ resonance. Unbinned maximum likelihood fits to the proper-time distributions of fully reconstructed charm mesons yield the following results for the lifetimes and ratios;

$$\begin{aligned}
\tau(D^0) &= 415.0 \pm 1.7 \pm 1.9 \text{ fs}, \\
\tau(D^+) &= 1037.2^{+12.4+6.0}_{-12.2-6.8} \text{ fs}, \\
\tau(D_s^+) &= 485.7^{+7.9+4.0}_{-7.8-5.1} \text{ fs}, \\
\tau(D^+)/\tau(D^0) &= 2.50 \pm 0.03 \pm 0.02, \\
\tau(D_s^+)/\tau(D^0) &= 1.17 \pm 0.02 \pm 0.01,
\end{aligned}$$

where the first error is statistical and the second systematic. Table 8.1 and Figure 8.1 compare our results with the PDG averages [6] and recent measurements. Our lifetime measurements represent substantial improvements over the best previous measurements [21, 58, 59, 60, 61]. The ratio $\tau(D_s^+)/\tau(D^0)$ is more than 7.5 standard deviations larger than 1, based on the Belle data alone, and significantly larger than the range of estimates from the theoretical studies, 1.00–1.07 [7]. This result will act as a constraint on creation of new QCD models.

Experimentally the precision of the D^+ and D_s^+ lifetime measurements can be improved with more data. The precision of the D^0 lifetime measurement is limited by the systematic error, and therefore it is necessary to reduce systematic errors by further studies.

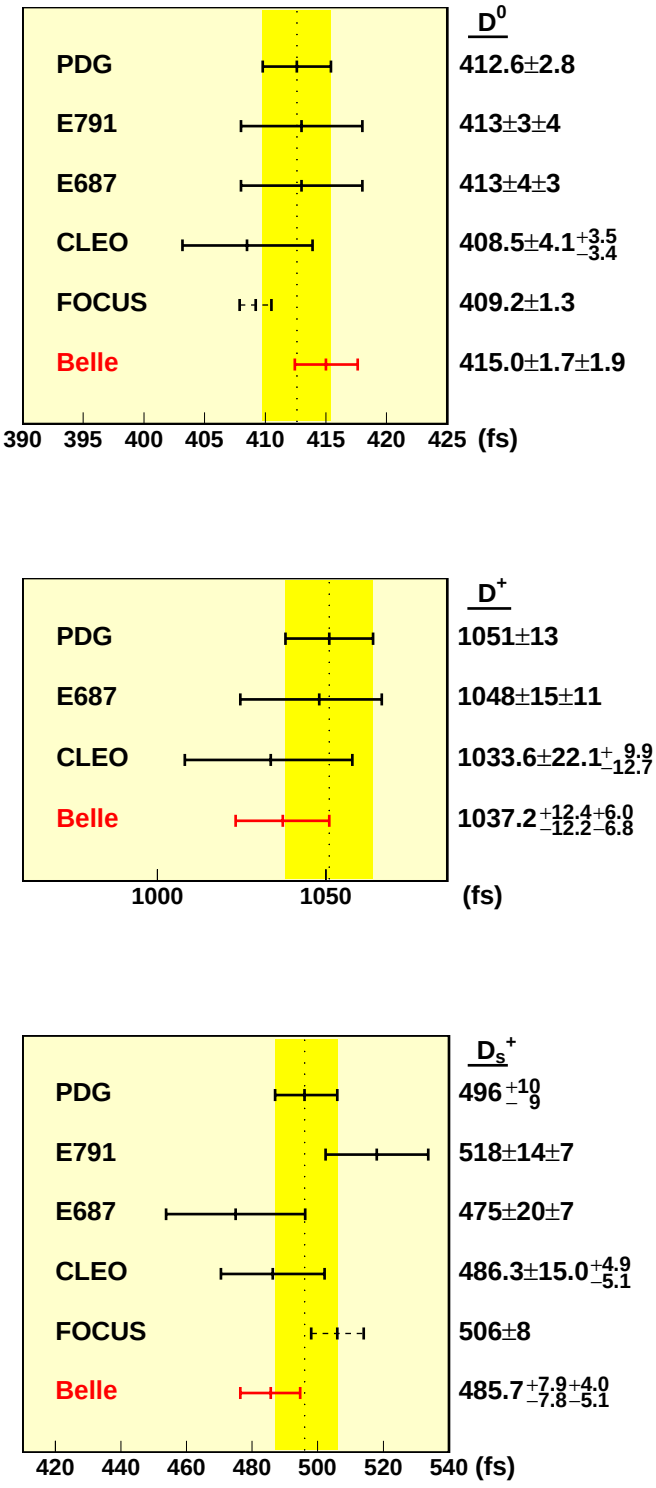


Figure 8.1: Comparison of our lifetime results with the PDG averages and recent measurements. No systematic errors are given for the FOCUS results.

Table 8.1: Comparison of our lifetime results with the PDG averages and recent measurements.

Experiment	$\tau(D^0)$ fs	$\tau(D^+)$ fs	$\tau(D_s^+)$ fs
PDG	412.6 ± 2.8	1051 ± 13	496^{+10}_{-9}
E791	$(413 \pm 3 \pm 4)^\dagger$	–	$(518 \pm 14 \pm 7)^\dagger$
E687	$(413 \pm 4 \pm 3)^\dagger$	$(1048 \pm 15 \pm 11)^\dagger$	$(475 \pm 20 \pm 7)^\dagger$
CLEO	$(408.5 \pm 4.1^{+3.5}_{-3.4})^\dagger$	$(1033.6 \pm 22.1^{+9.9}_{-12.7})^\dagger$	$(486.3 \pm 15.0^{+4.9}_{-5.1})^\dagger$
FOCUS	$409.2 \pm 1.3^\ddagger$	–	$506 \pm 8^\ddagger$
Belle	$415.0 \pm 1.7 \pm 1.9$	$1037.2^{+12.4+6.0}_{-12.2-6.8}$	$485.7^{+7.9+4.0}_{-7.8-5.1}$

†This result is included in the PDG average.

‡No systematic error is given.

Using a larger data sample corresponding to an integrated luminosity of 23.4 fb^{-1} we have measured the D^0-$\overline{D}^0$ mixing parameter y_{CP} using the $D^0 \to K^-\pi^+$ and $D^0 \to K^-K^+$ decays,

$$y_{CP} = -0.005 \pm 0.010^{+0.007}_{-0.008},$$

where the first error is statistical and the second systematic. In order to calculate a 95% confidence interval, we convolute the likelihood function with an asymmetric Gaussian whose standard deviations are given by the above systematic errors [62]. We thereby obtain a 95% confidence interval of

$$-0.030 < y_{CP} < 0.020.$$

Table 8.2 and Figures 8.3 compare our y_{CP} result with the PDG average and recent measurements. Our y_{CP} measurement has a better statistical precision than the FOCUS result [8] and, largely independent systematic errors. Our y_{CP} result is consistent with zero (i.e., the absence of D^0-$\overline{D}^0$ mixing) within 1σ. In addition, the average of the y_{CP} measurements shown in Table 8.2, is calculated to be 0.006 ± 0.008, which supports the prediction of the Standard Model, that is, a small rate of D^0-$\overline{D}^0$ mixing. The precision of the y_{CP} value can be also improved as we collect more data.

Table 8.2: Comparison of our y_{CP} result with the PDG average and recent measurements.

Experiment	y_{CP}
PDG	$0.0342 \pm 0.0139 \pm 0.0074$
E791	$0.008 \pm 0.029 \pm 0.010$
CLEO	$-0.011 \pm 0.025 \pm 0.014$
FOCUS	$(0.0342 \pm 0.0139 \pm 0.0074)^{\dagger}$
BABAR	$-0.010 \pm 0.022 \pm 0.017$
Belle	$-0.005 \pm 0.010^{+0.007}_{-0.008}$

†This result is included in the PDG average.

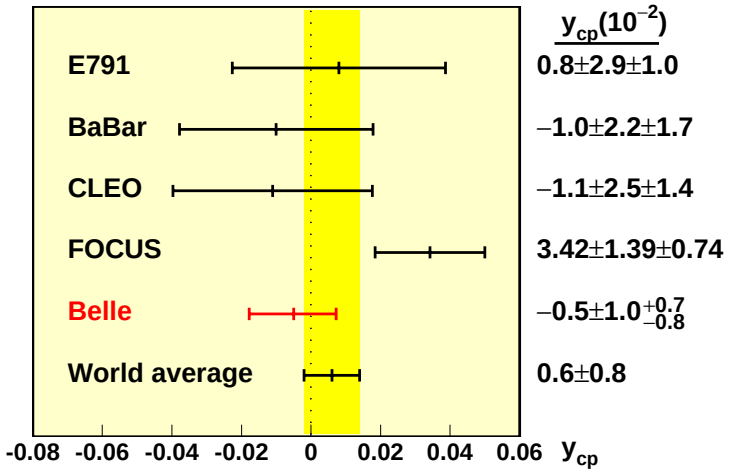


Figure 8.2: Comparison of our y_{CP} result with the recent measurements. The value of the world average is obtained from these 5 experimental results.

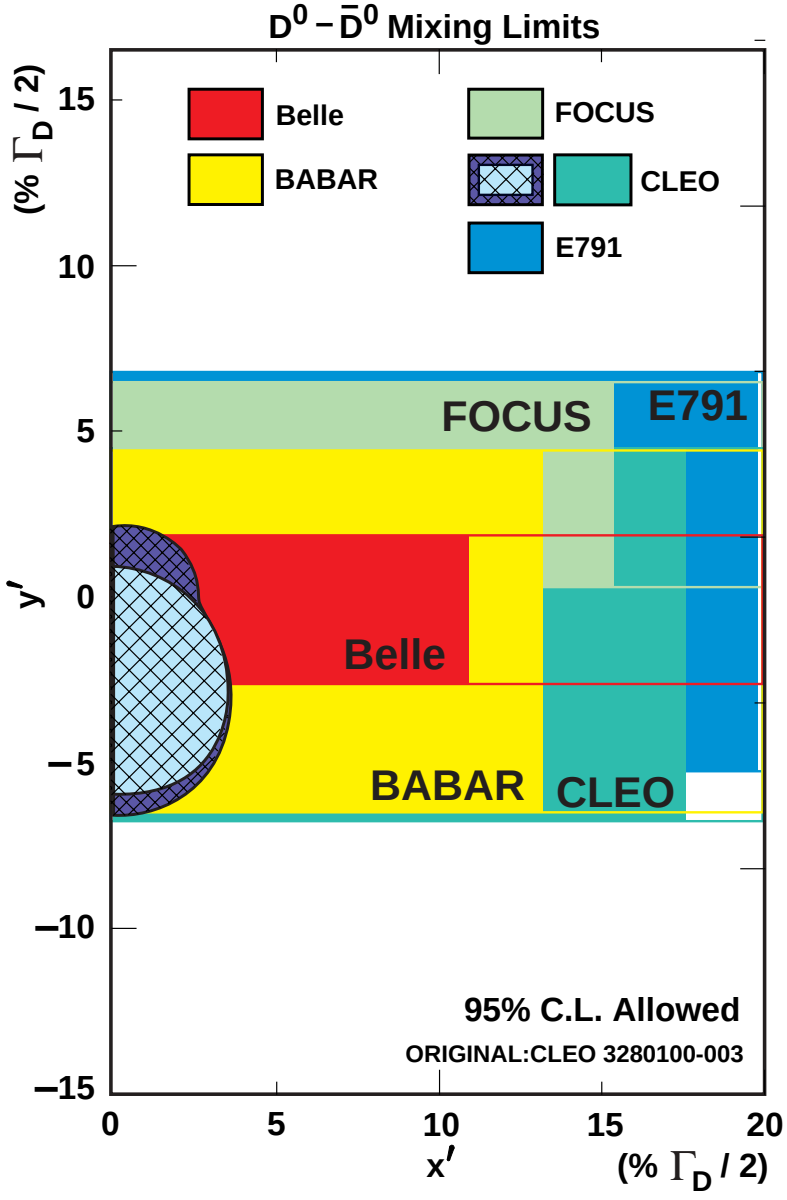


Figure 8.3: Comparison of our y_{CP} result with other recent D^0-$\overline{D}^0$ mixing measurements. Belle, BABAR, FOCUS, CLEO and E791 measurements of y_{CP} provide no information regarding x', and are plotted with the assumptions that $\delta = 0$ and that CP is conserved.

Appendix A

Charged Track Reconstruction

We describe the flow and algorithm of the charged track reconstruction in the following sections.

A.1 Overview

The tracking program reconstructs trajectories of charged particles in a magnetic field. The curvature of each trajectory gives information on the momentum of the charged particle. Figure A.1 shows the flow of the tracking in the Belle software library.

In the first step, most of charged tracks are found in the $r\phi$ plane using axial wires of the CDC, and then are reconstructed in the three dimensional space using both stereo and axial wires of the CDC. The remaining charged tracks are found in the $r\phi$ plane using axial plus stereo wires of the CDC or axial wires plus hit clusters on the SVD. Track parameters are then determined by a three dimensional fit assuming a uniform magnetic field, no dE/dx, no multiple scattering and so on.

In the next step, a refit is done using the Kalman Filter technique [40] taking into account a non-uniform magnetic field, dE/dx, multiple scattering and so on. SVD hit clusters are associated with the reconstructed charged tracks in the CDC using the same technique to improve the resolution of the track parameters.

