Measurement of CP-Violating Asymmetries in the Neutral B Meson Decaying to the $\rho\pi$ State Using a Time-Dependent Dalitz Plot Analysis

Akito Kusaka
Department of Physics, University of Tokyo

December 2006

Abstract

In the standard model (SM), CP violation arises from an irreducible phase in the Cabibbo-Kobayashi-Maskawa (CKM) matrix. A Dalitz plot analysis of the decay $B^0 \to \rho\pi \to \pi^+\pi^-\pi^0$ offers a unique way to determine the angle ϕ_2 in the CKM unitarity triangle without discrete ambiguities (for ϕ_2 in the range between 0 and π), which cannot be obtained from analyses of other modes sensitive to ϕ_2 such as $B \to \pi\pi$ or $B \to \rho\rho$. The Dalitz plot analysis uses isospin and takes into account a possible contamination from $b \to d$ penguin transitions. In addition, using measurements of the related charged decay modes $B^+ \to \rho^+\pi^0$ and $B^+ \to \rho^0\pi^+$ provides further improvement of the ϕ_2 determination.

In this thesis, we present the result of time-dependent Dalitz plot analysis in $B^0 \to \pi^+\pi^-\pi^0$ decays and a constraint on ϕ_2. We use a $414\,\mathrm{fb}^{-1}$ data sample that contains $449 \times 10^6 B\bar{B}$ pairs collected on the $\Upsilon(4S)$ resonance. The data were taken at the KEKB collider and collected with the Belle detector.

By the Dalitz plot analysis, we constrain the relative sizes and phases of the complex amplitudes of $B^0(\overline{B}^0) \to \rho^+\pi^-$, $\rho^-\pi^+$, and $\rho^0\pi^0$ decays, which are denoted by ${}^{(}\overline{A}{}^{)}_{+}$, ${}^{(}\overline{A}{}^{)}_{-}$, and ${}^{(}\overline{A}{}^{)}_{0}$, respectively. The amplitudes are related to ϕ_2 through an isospin relation by

$$e^{+2i\phi_2} = \frac{\overline{A}^+ + \overline{A}^- + 2\overline{A}^0}{A^+ + A^- + 2A^0} \,.$$

Combining our analysis with information on charged B decay modes, we perform a full Dalitz and isospin analysis for the first time and obtain a constraint on the CKM angle ϕ_2,

$$68^\circ < \phi_2 < 95^\circ \,,$$

as the 68.3% confidence interval consistent with the standard model (SM). A large SM-disfavored region also remains. This result is combined with the other measurements from $B \to \pi\pi$ and $B \to \rho\rho$, and its consistency with the SM expectation is examined; we confirm they are consistent with each other at a precision of $\sim 7^\circ$.

Acknowledgments

First, I would like to express my great appreciation to my supervisor, Prof. H. Aihara, for giving me the opportunity to work on this research and his great supervision.

I greatly appreciate Prof. M. Iwasaki (Abe) and Dr. H. Kakuno for their powerful supports and advices on my research works.

I am grateful to the members, students, and OB's in our group at University of Tokyo: Dr. T. Abe, Dr. N. Hastings, Dr. T. Uchida, Mr. K. Tanabe, Ms. Y. Nakahama, Mr. H. Nakayama, Mr. K. Yamada, Mr. H. Miyatake, Dr. T. Higuchi, Dr. J. Tanaka, Dr. M. Yokoyama, Dr. T. Nakadaira, Dr. T. Tomura, Mr. H. Kawai, Mr. N. Uozaki, Mr. Y. Yamashita, Mr. R. Ishida, and Mr. K. Ito. They have been good teachers, capable colleagues, and nice drinking companies.

I would like to express my gratitude to all the members of the CP-fit, ϕ_2, and SVD groups. I thank Prof. T. E. Browder, Prof. M. Hazumi, Dr. H. Ishino, Prof. A. S. Kuzmin, Prof. Y. Sakai, Dr. K. Trabelsi, and Dr. C. C. Wang, who are deeply involved in this analysis.

I appreciate all the members of the Belle collaboration and the KEKB accelerator group for keeping the good performance of the experimental apparatus, great contributions to the development of analysis frameworks, and the high luminosity by the excellent operation.

Finally, I am particularly indebted to my family; I am grateful to my parents for their support and thank my brothers for cheering me up. Without them, I could not accomplish this work.

Contribution of the Author

This thesis was written based on the experimental activities of the Belle members in collaboration with the author. The experimental apparatus described in chapter 3 is developed, integrated, and managed by the Belle collaboration and the KEKB accelerator group, where the author contributed to the integration of the new silicon vertex detector installed in the summer of 2003 and has been responsible for its slow control system. The tools described in Sec. 1, 2, and 3 of chapter 4 are developed by several subgroups of the Belle collaboration and used in various analyzes in common. The author has worked on the study and development of the vertexing algorithm for the new silicon vertex detector and the calibration of the vertexing and flavor tagging tools. The analysis described in chapter 5 and the following chapters are all dedicated to the $B^0 \to (\rho\pi)^0$ time-dependent Dalitz analysis; the author is the primary contributor to this analysis and is responsible for all the contents described there.

Contents

Chapter 1

Introduction

CP violation in the quark sector, now established as a part of the standard model (SM), is one of the best probes to search for the physics beyond the SM, and B-factories are the most exciting playground to investigate the physics of CP violation.

To establish a model that describes CP violation had been a forefront of the particle physics. CP symmetry, which corresponds to the symmetry between matter and anti-matter, had been believed to be conserved for a long time, until the first CP violation in K^0 decays was observed in 1964 [1]. In 1973, Kobayashi and Maskawa proposed a model that implements the CP-violating effect into the SM as an irreducible phase in the quark mixing matrix, or Cabibbo-Kobayashi-Maskawa (CKM) matrix [2, 3]. A profound prediction of this model is the indication of the quarks of third generation, which were yet to be discovered. The b (bottom) and t (top) quarks were discovered in 1977 [4] and 1995 [5, 6], respectively. These discoveries suggested the CKM model was indeed the origin of CP violation. Two conclusive observations coincidently came at the turn of the century: the observation of the direct CP violation in K^0 decays (1999) [7, 8] and the observation of CP violation in B^0 decays measured in B-factories (2001) [9, 10]. The former killed the "Super-Weak" model, the only compelling alternative of the CKM model, and the latter confirmed that the CKM picture describes the CP violation in b-quark decays in a unified way as well as s-quark decays.

Now, with the CKM model established as a part of the SM, we have entered a new era of the physics of CP violation; we use it as a probe sensitive to the physics beyond the SM, or new physics. A reason why it is useful to detect the new physics effect is that there are a large number of observables related to the CP violation, while the CKM model has only four degrees of freedom to describe the CP violating structure of the SM; to examine whether all the observables are universally described by the CKM picture is also a quest for the new physics, since we expect to see deviations from the CKM expectations if there contributes the new physics effect. Another reason is that the CP-violating observables tends to come from higher-order contributions of the SM, such as box and loop diagrams, and thus some models expect large manifestation of the new physics effect in the observables.

Observables related to CP violation are schematically described by the so called CKM triangle, or the Unitarity Triangle (Figure 1.1). With the length of the base normalized to be unity, the triangle is fully characterized by two degrees of freedom, e.g., two angles. Observables correspond to the angles, lengths of the sides, or their products or linear combinations. The three angles of the triangle, ϕ_1, ϕ_2, and ϕ_3, are of special importance, since they manifest themselves as CP-violating effects in the measurements. A variety of observables allows us to check if a single triangle can consistently describe all of them, which is a crucial test of the SM. A typical way of the consistency check is to compare direct measurements with the indirect measurement, the expectation from the CKM picture based on other measurements. For example, having independent measurements of the three angles, $\phi_1^{\text{meas.}}$, $\phi_2^{\text{meas.}}$, and $\phi_3^{\text{meas.}}$,

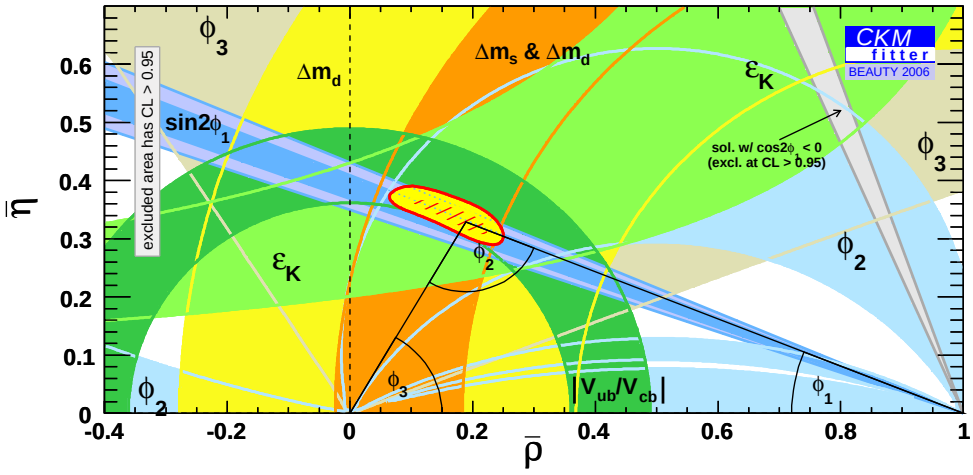


Figure 1.1: CKM triangle and constraints on it from various measurement, as of the summer of 2006, provided by CKMfitter group [11].

we can compare the direct measurement of ϕ_2, $\phi_2^{\rm dir.} \equiv \phi_2^{\rm meas.}$, with the expectation from the CKM triangle ansatz, $\phi_2^{\rm ind.} \equiv \pi - \phi_1^{\rm meas.} - \phi_3^{\rm meas.}$.

In B^0 meson decays[1], large CP-violating effect can be observed in the time-dependent decay rates, which was originally predicted by Carter, Bigi, and Sanda [12, 13, 14]. There are two paths for B^0 to decay to a final state f: a path where B^0 directly decays to f, $B^0 \to f$, and another where B^0 once oscillates to $\overline{B}^0$ and then decays to f, $B^0 \to \overline{B}^0 \to f$. The interference between the two paths can modulate the time-dependent decay rate from $\exp(-t/\tau_{B^0})$, where t is the proper time for the decaying B^0. The CP-violating effect can appear in the pattern of this modulation; the patterns of the modulation are different for the process $B^0 \to (\overline{B}^0) \to f$ and its CP conjugate process $\overline{B}^0 \to (B^0) \to \overline{f}$, when the process of interest violates CP. However, there are two technical difficulties in the measurement of the time-dependent CP-violating effect. One is the determination of the B^0 flavor, whether B^0 or $\overline{B}^0$, at the beginning of the time-evolution. Another is the measurement of the proper time t, which is typically

$$\tau_{B^0} \sim \Delta m_d^{-1} \sim 1\,{\rm ps}\,,$$

and too short to be precisely measured as time.

We overcome these difficulties by 1) exploiting the coherence of $B^0\overline{B}^0$ produced in pairs, and 2) boosting the pairs with respect to the laboratory frame. At B-factories, B mesons are created in a coherent state via the decay chain of $e^+e^- \to \Upsilon(4S) \to B^0\overline{B}^0$. Among various cases of the subsequent decays, we choose the events where one of the B^0 (B_{CP}) decaying to the final state of interest, f, irrespective of the decay of another B^0 ($B_{\rm tag}$), as shown in Fig. 1.2. Since B^0 decays are dominated by the flavor-specific decay modes that distinguish the flavor charge of B^0, $B_{\rm tag}$ is most likely to decay to a flavor-specific final state. The coherence of the $B^0\overline{B}^0$ pair allows us to use this $B_{\rm tag}$ decay to *tag* the b-flavor of the B_{CP} at the beginning of the time-evolution; B_{CP} is *tagged* to be $\overline{B}^0$ (B^0) at the time of $B_{\rm tag}$ decay, t_2, in case of the $B_{\rm tag}$ decaying to the final state of a B^0 ($\overline{B}^0$) flavor eigenstate. We then push the start button of the stopwatch at the time t_2 and track the time-evolution of the B_{CP} beginning from $\overline{B}^0$ (B^0), until the time of its decay, t_1, when we push the stop button; the proper time difference $\Delta t \equiv t_1 - t_2$ is the quantity that parameterizes the time-evolution of B_{CP}. To measure the Δt

[1] Throughout this thesis, the inclusion of the charge conjugate decay mode is implied unless otherwise stated.

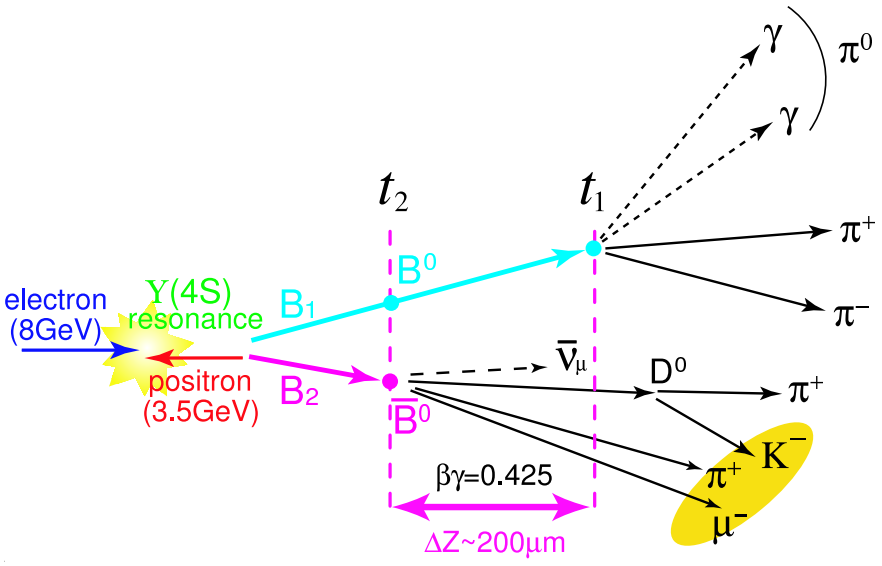


Figure 1.2: Illustration of event topology in energy-asymmetric B-factories. In this figure, B_1 and B_2 corresond to the B_{CP} and $B_{\rm tag}$, respectively, and $\pi^+\pi^-\pi^0$ is chosen to be the final state of interest, f, as an example.