In the final step, a trial is made to remove redundant tracks. For example, a low $p_{\rm t}$ pion may be reconstructed as two tracks (positive and negative charged tracks) in the case where it curls up in the detector. Comparing track parameters of the two tracks, the redundant track is removed.

Here we describe only the first part, that is, the finding of charged tracks. We have developed the Belle tracking program by considering the momentum range. Most of charged tracks are found by the "conformal finder" whose name comes from the algorithm of this finder. In this algorithm, the efficiency to find tracks degrades as track momenta ($p_{\rm t}$) become lower. We have, therefore, developed another finder, "curl finder" to find tracks with low $p_{\rm t}$. Since the curl finder is applied after the conformal finder, the axial wires which are associated with tracks by the conformal finder are ignored in the curl finder. We have also developed an additional tracking program, "PM (Pattern Matching) finder", to

find remaining tracks with low p_t. Both SVD hit clusters and CDC hit wires are used in this finder, while the conformal and curl finders only use CDC hit wires.

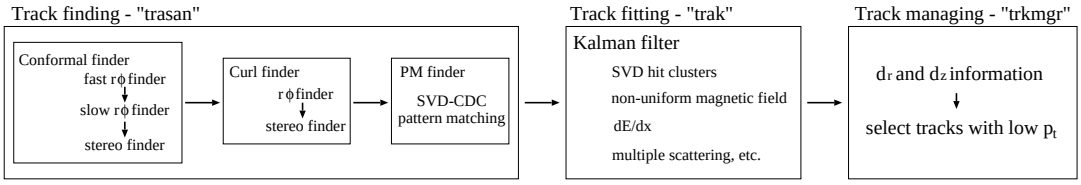


Figure A.1: Flow of the charged track reconstruction. Each program is called "trasan", "trak" and "trkmgr" in the Belle software library.

A.2 Fast Tracking: Conformal Finder

The conformal finder consists of two sub-finders: $r\phi$ and stereo finders. The former finds track candidates in the $r\phi$ plane and then the latter finds the corresponding stereo hit wires to reconstruct tracks in the three dimensional space. The $r\phi$ finder also consists of two sub-finders: fast and slow $r\phi$ finders. The conformal finder is required to be fast in speed in order to save CPU resources. The fast $r\phi$ finder, therefore, finds tracks with tighter criteria and fewer iterations than the slow one.

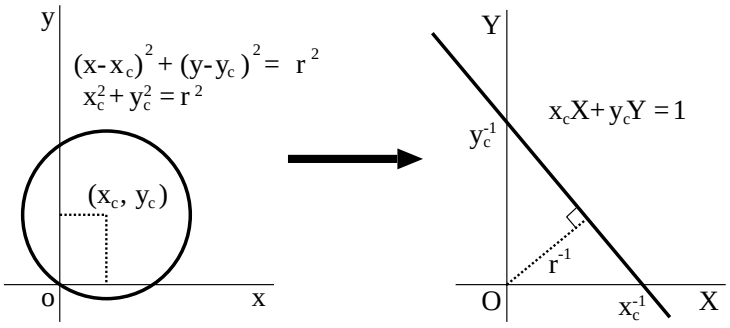


Figure A.2: Conformal transformation. A circle $(x - x_c)^2 + (y - y_c)^2 = r^2$ which passes through the origin is transformed into a line $x_c X + y_c Y = 1$.

The algorithm of the $r\phi$ finders is based on the conformal transformation. A hit position (x, y) in the xy plane ($r\phi$ plane) is transformed into a position (X, Y) in the conformal plane by the conformal transformation as,

$$X = \frac{2x}{x^2 + y^2}, \qquad Y = \frac{2y}{x^2 + y^2}.$$

By the conformal transformation, a circle which passes through the origin $(0, 0)$ is transformed into a line. The inverse of the distance between the line and the origin in the conformal plane corresponds to the radius of the circle as shown in Figure A.2.

(a) Axial hit wires (b) Axial and stereo hit wires

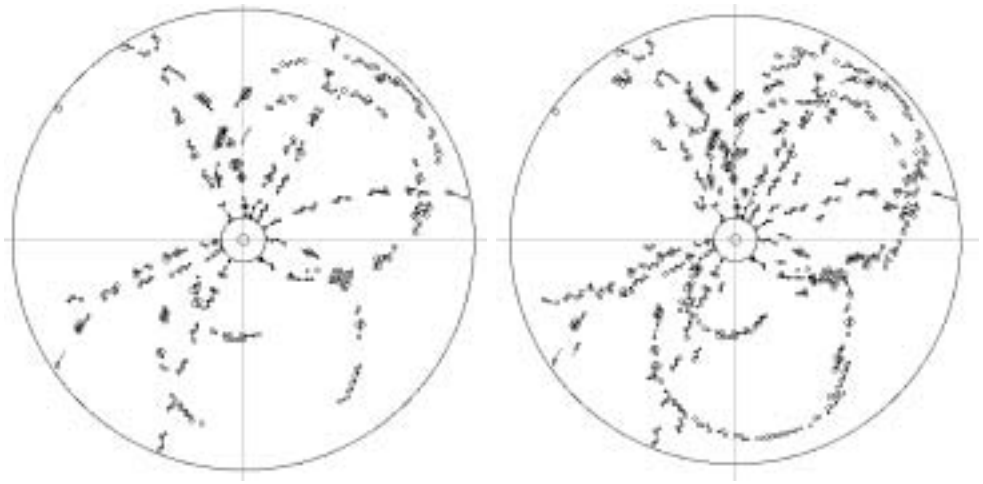


Figure A.3: Hit wires in the CDC: (a) axial hit wires, (b) axial and stereo hit wires. Both figures are obtained from the same event.

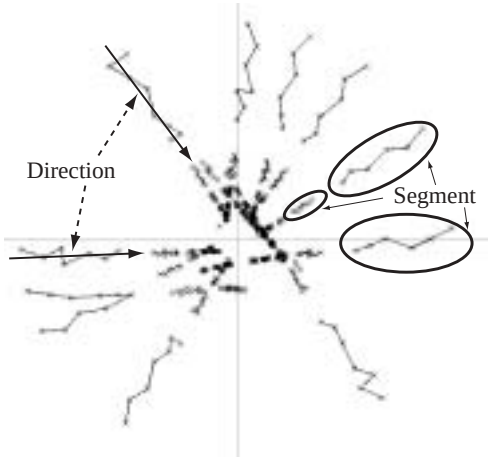


Figure A.4: Axial hit wires in the conformal plane. The cross points represent hit wires. The segment consists of the hit wires that are connected with lines in this figure and has a direction vector which points to the origin.

Figures A.3(a) and A.3(b) show CDC hit wires for a simulated $B^0 \to D^{*-} l^+ \nu_l$ event. After the conformal transformation, the position of the axial hit wires is transformed as shown in Figure A.4. Basically both $r\phi$ finders find tracks by searching hit wires with similar $\varphi = \tan^{-1} Y/X$ values in the conformal plane. The axial wires in the same superlayer are grouped according to the φ information as shown in Figures A.4 and A.5. These groups are called "segments." Next the segments of different superlayers are connected to form a track. We use a "direction" of the segment to connect the segments of different superlayers. As shown in Figure A.4, the segment points to the origin if it

is considered to be a line. The hit wires near the interaction point are transformed into positions far away from the origin in the conformal plane. We define a vector pointing to the origin for a segment. Using these vectors, we connect the segments of the different superlayers as shown in Figure A.5. Then we perform a circle fit for hit wires of the grouped segments in the $r\phi$ plane. In the fast $r\phi$ finder, we use "good" segments which easily reconstruct a circle in the $r\phi$ plane, that is, no unnecessary iteration exists to save CPU resources. In the slow $r\phi$ finder, we apply looser criteria to connect segments and repeat a circle fit to hit wires in the connected segments until we obtain a circle in the $r\phi$ plane.

Since the method in the stereo finder is almost the same as that of the curl finder, we describe it in the next section.

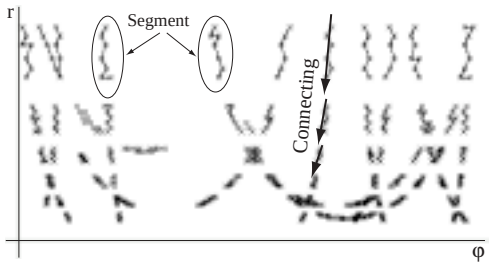


Figure A.5: Axial hit wires in the conformal plane.

A.3 Low p_t Tracking: Curl Finder

The curl finder consists of two sub-finders: $r\phi$ and stereo finders [63]. Their purpose is the same as that of the conformal finder.

As the track's p_t is lower the distance between the line and origin in the conformal plane is larger (Figure A.2). Consequently it is more difficult to find tracks using φ information in the conformal plane. The $r\phi$ finder, therefore, finds track candidates directly in the $r\phi$ plane by using "segments." A segment consists of consecutive hit wires in the same superlayer. Since a track with low p_t often makes consecutive hits in one layer of the superlayer, we start to reconstruct a track circle from the segment with the most consecutive hits assuming that the circle passes through the origin. Then we append axial hits in other superlayers one after another which are consistent with this circle.

The stereo finder reconstructs track candidates in the "sz" plane from stereo hits, where "s" is the signed arc length of the trajectory in the $r\phi$ plane. In the "sz" plane, a track trajectory becomes a line as shown in Figure A.6. First the stereo hit wires consistent with a track candidate in question are selected in the $r\phi$ plane using the distance between $r\phi$ position of the stereo wire at $z = 0$[a] and the circle. Next hit

[a]When a stereo wire is projected into the $r\phi$ plane, it becomes a line due to its slant.

positions (s, z) are calculated where the drift circle and track circle touch as shown in Figure A.7.

There are two solutions in (s, z); one is true and the other is fake. Some of the fake hits are rejected using geometrical information, but some fake hits remain. Furthermore there are sometimes noise hits in the selected stereo hits. In order to reject those "bad" hits, iterations of the line fit are performed and the best trajectory with the minimum $\chi^2/N.D.F.$ of the fit is found as shown in Figure A.8. Finally track parameters are re-determined by a three dimensional fit using those selected axial and stereo hits, assuming a helical trajectory although it is not exactly correct due to ionization loss (dE/dx), multiple scattering and so on.

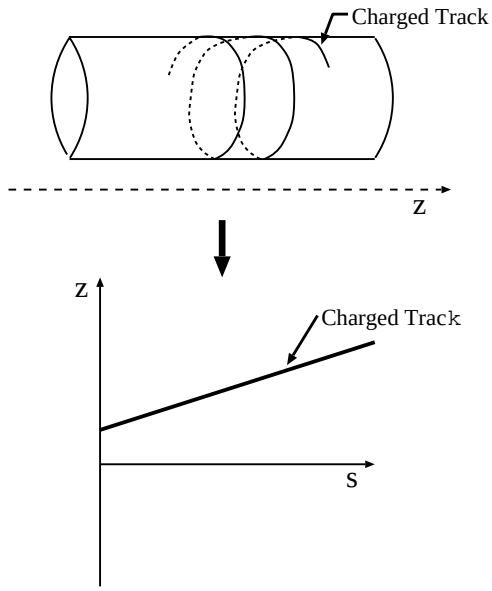


Figure A.6: Expansion of the cylinder onto a plane. The horizontal axis is called "s" and the vertical axis corresponds to the z-coordinate.

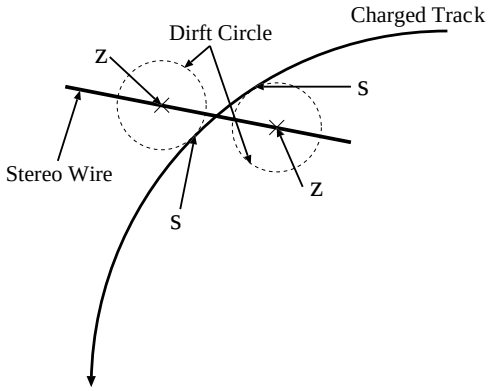


Figure A.7: Determination of "s" and z position of stereo wires.

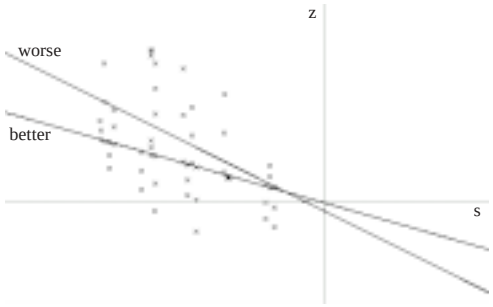


Figure A.8: Stereo hits on the sz plane. The two lines are obtained from these stereo hit wires. One ("better") with more number of associated hit wires is better than the other ("worse").

A.4 Low $p_{\rm t}$ Tracking: PM Finder

The additional low $p_{\rm t}$ tracking is based on the method of the SVD–CDC hit pattern matching. Here SVD hit information is also used in the track finding in order to achieve higher tracking efficiency.

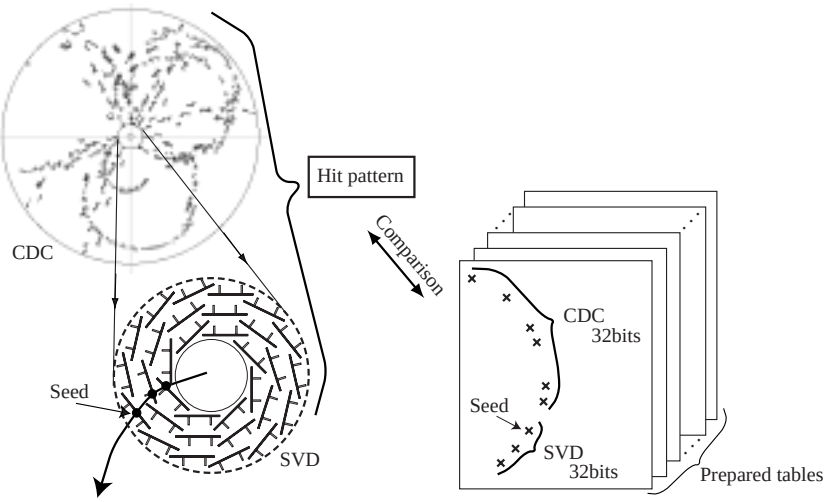


Figure A.9: Method of the hit pattern matching in the PM finder. We compare hit patterns calculated from the CDC hit wires and the SVD hit clusters with the prepared tables obtained from the Monte Carlo simulation.

We find tracks by comparing the SVD–CDC hit patterns with the prepared hit patterns which are generated from the Monte Carlo simulation as shown in Figure A.9. To find tracks with low p_t, we use 3 layers of the SVD and the inner 4 axial superlayers of the CDC. Since the first and second superlayers of the CDC are divided into 2 layers in the radial direction, 6 layers in total are used in the CDC. Each SVD and CDC layer is divided into 64 and 32 clusters in ϕ direction, respectively; in total 384 ($= 3 \times 64 + 6 \times 32$) clusters are defined to make hit patterns. A track is represented as two `unsigned` integers, where one (32 bits) is for the CDC and the other (32 bits) is for the SVD. Two `unsigned` integers are large enough to represent a track with a requirement that a track must pass through a special cluster (referred to as "seed" cluster). We use a cluster in the third layer of the SVD as a seed cluster. We find a track candidate by requiring cluster matching between data and the prepared table. For example it is required that the number of matching clusters in the CDC be larger than 0.8 times the number of clusters in the CDC, and also that the number of matching clusters in the SVD be 2 or more. Then $r\phi$-related track parameters are calculated from the axial hits, and z-related track parameters are form the SVD hits. Stereo hits are then selected using this track information. Finally track parameters are re-determined by a three dimensional fit using those selected axial and stereo hits and SVD hits.

Figure A.10 shows a result of the track finding. Some of the particles with low p_t are reconstructed as two separate tracks. The redundant track is removed by comparing the track parameters of the two tracks. The reconstructed tracks by the PM finder is required to have SVD hits, which are re-associated by the Kalman Filter technique in order to reject fake tracks.

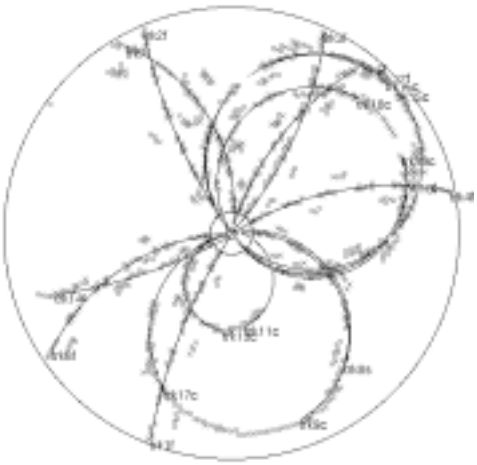


Figure A.10: Track candidates reconstructed by the track finding program. The character after "trk+*number*" represents which finder finds the track: "f" and "s" are by the fast and slow $r\phi$ finder of the conformal finder and "c" is by the curl or PM finder.

A.5 Performance of the Low $p_{\rm t}$ Tracking

Performance of the low $p_{\rm t}$ track finders is examined using 2×10^4 Monte Carlo events. Single pions are generated under the conditions that some effects such as multiple scattering and decay in flight are turned off (see Section 6.3) in this study. Figure A.11 shows the efficiency as a function of $p_{\rm t}$, where it is required that the difference between the reconstructed and generated pion momenta be less than 50 MeV/c. The solid and dashed histograms represent the efficiencies with and without the low $p_{\rm t}$ track finders (the curl and PM finders[b]), respectively. A substantial improvement can be seen in the range, $p_{\rm t} = 50 - 200$ MeV/c.

[b]The result with the curl finder is almost the same as that with the curl and PM finders. In the $B^0 \to D^{*-} l^+ \nu_l$ Monte Carlo events, the D^{*-} reconstruction efficiency is increased by 1% by the PM finder.

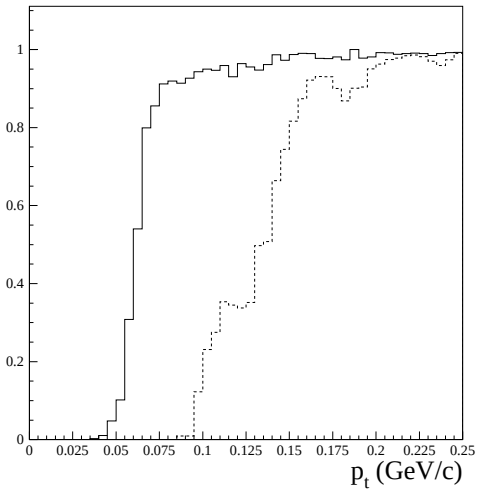


Figure A.11: Track finding efficiency as a function of p_{t} for single pion events. The solid and dashed histograms represent the efficiencies with and without the low p_{t} track finders, respectively.

Appendix B

Kaon Identification

The kaon identification [64] is carried out by combining information from three nearly-independent measurements.

- dE/dx measurement by the CDC ($p < 0.8$ or $p > 2.0$ GeV/c)
- TOF measurement ($p < 1.2$)
- Measurement of the number of photo-electrons in the ACC ($1.0 < p < 3.6$ in BACC and $0.7 < p < 2.4$ in EACC)

The mechanism for the particle identification is described in Section 3.

B.1 Probability

Particle identification is based on the likelihood of the detector response being due to an hypothesized signal particle species, compared to the likelihood for an assumed background particle species. This is expressed as a likelihood ratio

$$Prob(i:j) = \frac{P_i}{P_i + P_j}$$

where P_i is the PID likelihood calculated for the signal particle species and P_j for the background particle species. Index i and j can be any of five particle species, e (electron), μ (muon), π (pion), K (kaon) and p (proton). $Prob(i:j)$ is distributed on the interval [0,1] and we think of it as the "probability" that the measured particle belongs to the signal species i. When the particle ID system cannot distinguish between the hypotheses i and j, $Prob(i:j)$ is always set to 0.5.

Each PID likelihood P_i is a product of outputs from the three subdetectors,

$$P_i = P_i^{dE/dx} \times P_i^{\mathrm{TOF}} \times P_i^{\mathrm{ACC}}.$$

When information from one of the subdetectors is not available for the examined track, 0.5 is assigned for the subdetector PID likelihood for any particle species, $P_i^k = 0.5$ ($k = dE/dx$, TOF, ACC). This subdetector then has no effect in the combined likelihood ratio.

B.2 Calculation of PID likelihood

The PID likelihoods for the three subdetectors, $P_i^{dE/dx}$, P_i^{TOF}, P_i^{ACC}, are calculated as follows.

B.2.1 dE/dx

The χ^2 of dE/dx for each particle hypothesis is calculated as

$$\chi^2 = \left(\frac{(dE/dx)_{\mathrm{measured}} - (dE/dx)_{\mathrm{expected}}}{\sigma_{dE/dx}} \right)^2$$

where $(dE/dx)_{\mathrm{measured}}$ and $(dE/dx)_{\mathrm{expected}}$ are the measured and expected dE/dx values, respectively, and $\sigma_{dE/dx}$ is the expected resolution of the dE/dx measurement. Then the likelihood is calculated, assuming a Gaussian distribution:

$$P = \frac{e^{-\frac{1}{2}\chi^2}}{\sqrt{2\pi}\sigma_{dE/dx}}.$$

B.2.2 TOF

The χ^2 of the TOF is constructed by taking the difference between a two-vector containing the observed times in the two photomultiplier tubes (PMTs) of a counter and one containing the predicted times:

$$\Delta_i = t_i^{\mathrm{o}} - t_i^{\mathrm{p}}$$

where $i = 0, 1$ refers to the two ends of the counter. The 2×2 error matrix for Δ, which is called E, is evaluated by summing the contributions due to the uncorrelated uncertainty in the digitized times in the two phototubes and the correlated uncertainty due to the calculated time-of-flight from the tracking results. The χ^2 for one counter is then given by

$$\chi^2 = \Delta^T E^{-1} \Delta.$$

If a track passed through the edges of two counters, the counter χ^2's are summed to give the total χ^2 for the track ignoring the correlation between the times in the two counters. Then the likelihood is calculated by evaluating

$$P = \frac{e^{-\frac{1}{2}\chi^2}}{\prod_{i=1}^{N_{\mathrm{PMT}}} \sqrt{2\pi}\sigma_i}$$

where N_{PMT} is the number of PMT times that are included in the χ^2.

B.2.3 ACC

In contrast to dE/dx and the TOF counters, the ACC is basically an on-off device, where the observed signal (=number of photo-electrons, N_{pe}) is either zero, or distributed according to small-number statistics. For sub-threshold particles, the spectrum peaks at zero with a background tail due to δ-ray Čerenkov or light emission in the Goretex reflector. For above-threshold particles, the spectrum can be interpreted as s distribution of photo-electrons. Since it is very complex, it does not allow us to approximate the expected N_{pe} distribution with a simple function, by which the likelihood can be estimated from observed N_{pe} values.

The likelihood is given by comparing the observed N_{pe} with a threshold value, $N_{pe}^{\rm th}$, using the expected efficiency ε at the measured momentum:

$$P = \begin{cases} \varepsilon & (N_{pe} \geq N_{pe}^{\rm th}) \\ 1-\varepsilon & (N_{pe} < N_{pe}^{\rm th}). \end{cases}$$

The expected efficiency at the measured momentum for an assumed particle species, ε, is estimated by linear interpolation of efficient values, which are pre-stored in efficiency tables by evaluating at each momentum region. The threshold, $N_{pe}^{\rm th}$, is chosen so that the maximum separation between below-threshold and relativistic particles is obtained. The efficiency table and the threshold value are prepared based on Monte Carlo simulations for each type of counter module, that is, as a function of polar angular regions.

B.3 Kaon Efficiency and Pion Fake Rate

The validity of the K/π identification has been demonstrated using the charm decay, $D^{*+} \to D^0\pi_s^+$, followed by $D^0 \to K^-\pi^+$. The characteristic low momentum pion π_s^+ from the D^{*+} decay allows these decays to be selected with a good S/N ratio without relying on particle identification. Therefore, the detector performance can be directly probed with the daughter K and π mesons from the D decay, which can be tagged by their relative charge with respect to the low momentum pion.

Figure B.1 shows two dimensional plots of the likelihood ratio $Prob(K:\pi)$ and measured momenta for the kaon and pion tracks. The figure demonstrates the clear separation of kaons and pions up to $\sim$4 GeV/c.

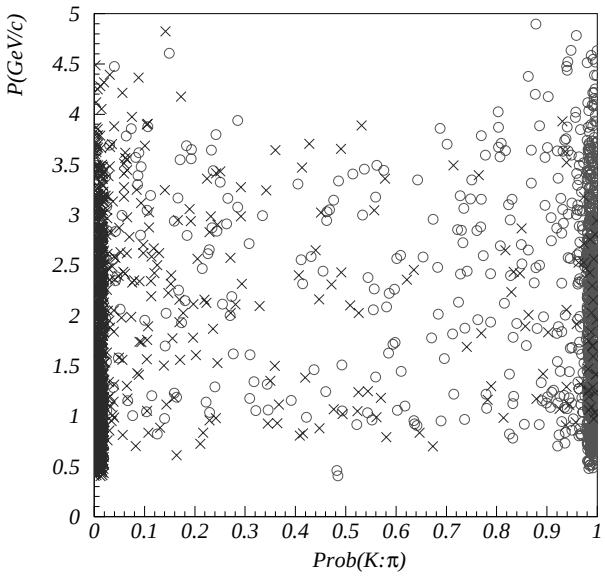


Figure B.1: Likelihood ratio $Prob(K:\pi)$ versus momenta for daughter tracks from $D^0 \to K^-\pi^+$ decays, tagged by the charged of the low momentum π_s^+'s. The open circles corresponds to K's and the cross points to π's.

The measured K efficiency (ε_{eff}) and π fake rate ($\varepsilon_{\text{fake}}$) are defined as

$$\varepsilon_{\text{eff(fake)}} = \frac{\text{number of } K(\pi) \text{ tracks (mis)identified } K}{\text{number of } K(\pi) \text{ tracks}}$$

where true $K(\pi)$ tracks are tagged by charge of the low momentum $\pi^{\pm}$. The measured K efficiency and π fake rate in the barrel region are plotted as function of the track momentum from 0.5 to 4.0 GeV/c in Figure B.2. The likelihood ratio cut, $Prob(K:\pi) \geq 0.6$, is applied in this figure. For most of the region, the measured K efficiency exceeds 80%, while the π fake rate is kept below 10%.