of order $\sim 1\,$ps, B-factories produce $\Upsilon(4S)$ in a boosted state, where the boost factor $(\beta\gamma)_\Upsilon$ is 0.425 at Belle/KEKB. Because

$$m_{\Upsilon(4S)} \simeq 2m_B \ ,$$

two B's decaying from $\Upsilon(4S)$ are almost at rest in the center of mass system of the $\Upsilon(4S)$ and thus they have the same velocity in laboratory frame. This allows us to relate the position difference of decay vertices of two B's in the boost direction, Δz, with the proper time difference, Δt, by

$$\Delta t \simeq \frac{\Delta z}{c(\beta\gamma)_\Upsilon} \ ,$$

as illustrated in Fig. 1.2. From the boost factor of 0.425 and the typical Δt of $\sim 1\,$ps, the typical Δz is estimated to be $\sim 200\,\mu$m, which is possible to be measured in a reasonable precision.

Among various observables measured in B-factories, the CKM angle ϕ_2 has a desirable feature for the SM test; the precisions of the direct measurement and the indirect measurement are at the similar level and reasonably good, being better than 10%. Recent observation of the B_s-$\overline{B}_s$ mixing [15] improves the indirect measurement of ϕ_2, which also motivates the improvement in the direct measurement of ϕ_2. The elements of the CKM matrix, V_{td}, V_{tb}, V_{ud}, and V_{ub}, defines the angle ϕ_2 as

$$\phi_2 \equiv \arg\left(\frac{V_{td}V_{tb}^*}{-V_{ud}V_{ub}^*}\right) \ ,$$

and thus the direct measurement of ϕ_2 is possible via the decay processes that involve these CKM factors. As shown in Fig. 1.3, the factor $V_{td}V_{tb}^*$ arises from B^0-$\overline{B}^0$ mixing and $V_{ud}V_{ub}^*$ appears in the decay processes that involve $b \to u$ transition, such as $B^0 \to \pi^+\pi^-$, $B^0 \to \rho^+\rho^-$, or $B^0 \to \rho^\pm\pi^\mp$.

The difficulty in the measurement of ϕ_2 through the decays with $b \to u$ transition is the possible contribution from the gluonic penguin diagram of $b \to d$ transition (Fig. 1.4). The

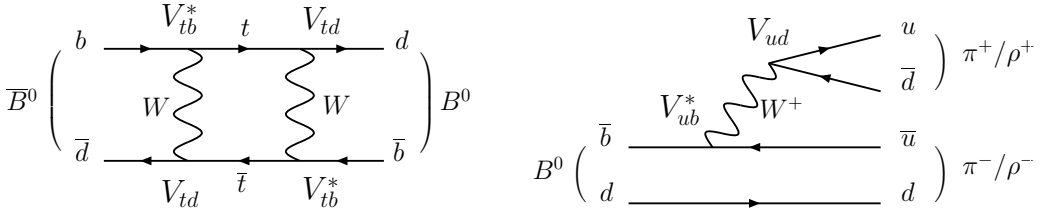


Figure 1.3: Feynman diagrams corresponding to the B^0-$\overline{B}^0$ mixing (left) and the B^0 decay via $b \to u$ transition (right).

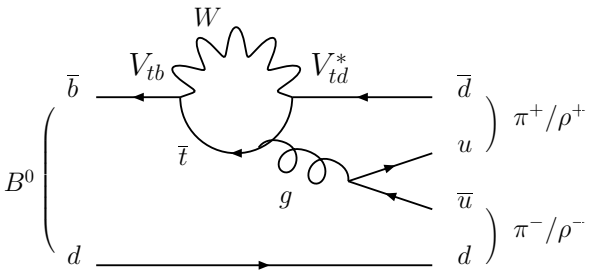


Figure 1.4: Gluonic penguin diagram of the B^0 decay via $b \to d$ transition.

contribution contaminates the measurement of ϕ_2, since the decays via this diagram yields the CKM factor $V_{tb}V_{td}^*$, which is different from that of the $b \to u$ transition. The isospin analysis [16] removes this contamination by incorporating the knowledge of other related processes; in the case of $B \to \pi\pi$, for example, it involves the decay modes $B^0 \to \pi^0\pi^0$ and $B^+ \to \pi^+\pi^0$ in addition to $B^0 \to \pi^+\pi^-$. Though it removes the contamination from penguin diagram in a model independent way, an additional four-fold discrete ambiguity arises. The discrete ambiguity is eight-fold in total, since there is an intrinsic two-fold ambiguity of the time-dependent CP-violation measurement in general[2]. Although the isospin analysis is also applicable to the case of $B \to \rho\pi$ [17, 18], the number of discrete ambiguities is much more than those of $B \to \pi\pi$ and $B \to \rho\rho$ due to the fact that the number of involved decay modes is larger; five decay modes, $B^0 \to \rho^+\pi^-$, $B^0 \to \rho^-\pi^+$, $B^0 \to \rho^0\pi^0$, $B^+ \to \rho^+\pi^0$, and $B^+ \to \rho^0\pi^+$, are involved in total. This makes it difficult to constrain ϕ_2 by the $B \to \rho\pi$ decay processes using the isospin analysis only.

In 1993, however, Snyder and Quinn pointed out that a time-dependent Dalitz plot analysis of the decay $B^0 \to \rho\pi \to \pi^+\pi^-\pi^0$ offers a unique way for the measurement of the angle ϕ_2 without discrete ambiguity (for ϕ_2 in the range between 0 and π), which cannot be obtained from the analyses of $B \to \pi\pi$ or $B \to \rho\rho$ [19]. The Dalitz plot analysis uses the isospin relation among the three B^0 decay modes, $B^0 \to \rho^+\pi^-$, $\rho^-\pi^+$, $\rho^0\pi^0$, and takes into account the possible contamination from the $b \to d$ penguin transitions. In addition, measurements of the related charged decay modes $B^+ \to \rho^+\pi^0$ and $B^+ \to \rho^0\pi^+$ provides further improvement of the ϕ_2 determination.