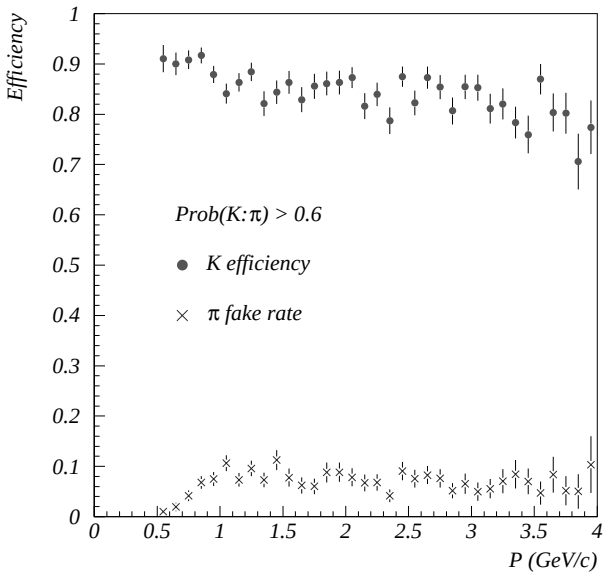


Figure B.2: K efficiency and π fake rate, measured with $D^{*+} \to D^0(K\pi)\pi_s^+$ decays, for the barrel region. The likelihood ratio cut $Prob(K:\pi) \geq 0.6$ is applied.

Appendix C

Interaction Point Profile

The interaction point (IP) profile is calculated run-by-run using the hadronic data described in Section 4.2.1 since the IP position changes during the collision of e^+e^-. Initially the primary event vertex is calculated using all charged tracks. Then the primary event vertex is recalculated using only the tracks coming from the initial primary vertex. Three-dimensional distribution of the primary vertex is fitted with 3-dimensional Gaussian to obtain the IP position and its size ($\boldsymbol{\sigma}_{\rm Hadron}$) which includes our detector resolution. The 3-dimensional fit takes into account the effect of the ± 11 mrad crossing angle of the beamlines. When a beam fill contains several runs, the IP position is recomputed for each run. In the case of very long runs, the IP position is also recalculated every $\sim$2000 event sample.

We know that our detector resolution for the y-direction is much larger than the IP size ($\boldsymbol{\sigma}_{\rm IP}$). For the IP size in the y-direction we use the beam size determined from the machine optics of the accelerator, which is shown in Figure C.1.

A $\sigma_{{\rm Hadron}X}$ and a $\sigma_{{\rm Hadron}Y}$ shown in Figure C.2 include our detector resolutions and the IP size. The IP size for the y-direction ($\sigma_{{\rm IP}y}$) is negligible in $\sigma_{{\rm Hadron}Y}$ because it is small ($\sim$ 3 μm) enough as described above. We assume that a $\sigma_{{\rm Hadron}Y}$ represents our detector resolution ($\sim$ 70 μm). Therefore we can obtain the IP size for the x-direction ($\sigma_{{\rm IP}x}$) by

$$\sigma^2_{{\rm IP}x} = \sigma^2_{{\rm Hadron}X} - \sigma^2_{{\rm Hadron}Y}.$$

A $\sigma_{{\rm Hadron}Z}$ ($\gtrsim$ 2 mm) also includes our detector resolution and the IP size. However a size of our detector resolution is negligible in the z-direction since the IP size is much larger than our detector resolution. A $\sigma_{{\rm Hadron}Z}$ is used as IP size for the z-direction ($\sigma_{{\rm IP}z}$).

The IP profile is calculated from the IP size and uncertainties of the IP position:

$$\boldsymbol{\sigma}^2_{\rm IPprofile} = \boldsymbol{\sigma}^2_{\rm IP} + \boldsymbol{\sigma}^2_{\rm IPposition}.$$

The results are shown in Figures 7.1.

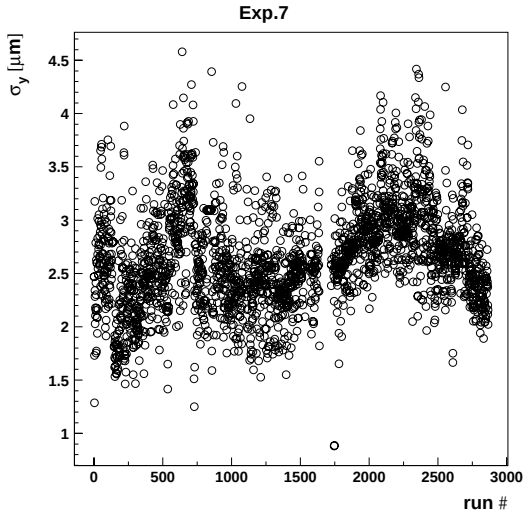


Figure C.1: IP size for the y-direction determined from the machine optics of the accelerator in Experiment 7.

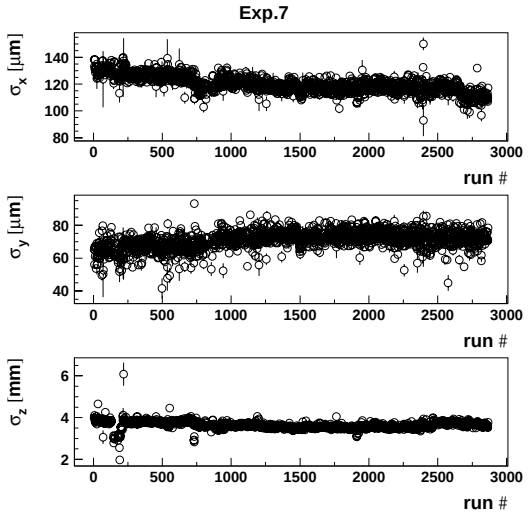


Figure C.2: IP size obtained from hadronic events in Experiment 7.

Appendix D

Kinematic Fitter

We describe the fitting technique to obtain the vertices which is based on the least square method using the Lagrange multiplier method [54].

D.1 Overview

The parameters before and after a fitting are represented by $\boldsymbol{\alpha}_0$ and $\boldsymbol{\alpha}$. The error matrix of $\boldsymbol{\alpha}_0$ is represented by $V_{\boldsymbol{\alpha}_0}$. In general, the χ^2 is written as

$$\chi^2 = (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0)^T V_{\boldsymbol{\alpha}_0}^{-1} (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0).$$

In the least square method we try to find the minimum χ^2 by varying the $\boldsymbol{\alpha}$. When we have no constraint for the parameter $\boldsymbol{\alpha}$, the χ^2 is minimized by $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$, that is, $\chi^2 = 0$.

Constraints can be written generally as $\boldsymbol{H}(\boldsymbol{\alpha}) = 0$. The constraints are incorporated using the method of the Lagrange multiplier in which the χ^2 is written as a sum of two terms,

$$\chi^2 = (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0)^T V_{\boldsymbol{\alpha}_0}^{-1} (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0) + 2\boldsymbol{\lambda}^T \boldsymbol{H}(\boldsymbol{\alpha}),$$

where $\boldsymbol{\lambda}$ is the set of Lagrange multipliers which is an unknown parameter. Since we have constraints, e.g. vertex constraint that tracks come from a common vertex, χ^2 is minimized with respect to not only $\boldsymbol{\alpha}$ but also $\boldsymbol{\lambda}$. This procedure is called the least square method using the Lagrange multiplier method.

In general, the minimization cannot be done analytically since the constraints are not linearized. However we can obtain a solution, where the χ^2 is minimized, analytically by expanding the constraints around a convenient point.

D.2 Vertex Fit

We describe the method of the fitting with constraints for unknown parameters. Then we apply it to the case of a vertex constraint.

A parameter $\boldsymbol{\alpha}$ represents a set of tracks and a parameter $\boldsymbol{v}$ represents vertex. A constraint equation $\boldsymbol{H}(\boldsymbol{\alpha}, \boldsymbol{v}) = 0$ can be expanded around $\boldsymbol{\alpha}_A$, $\boldsymbol{v}_A$ to give a linearized equation,

$$\begin{aligned}\boldsymbol{H}(\boldsymbol{\alpha}, \boldsymbol{v}) &\simeq \boldsymbol{H}(\boldsymbol{\alpha}_A, \boldsymbol{v}_A) + \frac{\partial \boldsymbol{H}(\boldsymbol{\alpha}_A, \boldsymbol{v}_A)}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha} - \boldsymbol{\alpha}_A) + \frac{\partial \boldsymbol{H}(\boldsymbol{\alpha}_A, \boldsymbol{v}_A)}{\partial \boldsymbol{v}}(\boldsymbol{v} - \boldsymbol{v}_A) \\ &\equiv \boldsymbol{d} + \boldsymbol{D}\delta\boldsymbol{\alpha} + \boldsymbol{E}\delta\boldsymbol{v}\end{aligned}$$

where $\delta\boldsymbol{\alpha} \equiv \boldsymbol{\alpha} - \boldsymbol{\alpha}_A$, $\delta\boldsymbol{v} \equiv \boldsymbol{v} - \boldsymbol{v}_A$, $\boldsymbol{E} \equiv \frac{\partial \boldsymbol{H}(\boldsymbol{\alpha}_A, \boldsymbol{v}_A)}{\partial \boldsymbol{v}}$, and $\boldsymbol{D} \equiv \frac{\partial \boldsymbol{H}(\boldsymbol{\alpha}_A, \boldsymbol{v}_A)}{\partial \boldsymbol{\alpha}}$.

The covariance matrix for the $\boldsymbol{v}$ is represented by $\boldsymbol{L}$. The linearized χ^2 is

$$\begin{aligned}\chi^2 &= (\boldsymbol{\alpha} - \boldsymbol{\alpha}_0)^T \boldsymbol{V}_{\alpha_0}^{-1}(\boldsymbol{\alpha} - \boldsymbol{\alpha}_0) + (\boldsymbol{v} - \boldsymbol{v}_0)^T \boldsymbol{L}^{-1}(\boldsymbol{v} - \boldsymbol{v}_0) + 2\boldsymbol{\lambda}^T(\boldsymbol{D}\delta\boldsymbol{\alpha} + \boldsymbol{E}\delta\boldsymbol{v} + \boldsymbol{d}) \\ &\equiv (\tilde{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}_0})^T \boldsymbol{V}_{\tilde{\alpha_0}}^{-1}(\tilde{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}_0}) + 2\boldsymbol{\lambda}^T(\tilde{\boldsymbol{D}}\delta\tilde{\boldsymbol{\alpha}} + \boldsymbol{d})\end{aligned}$$

where the quantities with ˜ over them have the $\boldsymbol{v}$ information appended to them as follows:

$$(\delta)\tilde{\boldsymbol{\alpha}} \equiv \begin{pmatrix} (\delta)\boldsymbol{\alpha} \\ (\delta)\boldsymbol{v} \end{pmatrix}, \quad \tilde{\boldsymbol{D}} \equiv \begin{pmatrix} \boldsymbol{D} & \boldsymbol{E} \end{pmatrix}, \quad \boldsymbol{V}_{\tilde{\alpha_0}} \equiv \begin{pmatrix} \boldsymbol{V}_{\alpha_0} & 0 \\ 0 & \boldsymbol{L} \end{pmatrix}.$$

The solution can be found by minimizing the χ^2 with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\lambda}$.

$$\begin{aligned}\boldsymbol{V}_{\tilde{\alpha_0}}^{-1}(\tilde{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}_0}) + \tilde{\boldsymbol{D}}^T\boldsymbol{\lambda} &= 0 \\ \tilde{\boldsymbol{D}}\delta\tilde{\boldsymbol{\alpha}} + \boldsymbol{d} &= 0\end{aligned}$$

Therefore we obtain the equations as follows.

$$\begin{aligned}\tilde{\boldsymbol{\alpha}} &= \tilde{\boldsymbol{\alpha}_0} - V_{\tilde{\alpha}_0}\tilde{\boldsymbol{D}}^T\boldsymbol{\lambda} \\ \boldsymbol{\lambda} &= \boldsymbol{V}_{\tilde{D}}(\tilde{\boldsymbol{D}}\delta\tilde{\boldsymbol{\alpha}_0} + \boldsymbol{d}) = \boldsymbol{V}_{\tilde{D}}(\boldsymbol{D}\delta\boldsymbol{\alpha}_0 + \boldsymbol{E}\delta\boldsymbol{v}_0 + \boldsymbol{d}) \\ \boldsymbol{V}_{\tilde{D}} &\equiv (\tilde{\boldsymbol{D}}\boldsymbol{V}_{\tilde{\alpha_0}}\tilde{\boldsymbol{D}}^T)^{-1} = (\boldsymbol{D}\boldsymbol{V}_{\alpha_0}\boldsymbol{D}^T + \boldsymbol{E}\boldsymbol{L}\boldsymbol{E}^T)^{-1} \\ \boldsymbol{V}_{\tilde{\alpha}} &= \boldsymbol{V}_{\tilde{\alpha_0}} - \boldsymbol{V}_{\tilde{\alpha_0}}\tilde{\boldsymbol{D}}^T\boldsymbol{V}_{\tilde{D}}\tilde{\boldsymbol{D}}\boldsymbol{V}_{\tilde{\alpha_0}} \\ \chi^2 &= \boldsymbol{\lambda}^T\boldsymbol{V}_{\tilde{D}}^{-1}\boldsymbol{\lambda} = \boldsymbol{\lambda}^T(\tilde{\boldsymbol{D}}\delta\tilde{\boldsymbol{\alpha}_0} + \boldsymbol{d}) \\ &= \boldsymbol{\lambda}^T(\boldsymbol{D}\delta\boldsymbol{\alpha}_0 + \boldsymbol{E}\delta\boldsymbol{v}_0 + \boldsymbol{d})\end{aligned}$$