The uniqueness of the time-dependent Dalitz plot analysis is from the measurement related to the interference between $B^0 \to \rho^+\pi^-$, $B^0 \to \rho^-\pi^+$, and $B^0 \to \rho^0\pi^0$. Since the final state is $\pi^+\pi^-\pi^0$ for all of the $\rho^+\pi^-$, $\rho^-\pi^+$, and $\rho^0\pi^0$, their kinematic overlaps give rise to the

[2]The two-fold ambiguity comes from the fact that the observable is proportional to $\sin 2\phi$, where ϕ is the weak phase related to the decay mode of interest, and thus one cannot discriminate ϕ and $\pi - \phi$.

interferences among them. The measurement of the interferences and their time-dependent decay rates offer us the information on the relative phases among the complex decay amplitudes of the six related processes: $B^0 \to \rho^+\pi^-$, $\rho^-\pi^+$, $\rho^0\pi^0$ and their charge conjugates. The isospin relation connects the amplitudes with ϕ_2 as

$$e^{+i2\phi_2} = \frac{q}{p}\frac{A(\overline{B}^0 \to \rho^+\pi^-) + A(\overline{B}^0 \to \rho^-\pi^+) + 2A(\overline{B}^0 \to \rho^0\pi^0)}{A(B^0 \to \rho^+\pi^-) + A(B^0 \to \rho^-\pi^+) + 2A(B^0 \to \rho^0\pi^0)} ,$$

and thus the information on the relative phases plays an essential role in constraining ϕ_2.

In this thesis, we present a direct measurement of ϕ_2 using a unique method of time-dependent Dalitz plot analysis in the $B^0 \to \rho\pi \to \pi^+\pi^-\pi^0$ decay process. The analysis was performed on a $414\,\mathrm{fb}^{-1}$ data sample that contains $449 \times 10^6 B\bar{B}$ pairs collected on the $\Upsilon(4S)$ resonance. The data were taken at the KEKB collider [20] and collected with the Belle detector [21]. Based on the obtained result, we investigate the consistency between the direct and indirect measurements of ϕ_2.

Chapter 2

Phenomenology of CP Violation in $B^0 \to \pi^+\pi^-\pi^0$ Decay

1 Kobayashi-Maskawa Mechanism

In this section, we describe how the CP violation manifests itself in the quark sector. Denoting the u-type quarks (u, c, t) and d-type quarks (d, s, b) by U and D, respectively, the charged-current weak interaction Lagrangian in the quark sector, $\mathcal{L}_{\text{C.C.}}$, is written as

$$\mathcal{L}_{\text{C.C.}} = -\frac{g}{\sqrt{2}}\left[\left(\overline{U}_L\gamma^\mu \boldsymbol{V} D_L\right) W_\mu^+ + \left(\overline{D}_L\gamma^\mu \boldsymbol{V}^\dagger U_L\right) W_\mu^-\right] , \tag{2.1}$$

where $\boldsymbol{V}$ is the quark mixing matrix, or Cabibo-Kobayashi-Maskawa (CKM) matrix [2, 3]. It describes the relation between the mass eigenstates, (d, s, b), and the eigenstates of charged-current weak interaction, (d', s', b'):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \boldsymbol{V} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{2.2}$$

Since

$$(CP)\, W_\mu^+ \,(CP)^\dagger = -W_\mu^- \tag{2.3}$$

and

$$(CP)\, \overline{\psi}_1\gamma^\mu\psi_2 \,(CP)^\dagger = -\overline{\psi}_2\gamma^\mu\psi_1 , \tag{2.4}$$

where ψ_i is a fermion state in general,

$$(CP)\, \mathcal{L}_{\text{C.C.}} \,(CP)^\dagger = -\frac{g}{\sqrt{2}}\left[\left(\overline{U}_L\gamma^\mu \boldsymbol{V}^* D_L\right) W_\mu^+ + \left(\overline{D}_L\gamma^\mu (\boldsymbol{V}^*)^\dagger U_L\right) W_\mu^-\right] , \tag{2.5}$$

and thus $\mathcal{L}_{\text{C.C.}}$ is invariant under the CP transformation only when $\boldsymbol{V} = \boldsymbol{V}^*$, i.e., all the elements of $\boldsymbol{V}$ are real numbers.

To calculate the number of possible complex phase in the matrix, we start by counting the degrees of freedom in the $N \times N$ matrix $\boldsymbol{V}$. A complex $N \times N$ matrix in general has $2N^2$ degrees of freedom. The unitarity of the matrix $\boldsymbol{V}$ requires

$$\sum_j V_{ij} V_{kj}^* = \delta_{ik} , \tag{2.6}$$

corresponding to N^2 constraints. Having N dimensions in $\boldsymbol{V}$, total number of quarks is $2N$, N for each of u-types and d-types. In quantum mechanics, the phases of the quarks are meaningless and thus we can use the meaningless phases to remove complex phases in $\boldsymbol{V}$. Since the overall phase does not change $\boldsymbol{V}$, the redefinition of the quark phase removes $2N-1$ phases from $\boldsymbol{V}$. Thus, the degrees of freedom in $\boldsymbol{V}$ are

$$2N^2 - N^2 - (2N-1) = N^2 - 2N + 1 = (N-1)^2 \,. \tag{2.7}$$

This degrees of freedom can be classified into two categories:

- Real-valued degrees of freedom, corresponding to the rotation of N-element vectors, and
- Complex phases.

The former corresponds to the degrees of freedom in an $N \times N$ real-valued orthogonal matrix $\boldsymbol{R}$. The number of elements of $\boldsymbol{R}$ is N^2; the requirement of orthogonality yields

$$\sum_j R_{ij} R_{kj} = \delta_{ik} \quad (i < k) \,, \tag{2.8}$$

corresponding to $N(N+1)/2$ constraints; the degrees of freedom in $\boldsymbol{R}$ is thus

$$N^2 - \frac{1}{2}N(N+1) = \frac{1}{2}N(N-1) \,. \tag{2.9}$$

Consequently, the number of possible complex phase in $\boldsymbol{V}$ is

$$(N-1)^2 - \frac{1}{2}N(N-1) = \frac{1}{2}(N-1)(N-2) \,. \tag{2.10}$$

When $N = 3$, the CKM matrix $\boldsymbol{V}$ can have a single complex phase and CP can be violated in case the phase is non-zero.

Since degrees of freedom in $\boldsymbol{V}$ are four for $N = 3$, it is convenient to parameterize $\boldsymbol{V}$ with four parameters. The most popular parameterization is the Wolfenstein parameterization [22], which is a power-series expansion in the real parameter $\lambda \equiv \sin\theta_C$, where θ_C is Cabibbo angle [2]:

$$\boldsymbol{V} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \,. \tag{2.11}$$

Here, A, ρ, and η are real-valued parameters of order one.