The expressions for $\boldsymbol{\alpha}, \boldsymbol{v}, \boldsymbol{V}_\alpha, cov(\boldsymbol{v}, \boldsymbol{\alpha})$ and $\boldsymbol{V}_v$ can be extracted from the full solution using

$$\delta\tilde{\boldsymbol{\alpha}} = \begin{pmatrix} \boldsymbol{\alpha} - \boldsymbol{\alpha}_A \\ \boldsymbol{v} - \boldsymbol{v}_A \end{pmatrix}, \boldsymbol{V}_{\delta\tilde{\alpha}} = \boldsymbol{V}_{\tilde{\alpha}} = \begin{pmatrix} \boldsymbol{V}_\alpha & cov(\boldsymbol{\alpha}, \boldsymbol{v}) \\ cov(\boldsymbol{v}, \boldsymbol{\alpha}) & \boldsymbol{V}_v \end{pmatrix}, \boldsymbol{V}_{\delta\tilde{\alpha_0}} = \boldsymbol{V}_{\tilde{\alpha_0}},$$

which gives

$$
\begin{aligned}
\boldsymbol{\alpha} &= \boldsymbol{\alpha}_0 - \boldsymbol{V}_{\alpha_0}\boldsymbol{D}^T\boldsymbol{\lambda} \\
\boldsymbol{v} &= \boldsymbol{v}_0 - \boldsymbol{L}\boldsymbol{E}^T\boldsymbol{\lambda} \\
\boldsymbol{V}_{\alpha} &= \boldsymbol{V}_{\alpha_0} - \boldsymbol{V}_{\alpha_0}\boldsymbol{D}^T\boldsymbol{V}_{\tilde{D}}\boldsymbol{D}\boldsymbol{V}_{\alpha_0} \\
\boldsymbol{V}_{v} &= \boldsymbol{L} - \boldsymbol{L}\boldsymbol{E}^T\boldsymbol{V}_{\tilde{D}}\boldsymbol{E}\boldsymbol{L} \\
cov(\boldsymbol{v},\boldsymbol{\alpha}) &= -\boldsymbol{L}\boldsymbol{E}^T\boldsymbol{V}_{\tilde{D}}\boldsymbol{D}\boldsymbol{V}_{\alpha_0} \\
\chi^2 &= \boldsymbol{\lambda}^T(\boldsymbol{D}\delta\boldsymbol{\alpha}_0 + \boldsymbol{E}\delta\boldsymbol{v}_0 + \boldsymbol{d}) \\
\boldsymbol{\lambda} &= \boldsymbol{V}_{\tilde{D}}(\boldsymbol{D}\delta\boldsymbol{\alpha}_0 + \boldsymbol{E}\delta\boldsymbol{v}_0 + \boldsymbol{d}) \\
\boldsymbol{V}_{\tilde{D}} &= (\boldsymbol{D}\boldsymbol{V}_{\alpha_0}\boldsymbol{D}^T + \boldsymbol{E}\boldsymbol{L}\boldsymbol{E}^T)^{-1} \\
\boldsymbol{V}_{D} &\equiv (\boldsymbol{D}\boldsymbol{V}_{\alpha_0}\boldsymbol{D}^T)^{-1} \\
\boldsymbol{V}_{E} &\equiv (\boldsymbol{E}^T\boldsymbol{V}_{D}\boldsymbol{E})^{-1} \\
\delta\boldsymbol{\alpha}_0 &\equiv \boldsymbol{\alpha}_0 - \boldsymbol{\alpha}_A \\
\delta\boldsymbol{v}_0 &\equiv \boldsymbol{v}_0 - \boldsymbol{v}_A.
\end{aligned}
$$

Since $\boldsymbol{v}$ is unknown parameters, we set $\boldsymbol{L}$ to ∞ and then obtain

$$
\begin{aligned}
\boldsymbol{\alpha} &= \boldsymbol{\alpha}_0 - \boldsymbol{V}_{\alpha_0}\boldsymbol{D}^T\boldsymbol{\lambda} \\
\boldsymbol{v} &\simeq \boldsymbol{v}_A - \boldsymbol{V}_{E}\boldsymbol{E}^T\boldsymbol{\lambda}_0 \\
\boldsymbol{V}_{\alpha} &= \boldsymbol{V}_{\alpha_0} - \boldsymbol{V}_{\alpha_0}\boldsymbol{D}^T\boldsymbol{V}_{\tilde{D}}\boldsymbol{D}\boldsymbol{V}_{\alpha_0} \\
\boldsymbol{V}_{v} &\simeq \boldsymbol{V}_{E} \\
cov(\boldsymbol{v},\boldsymbol{\alpha}) &\simeq -\boldsymbol{V}_{E}\boldsymbol{E}^T\boldsymbol{V}_{D}\boldsymbol{D}\boldsymbol{V}_{\alpha_0} \\
\chi^2 &\simeq \boldsymbol{\lambda}_0^T(\boldsymbol{D}\delta\boldsymbol{\alpha}_0 + \boldsymbol{E}\delta\boldsymbol{v}_0 + \boldsymbol{d}) \\
\boldsymbol{\lambda} &\simeq \boldsymbol{\lambda}_0 - \boldsymbol{V}_{D}\boldsymbol{E}\boldsymbol{V}_{E}\boldsymbol{E}^T\boldsymbol{\lambda}_0 \\
\boldsymbol{V}_{\tilde{D}} &\simeq \boldsymbol{V}_{D} - \boldsymbol{V}_{D}\boldsymbol{E}\boldsymbol{V}_{E}\boldsymbol{E}^T\boldsymbol{V}_{D} \\
\boldsymbol{V}_{D} &\equiv (\boldsymbol{D}\boldsymbol{V}_{\alpha_0}\boldsymbol{D}^T)^{-1} \\
\boldsymbol{V}_{E} &\equiv (\boldsymbol{E}^T\boldsymbol{V}_{D}\boldsymbol{E})^{-1} \\
\boldsymbol{\lambda}_0 &\equiv \boldsymbol{V}_{D}(\boldsymbol{D}\delta\boldsymbol{\alpha}_0 + \boldsymbol{d}) \\
\delta\boldsymbol{\alpha}_0 &\equiv \boldsymbol{\alpha}_0 - \boldsymbol{\alpha}_A \\
\delta\boldsymbol{v}_0 &\equiv \boldsymbol{v}_0 - \boldsymbol{v}_A.
\end{aligned}
$$

In general, the solutions with $\boldsymbol{L} \to \infty$ are used for the vertex fit. On the other hand, when we know vertex information, e.g. IP profile, the solutions with the finite $\boldsymbol{L}$ can be used to find a vertex.

We show the concrete matrices $\boldsymbol{D}, \boldsymbol{E}, \boldsymbol{d}$ to obtain vertices. When all tracks come from a common vertex $\boldsymbol{v} = (v_x, v_y, v_z)$, each track i, which has momentum $\boldsymbol{p} = (p_x, p_y, p_z)$ at the position $\boldsymbol{x} = (x, y, z)$, can be represented by

$$p_{xi}\Delta y_i - p_{yi}\Delta x_i - \frac{a_i}{2}(\Delta x_i^2 + \Delta y_i^2) = 0$$
$$\Delta z_i p_{\mathrm{t}i} - \frac{p_{zi}p_{\mathrm{t}i}}{a_i}\sin^{-1}[a_i(p_{xi}\Delta x_i + p_{yi}\Delta y_i)/p_{\mathrm{t}i}^2] = 0$$

where $p_{\mathrm{t}} = \sqrt{p_x^2 + p_y^2}$, $\Delta x_i = v_x - x_i$ etc. In case of the neutral track,

$$p_{xi}\Delta y_i - p_{yi}\Delta x_i = 0$$
$$\Delta z_i p_{\mathrm{t}i} - \frac{p_{zi}}{p_{\mathrm{t}i}}(p_{xi}\Delta x_i + p_{yi}\Delta y_i) = 0.$$

From these constraints, $\boldsymbol{D}_i$, $\boldsymbol{E}_i$ and $\boldsymbol{d}_i$ can be calculated as,

$$\boldsymbol{D}_i = \begin{pmatrix} \Delta y_i & -\Delta x_i & 0 \\ -p_{zi}p_{\mathrm{t}i}S_iR_{xi} + U_i\frac{p_{xi}}{p_{\mathrm{t}i}} & -p_{zi}p_{\mathrm{t}i}S_iR_{yi} + U_i\frac{p_{yi}}{p_{\mathrm{t}i}} & -\frac{p_{\mathrm{t}i}sin^{-1}B_i}{a_i} \\ p_{yi} + a_i\Delta x_i & -p_{xi} + a_i\Delta y_i & 0 \\ p_{xi}p_{zi}p_{\mathrm{t}i}S_i & p_{yi}p_{zi}p_{\mathrm{t}i}S_i & -p_{\mathrm{t}i} \end{pmatrix}$$

$$\boldsymbol{E}_i = \begin{pmatrix} -p_{yi} - a_i\Delta x_i & p_{xi} - a_i\Delta y_i & 0 \\ -p_{xi}p_{zi}p_{\mathrm{t}i}S_i & -p_{yi}p_{zi}p_{\mathrm{t}i}S_i & p_{\mathrm{t}i} \end{pmatrix}$$

$$\boldsymbol{d}_i = \begin{pmatrix} A_{1i} - \frac{a_i}{2}(\Delta x_i^2 + \Delta y_i^2) \\ U_i p_{\mathrm{t}i} \end{pmatrix}.$$

S_i etc. are defined as follows[a].

$$A_{1i} = p_{xi}\Delta y_i - p_{yi}\Delta x_i$$
$$A_{2i} = p_{xi}\Delta x_i + p_{yi}\Delta y_i$$
$$B_i = a_i A_{2i}/p_{\mathrm{t}i}^2$$
$$R_{xi} = \Delta x_i - 2p_{xi}A_{2i}/p_{\mathrm{t}i}^2$$
$$R_{yi} = \Delta y_i - 2p_{yi}A_{2i}/p_{\mathrm{t}i}^2$$
$$S_i = \frac{1}{p_{\mathrm{t}i}^2\sqrt{1 - B_i^2}}$$
$$U_i = \Delta z_i - \frac{p_{zi}}{a_i}\sin^{-1} B_i$$

We can obtain vertices using matrices $\boldsymbol{D}$, $\boldsymbol{E}$, $\boldsymbol{d}$ with a few iterations since we expand the constraints around a convenient point.

D.3 Check

We check whether the decay and production vertices are properly obtained from the vertex fit using the MC sample of $D^0 \to K^-\pi^+$. Figures D.1 and D.2 show the difference

[a]Neutral track: $a = 0 \to B = 0 \to \sin^{-1} B = 0$ and $U_i = \Delta z_i - \frac{p_{zi}A_{2i}}{p_{\mathrm{t}i}^2}$

between the vertex obtained by the vertex fit and the true vertex from the generator information. We fit the distributions with Gaussians and the mean values are listed in Table D.1. The mean values are consistent with zero within two standard deviations and are much smaller than the errors of the D decay length (Section 4.3.2).

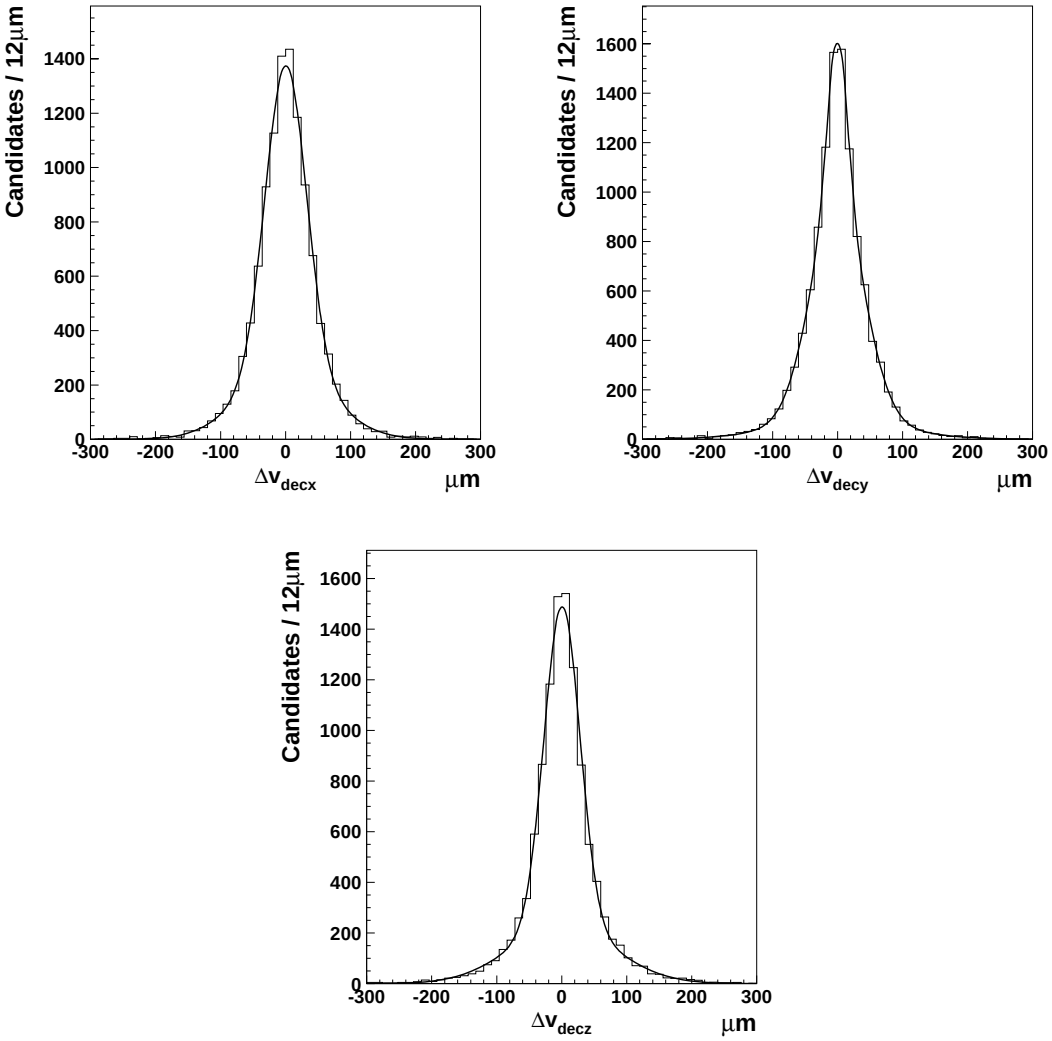


Figure D.1: Difference Δv between the decay vertex obtained by the vertex fit and the true decay vertex.