Among the unitarity conditions of Eq. (2.6), those with $i \neq k$ describe triangles in complex

plane:

$$\underbrace{V_{ud}V_{cd}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{us}V_{cs}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{ub}V_{cb}^*}_{\mathcal{O}(\lambda^5)} = 0\,, \tag{2.12}$$

$$\underbrace{V_{ud}V_{us}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{cd}V_{cs}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{td}V_{ts}^*}_{\mathcal{O}(\lambda^5)} = 0\,, \tag{2.13}$$

$$\underbrace{V_{td}V_{cd}^*}_{\mathcal{O}(\lambda^4)} + \underbrace{V_{ts}V_{cs}^*}_{\mathcal{O}(\lambda^2)} + \underbrace{V_{tb}V_{cb}^*}_{\mathcal{O}(\lambda^2)} = 0\,, \tag{2.14}$$

$$\underbrace{V_{us}V_{ub}^*}_{\mathcal{O}(\lambda^4)} + \underbrace{V_{cs}V_{cb}^*}_{\mathcal{O}(\lambda^2)} + \underbrace{V_{ts}V_{tb}^*}_{\mathcal{O}(\lambda^2)} = 0\,, \tag{2.15}$$

$$\underbrace{V_{td}V_{ud}^*}_{\mathcal{O}(\lambda^3)} + \underbrace{V_{ts}V_{us}^*}_{\mathcal{O}(\lambda^3)} + \underbrace{V_{tb}V_{ub}^*}_{\mathcal{O}(\lambda^3)} = 0\,, \tag{2.16}$$

$$\underbrace{V_{ud}V_{ub}^*}_{\mathcal{O}(\lambda^3)} + \underbrace{V_{cd}V_{cb}^*}_{\mathcal{O}(\lambda^3)} + \underbrace{V_{td}V_{tb}^*}_{\mathcal{O}(\lambda^3)} = 0\,. \tag{2.17}$$

As illustrated in Fig. 2.1, the shapes of the six unitarity triangles can be characterized by the dependence of the side lengths on the order of λ. The first four triangles are squashed; the first and second (third and fourth) triangles have the sides that are order λ^4 (λ^2) smaller than the others. The fifth and sixth have three sides of the same order of λ^3, which implies the CP violating effect can be large when the physics process of interest is related to the these two triangles. In particular, Eq. (2.17) is related to B meson decays and called the "Unitarity Triangle," and thus large CP violation is expected in B decays. As shown in Fig. 2.2, the angles of the Unitarity Triangle are denoted by ϕ_1, ϕ_2, and ϕ_3[1]:

$$\phi_1 \equiv \pi - \arg\left(\frac{-V_{td}V_{tb}^*}{-V_{cd}V_{cb}^*}\right)\,, \tag{2.18}$$

$$\phi_2 \equiv \arg\left(\frac{V_{td}V_{tb}^*}{-V_{ud}V_{ub}^*}\right)\,, \tag{2.19}$$

$$\phi_3 \equiv \arg\left(\frac{V_{ud}V_{ub}^*}{-V_{cd}V_{cb}^*}\right)\,. \tag{2.20}$$

The measurement of these angles as well as the lengths of the sides of the Unitarity Triangle are the crucial test of the CKM picture of the CP violation.

2 Neutral Meson System

2-1 Time Evolution

Suppose we have a neutral meson P^0 and its CP conjugate $\overline{P}^0$, whose eigenstates are denoted by $\left|P^0\right\rangle$ and $\left|\overline{P}^0\right\rangle$, respectively. A state $\left|\Psi\right\rangle$ is described in terms of their linear combination

$$\left|\Psi\right\rangle = a\left|P^0\right\rangle + b\left|\overline{P}^0\right\rangle\,. \tag{2.21}$$

The Shrödinger equation for the state $\left|\Psi\right\rangle$ is

$$i\hbar\frac{\partial}{\partial t}\left|\Psi\right\rangle = \hat{H}\left|\Psi\right\rangle\,, \tag{2.22}$$

[1] Another naming convention, $\beta(=\phi_1)$, $\alpha(=\phi_2)$, and $\gamma(=\phi_3)$, is also used in the literature.

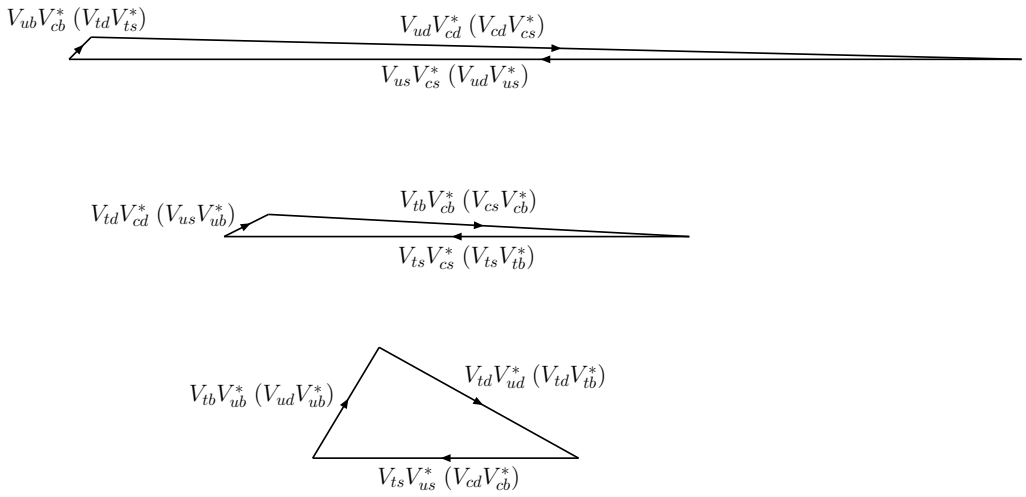


Figure 2.1: Schematic figures representing the unitarity relations. Top, middle, and bottom triangles correspond to equations (2.12)-(2.13), (2.14)-(2.15), and (2.16)-(2.17), respectively.