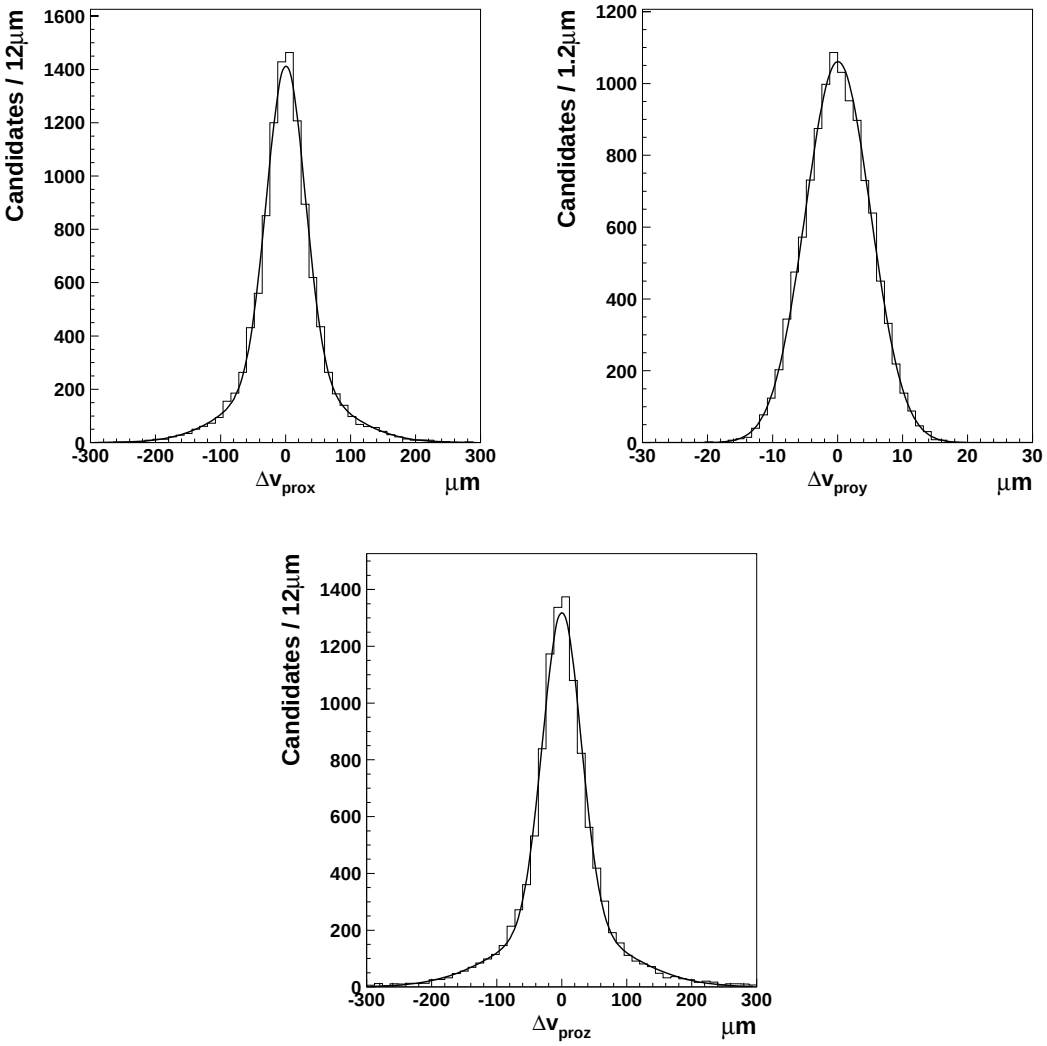


Figure D.2: Difference Δv between the decay vertex obtained by the production vertex fit and the true production vertex.

Table D.1: Mean values of the difference distributions Δv.

	x (μm)	y (μm)	z (μm)
Decay vertex	0.51 ± 0.40	-0.20 ± 0.36	0.49 ± 0.37
Production vertex	-0.64 ± 0.39	0.070 ± 0.048	0.28 ± 0.42

To demonstrate that the production vertex is properly obtained for real data, Figure D.3 shows the distribution of the difference between the production vertex of $D^0 \to K^-\pi^+$ and the position of the IP profile in Experiment 7. The root-mean-square (rms) of the distribution should be similar to the size of the IP profile (Fig 7.1). Table D.2 shows

the values of the rms of the distribution. As expected, they are similar to the the size of the IP profile. The shape of the IP profile is well reproduced by the distribution of production vertices.

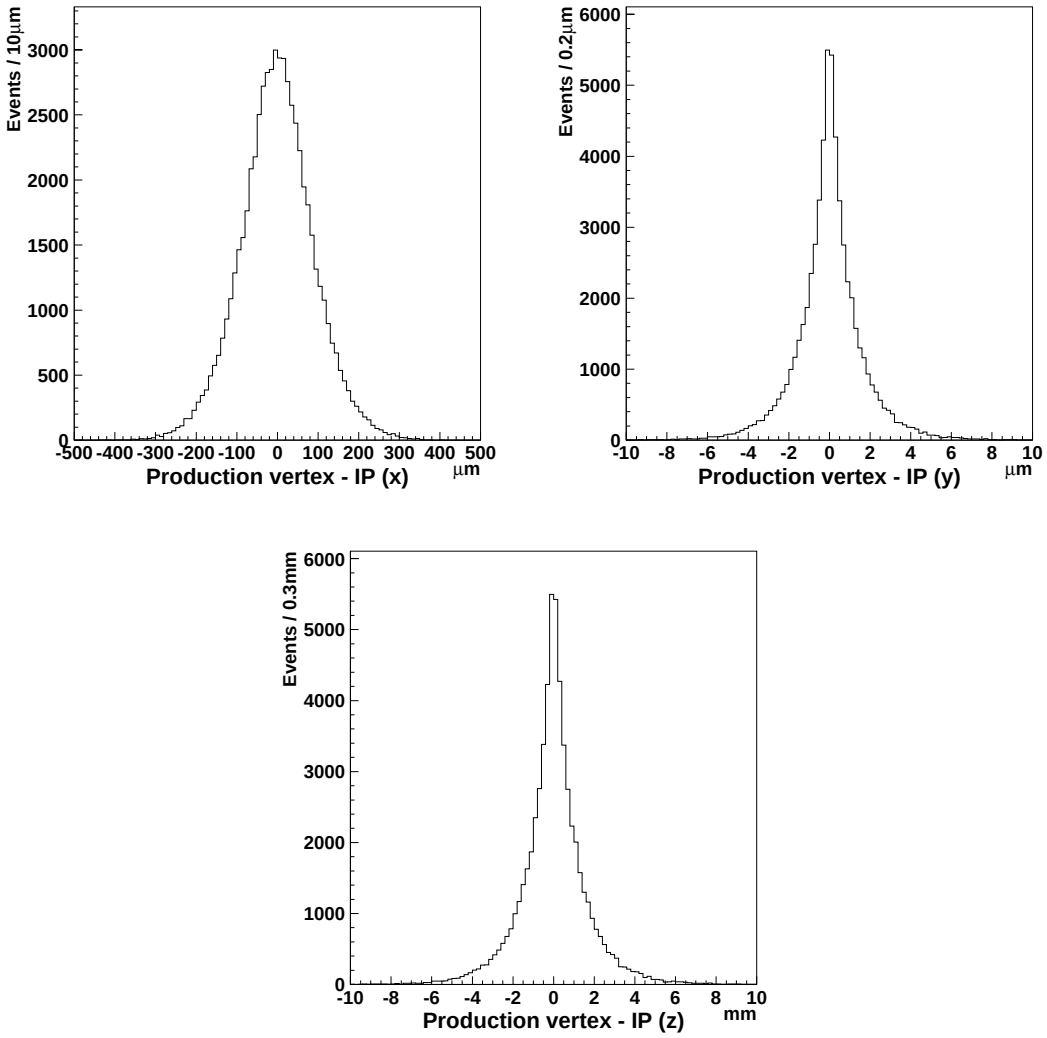


Figure D.3: Difference between the production vertex and the position of the IP profile.

Table D.2: Root-mean-square of the distributions shown in Fig D.3.

	x (μm)	y (μm)	z (mm)
rms	91	1.7	3.8

Appendix E

Calibration of Track Error with Cosmic Ray

We describe how to calibrate properly errors of the charged tracks reconstructed by our tracking programs.

Track errors which are obtained by our tracking programs can be tested by cosmic ray events that are reconstructed by the same tracking programs. A cosmic ray is reconstructed as two tracks by our tracking programs. In principle, two tracks must have the same track parameters at the same position, therefore the difference between the two track parameters should be zero and an error of the difference can be calculated and then can be tested by the "pull" distribution as follows.

In our experiment, a track is represented by helix parameters [57]:

$$(d_\rho, \phi_0, \kappa, d_z, \tan\lambda).$$

Two tracks of one cosmic ray are called "up-side" track and "down-side" track according to the position of track trajectory, that is, $y > 0$ and $y < 0$ respectively. We define the following variables whose distributions are called "pull" distributions.

$$\frac{d_\rho^{\rm up} + d_\rho^{\rm down}}{\sqrt{\sigma^2_{d_\rho^{\rm up}} + \sigma^2_{d_\rho^{\rm down}}}}$$

$$\frac{\phi_0^{\rm up} - \phi_0^{\rm down}}{\sqrt{\sigma^2_{\phi_0^{\rm up}} + \sigma^2_{\phi_0^{\rm down}}}}$$

$$\frac{\kappa^{\rm up} + \kappa^{\rm down}}{\sqrt{\sigma^2_{\kappa^{\rm up}} + \sigma^2_{\kappa^{\rm down}}}}$$

$$\frac{d_z^{\rm up} - d_z^{\rm down}}{\sqrt{\sigma^2_{d_z^{\rm up}} + \sigma^2_{d_z^{\rm down}}}}$$

$$\frac{\tan\lambda^{\rm up} + \tan\lambda^{\rm down}}{\sqrt{\sigma^2_{\tan\lambda^{\rm up}} + \sigma^2_{\tan\lambda^{\rm down}}}}$$

We check $\sigma_{\rm pull}$ of these pull distributions as a function of $p_{\rm t}$ and $\tan\lambda$, where a $\sigma_{\rm pull}$ is a width of the Gaussian when the pull distribution is fitted with a Gaussian. If our track errors are proper, the values of $\sigma_{\rm pull}$ have no dependence and are a flat at $\sigma_{\rm pull}$=1.

Figures E.1 and E.2 show $\sigma_{\rm pull}$ as a function of $p_{\rm t}$ and $\tan\lambda$ before the calibration. We can see clear dependences and results of $\sigma_{\rm pull} \neq 1$ in most distributions. We fit these dependences with the proper functions and then calibrated track errors. Figures E.3 and E.4 show the distributions after the calibration. To obtain them, we use a different data sample to confirm our calibrations. The dependences disappear and the value of $\sigma_{\rm pull}$ is almost 1. We use the calibrated track errors in our lifetime and y_{CP} measurements.