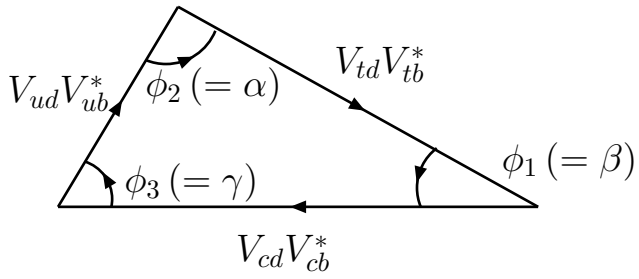


Figure 2.2: The "Unitarity Triangle."

where $\hat{H}$ is the Hamiltonian operator. The time evolution of $\left|\Psi\right\rangle$ is written in terms of those of the coefficients a and b

$$\left|\Psi(t)\right\rangle = a(t)\left|P^0\right\rangle + b(t)\left|\overline{P}^0\right\rangle \ , \tag{2.23}$$

and the Shrödinger equation becomes a differential equation of a vector $\vec{\Psi}^T(t) \equiv (a(t), b(t))$:

$$i\hbar\frac{\partial}{\partial t}\vec{\Psi}(t) = \mathscr{H}\vec{\Psi}(t) \ , \tag{2.24}$$

where

$$\mathscr{H} \equiv \begin{pmatrix} \left\langle P^0\right|\hat{H}\left|P^0\right\rangle & \left\langle P^0\right|\hat{H}\left|\overline{P}^0\right\rangle \\ \left\langle \overline{P}^0\right|\hat{H}\left|P^0\right\rangle & \left\langle \overline{P}^0\right|\hat{H}\left|\overline{P}^0\right\rangle \end{pmatrix} \ . \tag{2.25}$$

The 2×2 matrix $\mathscr{H}$ is given by

$$\mathscr{H} = \boldsymbol{M} - \frac{i}{2}\boldsymbol{\Gamma} \ , \tag{2.26}$$

where $\boldsymbol{M}$ and $\boldsymbol{\Gamma}$ are Hermitian matrices[2]. The CPT conservation requires the matrices to satisfy

$$M_{11} = M_{22} \ , \quad \text{and} \quad \Gamma_{11} = \Gamma_{22} \ . \tag{2.27}$$

In the followings, we assume the CPT conservation.

We define the eigenvectors of the matrix $\mathscr{H}$ as $(p,q)^T$ and $(p,-q)^T$, which satisfy

$$\frac{q}{p} = +\sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} \ , \tag{2.28}$$

$$|p|^2 + |q|^2 = 1 \ . \tag{2.29}$$

Here, the sign of the left hand side of Eq. (2.28) is just a convention[3]. The eigenstates of $\hat{H}$, or the mass eigenstates, are then

$$|P_L\rangle \equiv p\left|P^0\right\rangle + q\left|\overline{P}^0\right\rangle \ , \tag{2.30}$$

$$|P_H\rangle \equiv p\left|P^0\right\rangle - q\left|\overline{P}^0\right\rangle \ , \tag{2.31}$$

with corresponding eigenvalues of H_L and H_H. The eigenvalues are written with real-valued parameters M_L, M_H, Γ_L, and Γ_H as

$$H_L = M_L - \frac{i}{2}\Gamma_L \ , \tag{2.32}$$

$$H_H = M_H - \frac{i}{2}\Gamma_H \ , \tag{2.33}$$

which satisfy

$$M_L - \frac{i}{2}\Gamma_L = M_{11} - \frac{i}{2}\Gamma_{11} + \frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right) \ , \tag{2.34}$$

$$M_H - \frac{i}{2}\Gamma_H = M_{11} - \frac{i}{2}\Gamma_{11} - \frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right) \ . \tag{2.35}$$

[2]This means that the Hamiltonian matrix $\mathscr{H}$ is not Hermitian. This is because the state $\left|\Psi\right\rangle$ only includes P^0 and $\overline{P}^0$, and does not include the states of their decay products, leading to the fact that the probability $\left|\langle\Psi(t)|\Psi(t)\rangle\right|^2$ does not conserve once the decay occurs. The Hermitian property of Hamiltonian in general comes from the requirement of probability conservation. Thus, the matrix $\mathscr{H}$ here does not have to be Hermitian.

[3]This is related to which of (almost) CP even and odd states is heavier, or the sign of ΔM.

Equations (2.28), (2.34), and (2.35) give the following relations:

$$(\Delta M)^2 - \frac{1}{4}(\Delta\Gamma)^2 = 4|M_{12}|^2 - |\Gamma_{12}|^2 \,, \tag{2.36}$$

$$\Delta M \Delta\Gamma = 4\mathrm{Re}(M_{12}\Gamma_{12}^*) \,, \tag{2.37}$$

with

$$\Delta M \equiv M_H - M_L \,, \tag{2.38}$$

$$\Delta\Gamma \equiv \Gamma_H - \Gamma_L \,. \tag{2.39}$$

The time evolution of $|P_L\rangle$ and $|P_H\rangle$ are

$$|P_L(t)\rangle = e^{-i(M_L - i\Gamma_L/2)t}\,|P_L(0)\rangle \,, \tag{2.40}$$

$$|P_H(t)\rangle = e^{-i(M_H - i\Gamma_H/2)t}\,|P_H(0)\rangle \,. \tag{2.41}$$

The time evolution of the states $\left|P^0\right\rangle$ and $\left|\overline{P}^0\right\rangle$ is obtained from the equations (2.30), (2.31), (2.40), and (2.41), as

$$\begin{aligned}\left|P^0(t)\right\rangle &= \frac{1}{2p}\Big(|P_L(t)\rangle + |P_H(t)\rangle\Big) \\ &= f_+(t)\left|P^0(0)\right\rangle + \frac{q}{p}f_-(t)\left|\overline{P}^0(0)\right\rangle \,,\end{aligned} \tag{2.42}$$

$$\begin{aligned}\left|\overline{P}^0(t)\right\rangle &= \frac{1}{2q}\Big(|P_L(t)\rangle - |P_H(t)\rangle\Big) \\ &= \frac{p}{q}f_-(t)\left|P^0(0)\right\rangle + f_+(t)\left|\overline{P}^0(0)\right\rangle \,,\end{aligned} \tag{2.43}$$

where

$$f_\pm(t) \equiv \frac{1}{2}\Big[e^{-i(M_L - i\Gamma_L/2)t} \pm e^{-i(M_H - i\Gamma_H/2)t}\Big] \,. \tag{2.44}$$

2-2 CP violation

Provided there is a decay process of $P^0 \to f$, where f is a final state of the decay, its decay amplitude $A_f(t)$ is[4]

$$\begin{aligned}A_f(t) &= \langle f|\,\hat{H}\left|P^0(t)\right\rangle \\ &= \langle f|\,\hat{H}(t)\left|P^0(0)\right\rangle \\ &= \langle f|\,(CP)^\dagger(CP)\hat{H}(t)(CP)^\dagger(CP)\left|P^0(0)\right\rangle \\ &= \left\langle\overline{f}\right|(CP)\hat{H}(t)(CP)^\dagger\left|\overline{P}^0(0)\right\rangle \,.\end{aligned} \tag{2.47}$$

On the other hand, the amplitude $\overline{A}_{\overline{f}}(t)$ of the CP conjugate process, $\overline{P}^0 \to \overline{f}$, is

$$\begin{aligned}\overline{A}_{\overline{f}}(t) &= \left\langle\overline{f}\right|\hat{H}\left|\overline{P}^0(t)\right\rangle \,. \\ &= \left\langle\overline{f}\right|\hat{H}(t)\left|\overline{P}^0(0)\right\rangle \,.\end{aligned} \tag{2.48}$$

[4]Here, we take a convention of

$$CP\left|P^0(0)\right\rangle = \left|\overline{P}^0(0)\right\rangle \quad\text{and}\quad CP\left|\overline{P}^0(0)\right\rangle = \left|P^0(0)\right\rangle \,, \tag{2.45}$$

though in general the CP transformation can yield an unphysical phase ζ as

$$CP\left|P^0(0)\right\rangle = e^{+i\zeta}\left|\overline{P}^0(0)\right\rangle \quad\text{and}\quad CP\left|\overline{P}^0(0)\right\rangle = e^{-i\zeta}\left|P^0(0)\right\rangle \,. \tag{2.46}$$

Thus, comparison of the two amplitudes, $A_f(t)$ and $\overline{A}_{\overline{f}}(t)$, is sensitive to the difference between $\hat{H}(t)$ and $(CP)\hat{H}(t)(CP)^\dagger$; if we find difference between $A_f(t)$ and $\overline{A}_{\overline{f}}(t)$, $\hat{H}(t)$ violates CP symmetry.

With Eq. (2.42) and (2.43), the time-dependent amplitudes are calculated as

$$A_f(t) = f_+(t)A_f + \frac{q}{p} f_-(t)\overline{A}_{\overline{f}} \,, \tag{2.49}$$

$$\overline{A}_{\overline{f}}(t) = \frac{p}{q} f_-(t)A_f + f_+(t)\overline{A}_{\overline{f}} \,, \tag{2.50}$$

where

$$A_f \equiv \langle f| \hat{H} \left|P^0(0)\right\rangle \quad \text{and} \quad \overline{A}_{\overline{f}} \equiv \left\langle \overline{f}\right| \hat{H} \left|\overline{P}^0(0)\right\rangle \,. \tag{2.51}$$

The difference between the amplitudes is

$$\frac{\overline{A}_{\overline{f}}(t)}{A_f(t)} = \frac{\overline{A}_{\overline{f}}}{A_f} + \frac{\frac{p}{q}A_f f_-(t)}{f_+(t)A_f + \frac{q}{p}f_-(t)\overline{A}_{\overline{f}}} \left[1 - \left(\frac{q}{p}\frac{\overline{A}_{\overline{f}}}{A_f}\right)^2\right] \,. \tag{2.52}$$

Thus, the condition for the CP-violation is

$$\frac{q}{p}\frac{\overline{A}_{\overline{f}}}{A_f} \neq \pm 1 \,, \tag{2.53}$$

or

$$\left|\frac{\overline{A}_{\overline{f}}}{A_f}\right| \neq 1 \,. \tag{2.54}$$

Note that CP is conserved in the case of

$$\frac{q}{p}\frac{\overline{A}_{\overline{f}}}{A_f} = \pm 1 \,, \quad \left|\frac{\overline{A}_{\overline{f}}}{A_f}\right| = 1 \,, \quad \text{and} \quad \frac{\overline{A}_{\overline{f}}}{A_f} \neq 1 \,, \tag{2.55}$$

since the pure phase difference in the amplitudes is not an observable.

3 CP Violation in B Decays

3-1 Special Properties in B^0 System

In $B^0(\overline{B}^0)$[23] system, a negligible difference in the widths of the two mass eigenstates is expected[5]

$$\frac{\Delta\Gamma}{\Gamma} = \mathcal{O}(10^{-2}) \ll 1 \,, \tag{2.56}$$

where

$$\Gamma \equiv \frac{1}{2}\left(\Gamma_H + \Gamma_L\right) \,. \tag{2.57}$$

Ignoring the width difference, equation (2.44) becomes

$$f_+(t) = e^{-iMt}e^{-\Gamma t/2}\cos(\Delta m_d\, t/2) \,, \tag{2.58}$$

$$f_-(t) = e^{-iMt}e^{-\Gamma t/2}\, i\sin(\Delta m_d\, t/2) \,, \tag{2.59}$$

[5]This is because the difference $\Gamma_H - \Gamma_L$ is only produced by decay channels common to B^0 and $\overline{B}^0$. The small width difference is expected as a natural consequence of the fact that the decay width of B^0 and $\overline{B}^0$ are dominated by the decay channels to which B^0 or $\overline{B}^0$ only can decay exclusively.

where

$$M \equiv \frac{1}{2}\left(M_H + M_L\right) , \tag{2.60}$$

$$\Delta m_d \equiv \Delta M = M_H - M_L . \tag{2.61}$$

The mass difference measured by experiments is comparable with the decay width:

$$\frac{\Delta m_d}{\Gamma} \sim 1 . \tag{2.62}$$

With equations (2.56) and (2.62), we obtain

$$\frac{\Delta\Gamma}{\Delta m_d} = \mathcal{O}(10^{-2}) \ll 1 . \tag{2.63}$$

The mass hierarchy and GIM mechanism of the standard model gives the following relation:

$$\left|\frac{\Gamma_{12}}{M_{12}}\right| = \mathcal{O}\left(\frac{{m_b}^2}{{m_t}^2}\right) \sim 10^{-3} . \tag{2.64}$$

This can be roughly understood as follows. Since both M_{12} and Γ_{12} are related to the transition of $B^0 \leftrightarrow \overline{B}^0$, this can be understood by the box diagrams describing the B^0-$\overline{B}^0$ mixing. (Fig. 2.3) To M_{12}, virtual intermediate states contribute and all of the cases $q = u, c, t$ are allowed. Here, the contribution from $q = u, c$ are canceled out by GIM-cancellation and contribution of $q = t$ with large mass difference is dominant; and thus $|M_{12}| \sim {m_t}^2/{m_W}^2$. To Γ_{12}, on the other hand, the virtual intermediate state contribution is not allowed since $\mathbf{\Gamma}$ is related to decays of b quark, and thus $|\Gamma_{12}| \sim {m_b}^2/{m_W}^2$. The detailed discussion can be found elsewhere [24].

By the relations (2.56), (2.62), and (2.64), we obtain approximations of Eqs. (2.36) and (2.37):

$$\Delta m_d \simeq 2|M_{12}| , \tag{2.65}$$

$$\Delta\Gamma \simeq 2\mathrm{Re}(M_{12}\Gamma_{12}^*)/|M_{12}| , \tag{2.66}$$

$$\tag{2.67}$$

which lead to the approximated expression of Eq. (2.28):

$$\frac{q}{p} \simeq \frac{M_{12}^*}{|M_{12}|} . \tag{2.68}$$

As described above, M_{12} is related to the box diagrams of Fig. 2.3 with $q = t$. Thus, it is related to the CKM matrix elements as

$$M_{12} \propto (V_{tb}^* V_{td})^2 , \tag{2.69}$$

and thus

$$\frac{q}{p} \simeq \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} . \tag{2.70}$$

Note that following relation is satisfied

$$\left|\frac{p}{q}\right| \simeq 1 , \tag{2.71}$$

up to the precision of $\mathcal{O}(10^{-2})$.