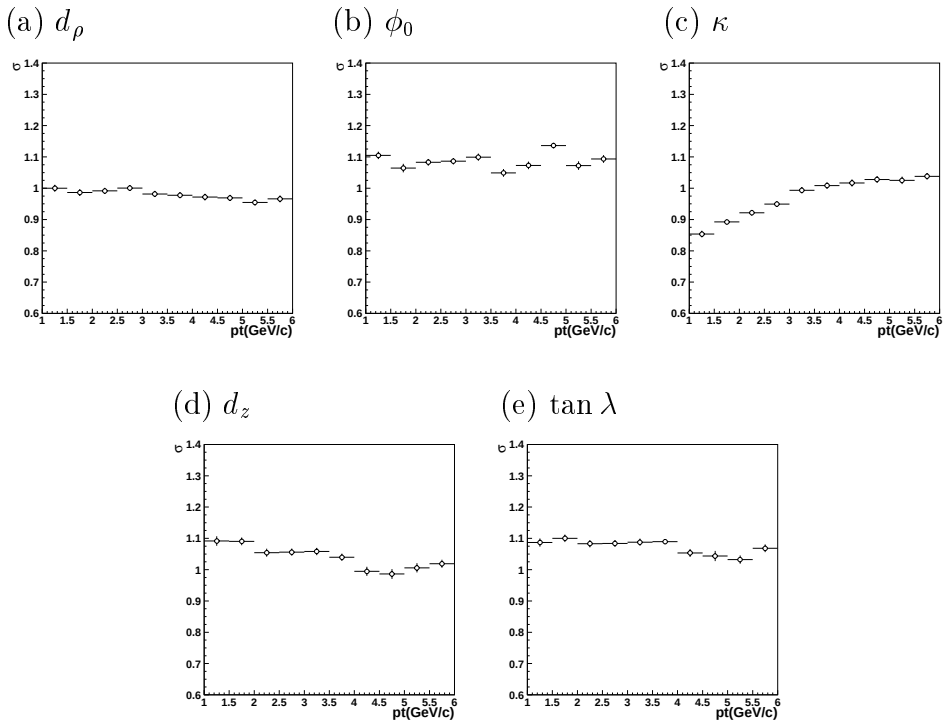


Figure E.1: σ of pull distributions as a function of $p_{\rm t}$ before the calibration: (a) d_ρ, (b) ϕ_0, (c) κ, (d) d_z and (e) $\tan\lambda$.

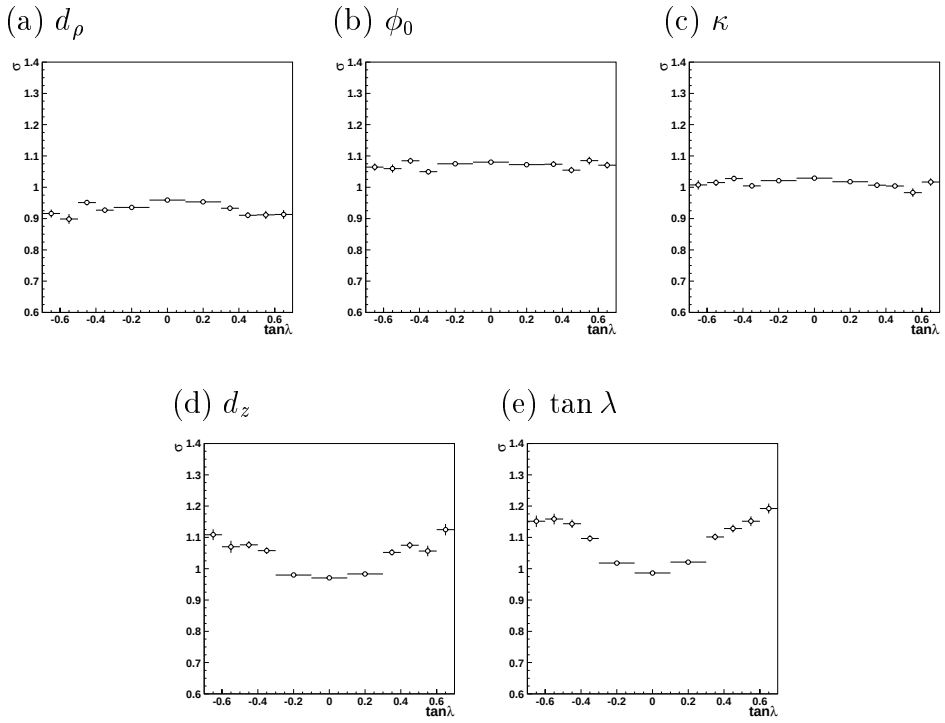


Figure E.2: σ of pull distributions as a function of $\tan\lambda$ before the calibration: (a) d_ρ, (b) ϕ_0, (c) κ, (d) d_z and (e) $\tan\lambda$.

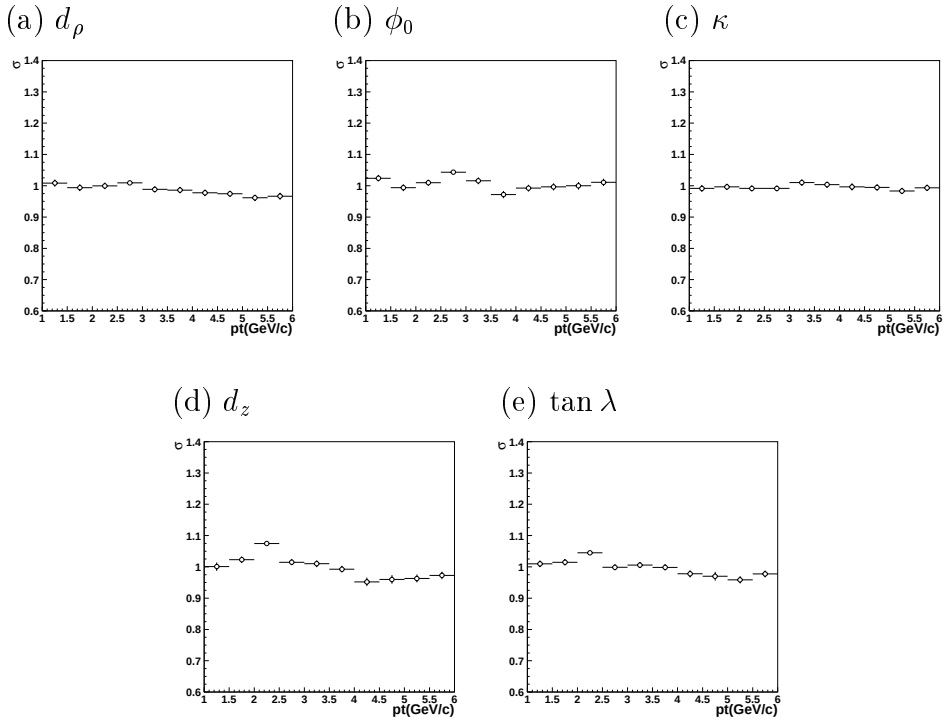


Figure E.3: σ of pull distributions as a function of p_t after the calibration: (a) d_ρ, (b) ϕ_0, (c) κ, (d) d_z and (e) $\tan\lambda$.

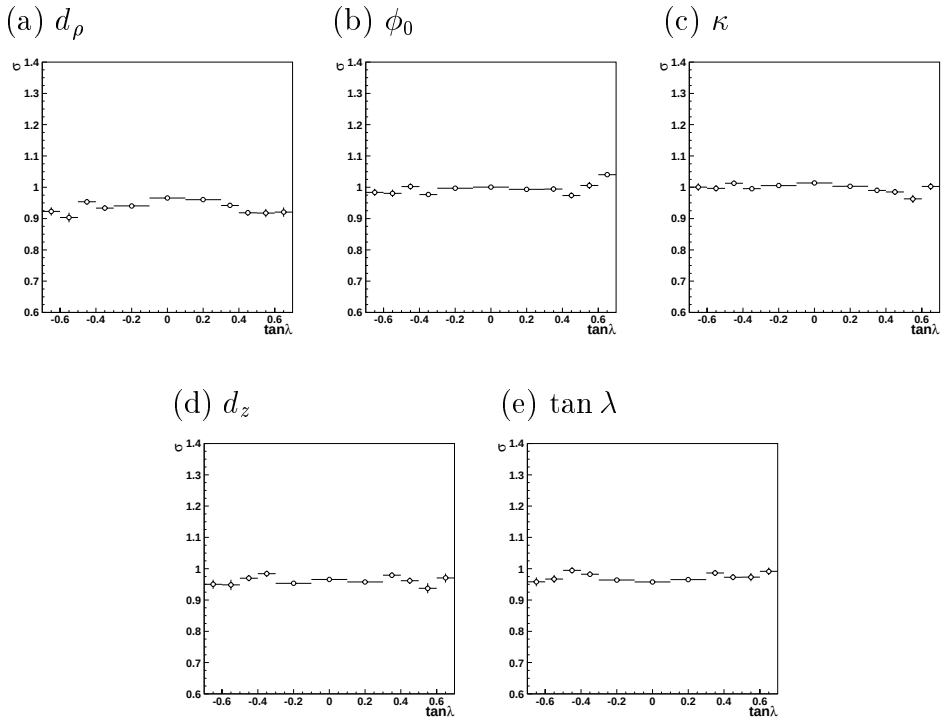


Figure E.4: σ of pull distributions as a function of $\tan\lambda$ after the calibration: (a) d_ρ, (b) ϕ_0, (c) κ, (d) d_z and (e) $\tan\lambda$.

Appendix F

Toy Monte Carlo Experiment

Our fit method is checked by performing many "toy Monte Carlo experiments." Three hundred sets of Monte Carlo data with the same statistics as the data are generated with the probability density function obtained from the lifetime or y_{CP} fit. Figures F.1 show the distributions of the difference between the fitted value and the input value in the toy Monte Carlo experiments and the solid curve is a result of a Gaussian fit. Table F.1 shows mean values of these distributions. Our fit method is not wrong since their mean values are consistent with zero. Figures F.2 show the distributions of the difference between the fitted value and the input value divided by the fitted error in the toy Monte Carlo experiments. These distributions are well represented by a fit with a Gaussian as shown in the figure by the overlaid solid curve. The standard deviation ($= \sigma$) of the Gaussian will be 1 if the error estimation is correct. Since our σ values are almost 1 as shown in Table F.2, our error estimations are reasonable.

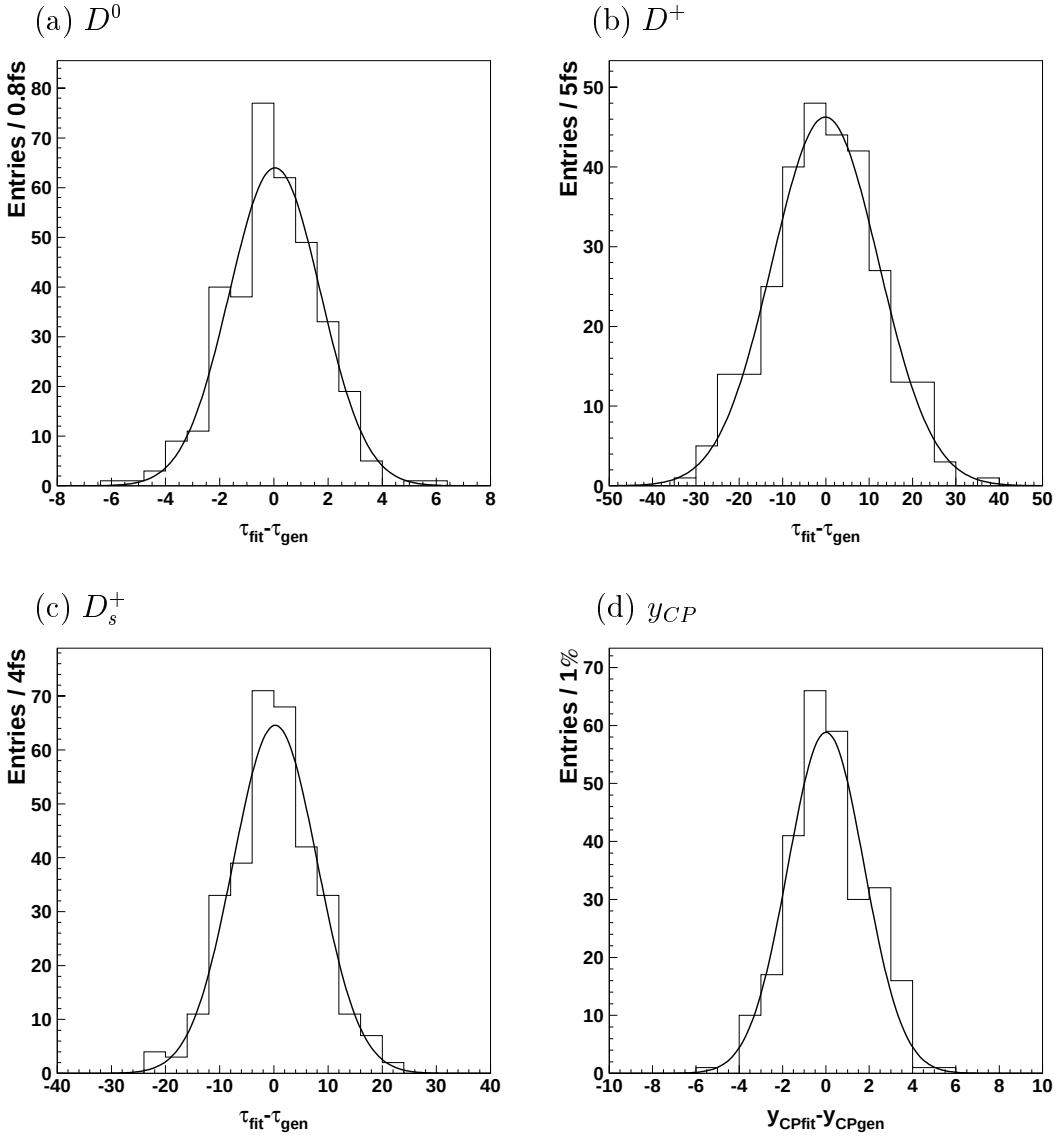


Figure F.1: Residual ($\tau_{\rm fit} - \tau_{\rm gen}$ or $y_{CP\rm fit} - y_{CP\rm gen}$) distributions obtained from the toy MC experiments: (a) D^0, (b) D^+, (c) D_s^+ and (d) y_{CP}.

Table F.1: Mean values of the difference distributions obtained from the toy MC experiments.

Measurement	"fitted−input" value
$D^0 \to K^-\pi^+$	0.04 ± 0.09 fs
D^+	-0.11 ± 0.75 fs
D_s^+	0.26 ± 0.44 fs
y_{CP}	$0.073 \pm 0.059\%$

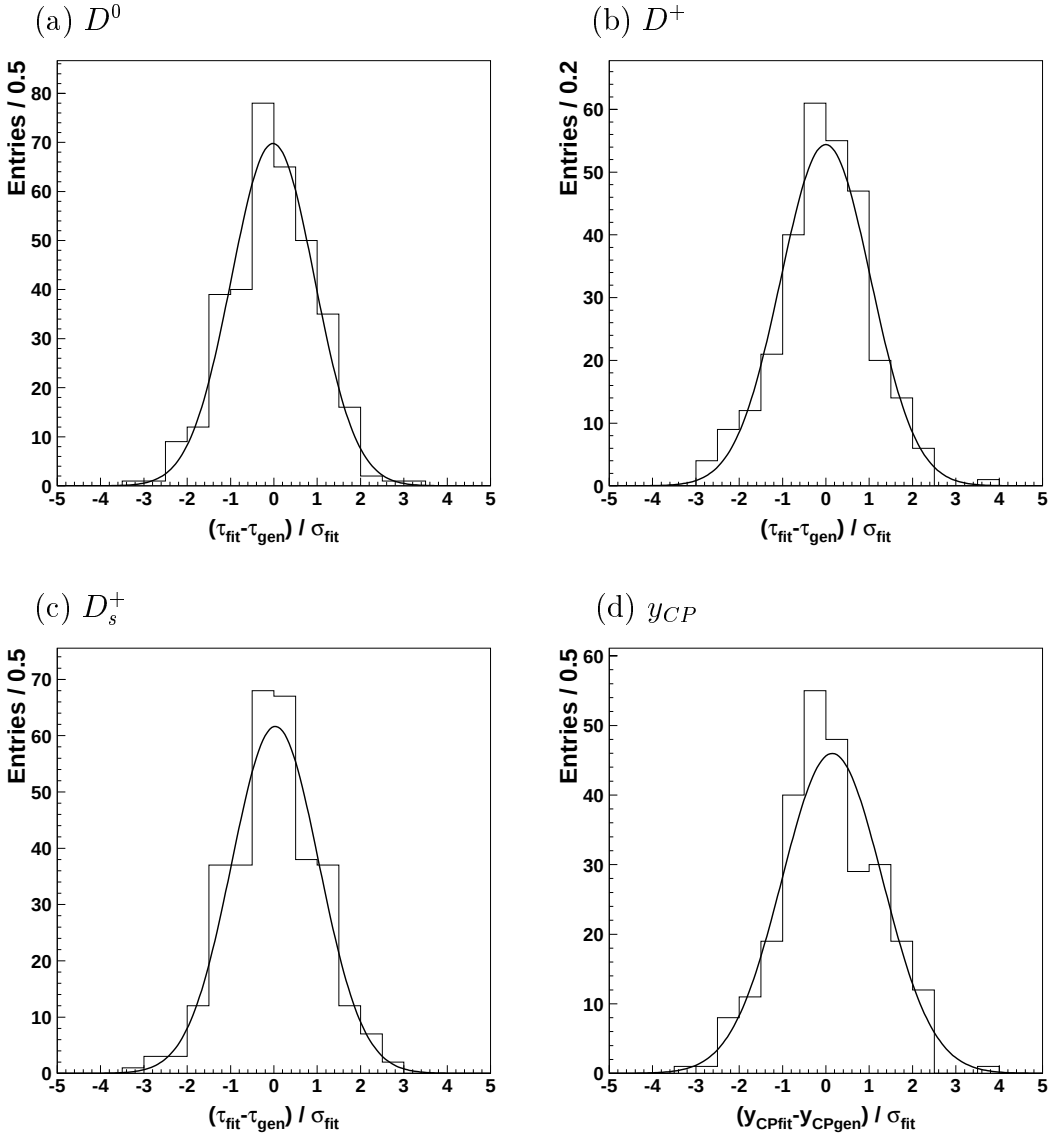


Figure F.2: Pull $((\tau_{\rm fit}-\tau_{\rm gen})/\sigma_{\rm fit}$ or $(y_{CP\rm fit}-y_{CP\rm gen})/\sigma_{\rm fit})$ distributions obtained from the toy MC experiments: (a) D^0, (b) D^+, (c) D_s^+ and (d) y_{CP}.

Table F.2: σ of pull distributions obtained from the toy MC experiments.

Measurement	σ
$D^0 \to K^-\pi^+$	0.96 ± 0.04
D^+	1.04 ± 0.06
D_s^+	1.01 ± 0.05
y_{CP}	1.03 ± 0.05

List of Tables

List of Figures

References

[1] S.L. Glashow, Nucl. Phys. **22**, 579 (1961); S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, *Elementary Particle Theory*, edited by N. Svartholm (Almqvist and Wiksell, Stockholm, 1968), p. 367.

[2] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[3] B. Aubert *et al.*, (BABAR) Phys. Rev. Lett. **87**, 091801 (2001); K. Abe *et al.*, (Belle) Phys. Rev. Lett. **87**, 091802 (2001).

[4] P.W. Higgs, Phys. Lett. **12**, 132 (1964), Phys. Rev. Lett. **13**, 508 (1964) and Phys. Rev. **145**, 1156 (1966); F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964); G.S. Guralnik, C.R. Hagen and T.W.B. Kibble, Phys. Rev. Lett. **13**, 585 (1964).

[5] W. Bacino *et al.*, Phys. Rev. Lett. **45**, 329 (1980); N. Ushida *et al.*, Phys. Rev. Lett. **45**, 1053 (1980); M. Roos *et al.*, Phys. Lett. B **111**, 1 (1982). The last one is the old review of particle properties by Particle Data Group. In this edition, D lifetimes are listed as $\tau(D^+) = 910^{+220}_{-150}$ fs, $\tau(D^0) = 480^{+240}_{-150}$ fs, and $\tau(D_s^+) = 220^{+280}_{-110}$ fs.

[6] Particle Data Group, D.E. Groom *et al.*, Eur. Phys. J. **C15**, 1 (2000) and 2001 off-year partial update for the 2002 edition available on the PDG WWW pages (`http://pdg.lbl.gov/`).

[7] G. Bellini, I.I. Bigi and P.J. Dornan, Phys. Rep. **289**, 1 (1997).

[8] J.M. Link *et al.*, (FOCUS) Phys. Lett. B **485**, 62 (2000).

[9] L. Gibbons, hep-ex/0107079, the workshop on "e^+e^- Physics at Intermediate Energies", SLAC, May 2001;
See `http://www.lns.cornell.edu/public/CLEO/spoke/CLEOc/`.

[10] In the Standard Model leptons are preferentially left-handed (anti-parallel of momentum and spin) and anti-leptons are preferentially right-handed (parallel of momentum and spin). In the decay of spin zero particle, angular momentum conservation forces the two outcoming fermions to have the same helicity, i.e., to be both left-handed or both right-handed. This means that one of the fermions is forced in the wrong helicity state, causing the decay to be suppressed.

[11] "D" represents "charm" or refers to D^0, D^+ and D_s^+ mesons throughout this thesis.

[12] I.I. Bigi, hep-ph/9612293, workshop on "Heavy Quarks at Fixed Target" (HQ96), Germany, Oct. 1996.

[13] I.I. Bigi, N.G. Uraltsev, Nucl. Phys. **B423**, 33 (1994).

[14] H.Y. Cheng and K.C. Yang, Phys. Rev. D **61**, 014008 (2000).

[15] For instance, one or several gluons (with the small momentum transfer) can be emitted before the W annihilation leaving the $c\overline{s}$ in a spin 1 state.

[16] E. Golowich, hep-ph/9706548, conference on "B Physics and CP Violation", Hawaii, March 1997.

[17] S. Bergmann, Y. Grossman, Z. Ligeti, Y. Nir and A.A. Petrov, Phys. Lett. B **486**, 418 (2000).

[18] H. Yamamoto, Belle note 463, unpublished.

[19] L. Wolfenstein, Phys. Rev. Lett. **75**, 2460 (1995). When the final state phases for $D^0 \to K^-\pi^+$ and $\overline{D}^0 \to K^-\pi^+$ are given by δ_1 and δ_2, the final state phase for $D^0 \to K^+\pi^-$ becomes δ_2 by the CP transformation, and by the interchange of the quarks s and d in the effective operators and in the final state, the decay amplitude $D^0 \to K^-\pi^+$ becomes the amplitude $D^0 \to K^+\pi^-$. Therefore in the SU(3) flavor symmetry approximation δ_1 is equal to δ_2.

[20] The value of ~ 0.06 is estimated from the following three measurements: (1) the braching fraction of $D^0 \to K^-\pi^+$ and $D^0 \to K^+\pi^-$ [6], (2) the wrong-sign decay rate $R_{\rm ws}$ (M. Grothe hep-ex/0112002), and (3) R_D in [24].

[21] E.M. Aitala *et al.*, (E791) Phys. Rev. Lett. **83**, 32 (1999).

[22] D. Cronin-Hennessy *et al.*, (CLEO) hep-ex/0102006.

[23] B. Aubert *et al.*, (BABAR) hep-ex/0109008.

[24] R. Godang *et al.*, (CLEO) Phys. Rev. Lett. **84**, 5038 (2000).

[25] H. Nelson, hep-ex/9908021, symposium on "Lepton and Photon Interactions at High Energies" (LP99), 1999.

[26] I.I. Bigi, N.G. Uraltsev, Nucl. Phys. **B592**, 92 (2001).

[27] KEKB accelerator group, KEKB B-Factory Design Report, KEK Report 95-7 (1995).

[28] R. Giles *et al.*, Phys. Rev. D **29**, 1285 (1985).

[29] H. Hayashii *et al.*, Phys. Lett. B **314**, 149 (1993).

[30] Belle Collaboration, Technical Design Report, KEK Report 95-1, (1995).

[31] A. Abashian *et al.*, KEK Progress Report 2000-4, to be submited to NIM.

[32] G. Alimonti *et al.*, Nucl. Inst. and Meth. **A453**, 71 (2000).

[33] H. Hirano *et al.*, Nucl. Inst. and Meth. **A455**, 294 (2000); M. Akatsu *et al.*, Nucl. Inst. and Meth. **A454**, 322 (2000).

[34] T. Iijima *et al.*, Nucl. Inst. and Meth. **A453**, 321 (2000).

[35] H. Kichimi *et al.*, Nucl. Inst. and Meth. **A453**, 315 (2000).

[36] H. Ikeda *et al.*, Nucl. Inst. and Meth. **A441**, 401 (2000).

[37] A. Abashian *et al.*, Nucl. Inst. and Meth. **A449**, 112 (2000).

[38] M. Yamaga, Doctor thesis, Tohoku University (2001).

[39] N. Tan *et al.*, IEEE Trans. Nucl. Sci. **48**, 900 (2001).

[40] R. Harr, IEEE Trans. Nucl. Sci. **42**, 134 (1995).

[41] R. Itoh and S. Ichizawa, Belle note 97, unpublished.

[42] H. Ozaki, Belle note 146, unpublished.

[43] R. Itoh, Belle note 161, unpublished.

[44] S. Nagayama, "Panther – User's guide", `http://belle.kek.jp/~nagayama/panther/`.

[45] See `http://www.lns.cornell.edu/public/CLEO/soft/QQ/`.

[46] R. Itoh, "QQ quick reference for Belle", `http://belle.kek.jp/~software/qq/html/belle_qq.html`.

[47] T. Sjöstrand, Computer Physics Commun. **82**, 74 (1994).

[48] B. Anderson *et al.*, Phys. Rep. **97**, 31 (1983).

[49] The detector response is simulated using GEANT, R. Brun *et al.*, GEANT 3.21, CERN Report No. DD/EE/84-1 (1987).

[50] B. Casey, Belle note 296, unpublished; B. Casey, Belle note 390, unpublished.

[51] F. Fang, Belle note 323, unpublished.

[52] Two Gaussians indicates two Gaussians with the same mean throughout this thesis.

[53] C. G. Fox, S. Wolfrain, Phys. Rev. Lett. **41**, 1581 (1978)

[54] J. Tanaka, Belle note 194, unpublished.

[55] J. Orear, CLNS 82/511.

[56] CERN Program Library Long Writeup D506.

[57] Y. Ohnishi, Belle note 148, unpublished.

[58] P. L. Frabetti *et al.*, (E687) Phys. Lett. B **323**, 459 (1994).

[59] P. L. Frabetti *et al.*, (E687) Phys. Rev. Lett. **71**, 827 (1993).

[60] E.M. Aitala *et al.*, (E791) Phys. Lett. B **445**, 449 (1999).

[61] G. Bonvicini *et al.*, (CLEO) Phys. Rev. Lett. **82**, 4586 (1999).

[62] We obtain a 95% confidence interval as follows. (1) The values of the likelihood function as a function of y_{CP} are convoluted with an asymmetric Gaussian function for the systematic errors. (2) We calculate values of y_{CP} at 2.5% and 97.5% of the total area of these convoluted likelihood values verse y_{CP} as bounds of the 95% confidence interval.

[63] J. Tanaka, Master thesis "Development of the Low Momentum Track Reconstruction Program and the Kinematic Fitter for the Belle Experiment", Department of Physics, University of Tokyo (1999).

[64] KID Group, Belle note 321, unpublished.