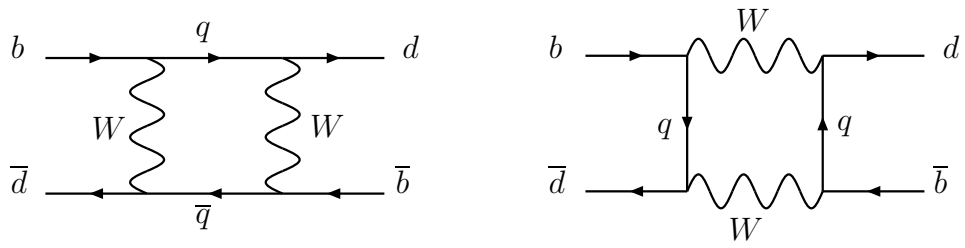


Figure 2.3: Box diagrams for the B^0-$\overline{B}^0$ mixing.

3-2 Time Evolution of a Coherent $B^0\overline{B}^0$ System

At B-factories, the B^0 mesons are produced from the decay of $\Upsilon(4S)$ in pairs. The two meson state in the center mass system of $\Upsilon(4S)$ is written as

$$|\Phi(\hat{n}_1,t_1;\hat{n}_2,t_2)\rangle = \frac{1}{\sqrt{1+a^2}}\Big[\big|B^0(\hat{n}_1,t_1)\big\rangle\big|\overline{B}^0(\hat{n}_2,t_2)\big\rangle + a\,\big|\overline{B}^0(\hat{n}_1,t_1)\big\rangle\big|B^0(\hat{n}_2,t_2)\big\rangle\Big]\delta(\hat{n}_1+\hat{n}_2)\ , \tag{2.72}$$

where $\hat{n}_{1,2}$ and $t_{1,2}$ are the flight direction and proper time of the each B meson, respectively, and $\delta(\hat{n}_1+\hat{n}_2)$ is from momentum conservation. Since $\Upsilon(4S)$ is a vector particle, the two meson state $|\Phi\rangle$ is required to be an eigenstate of P(parity) transformation with an eigenvalue of -1:

$$\begin{aligned} P\,|\Phi(\hat{n}_1,t_1;\hat{n}_2,t_2)\rangle &\equiv |\Phi(-\hat{n}_1,t_1;-\hat{n}_2,t_2)\rangle \\ &= |\Phi(\hat{n}_2,t_1;\hat{n}_1,t_2)\rangle \\ &= -\,|\Phi(\hat{n}_1,t_1;\hat{n}_2,t_2)\rangle\ , \end{aligned} \tag{2.73}$$

where $\delta(\hat{n}_1+\hat{n}_2)$ leads to the second equality. The solution for this equation is $a=-1$, and

$$|\Phi(\hat{n}_1,t_1;\hat{n}_2,t_2)\rangle = \frac{1}{\sqrt{2}}\Big[\big|B^0(\hat{n}_1,t_1)\big\rangle\big|\overline{B}^0(\hat{n}_2,t_2)\big\rangle - \big|\overline{B}^0(\hat{n}_1,t_1)\big\rangle\big|B^0(\hat{n}_2,t_2)\big\rangle\Big]\delta(\hat{n}_1+\hat{n}_2)\ . \tag{2.74}$$

From this equation and equations (2.42), (2.43), (2.58), and (2.59), the time evolution of the state $|\Phi\rangle$ is

$$\begin{aligned} |\Phi(\hat{n}_1,t_1;\hat{n}_2,t_2)\rangle =& \frac{1}{\sqrt{2}}e^{-i(M-i\Gamma/2)(t_1+t_2)}\Bigg[\cos\left(\frac{t_1-t_2}{2}\right)\Big(\big|{B^0}_1\big\rangle\big|{\overline{B}^0}_2\big\rangle - \big|{\overline{B}^0}_1\big\rangle\big|{B^0}_2\big\rangle\Big) \\ & - i\sin\left(\frac{t_1-t_2}{2}\right)\Big(\frac{p}{q}\big|{B^0}_1\big\rangle\big|{B^0}_2\big\rangle - \frac{q}{p}\big|{\overline{B}^0}_1\big\rangle\big|{\overline{B}^0}_2\big\rangle\Big)\Bigg]\ , \end{aligned} \tag{2.75}$$

with

$$\big|{B^0}_1\big\rangle \equiv \big|B^0(\hat{n}_1,0)\big\rangle\ , \tag{2.76}$$

$$\big|{B^0}_2\big\rangle \equiv \big|B^0(\hat{n}_2,0)\big\rangle\ , \tag{2.77}$$

$$\big|{\overline{B}^0}_1\big\rangle \equiv \big|\overline{B}^0(\hat{n}_1,0)\big\rangle\ , \tag{2.78}$$

$$\big|{\overline{B}^0}_2\big\rangle \equiv \big|\overline{B}^0(\hat{n}_2,0)\big\rangle\ . \tag{2.79}$$

The decay amplitude of $\Upsilon(4S) \to B^0\overline{B}^0 \to f_1 f_2$ is then

$$\begin{aligned} A(t_1,t_2) &\equiv \langle f_1 f_2 | \hat{H}_I | \Phi(\hat{n}_1,t_1;\hat{n}_2,t_2)\rangle \\ &= \frac{1}{\sqrt{2}} e^{-i(M-i\Gamma/2)(t_1+t_2)} \Bigg[\cos\left(\frac{t_1-t_2}{2}\right) \left(A_{f_1}\overline{A}_{f_2} - \overline{A}_{f_1}A_{f_2}\right) \\ &\qquad - i\sin\left(\frac{t_1-t_2}{2}\right)\left(\frac{p}{q}A_{f_1}A_{f_2} - \frac{q}{p}\overline{A}_{f_1}\overline{A}_{f_2}\right)\Bigg] , \end{aligned} \tag{2.80}$$

where B_1 decays to a final state f_1 at time t_1 and B_2 decays to f_2 at time t_2, respectively. Here, $\hat{H}_I$ is the interaction Hamiltonian and

$$A_{f_1} \equiv \langle f_1 | \hat{H}_I | B^0 \rangle , \tag{2.81}$$

$$A_{f_2} \equiv \langle f_2 | \hat{H}_I | B^0 \rangle , \tag{2.82}$$

$$\overline{A}_{f_1} \equiv \langle f_1 | \hat{H}_I | \overline{B}^0 \rangle , \tag{2.83}$$

$$\overline{A}_{f_2} \equiv \langle f_2 | \hat{H}_I | \overline{B}^0 \rangle . \tag{2.84}$$

In measuring CP violation, we choose f_1 to be a decay mode of interest and f_2 to be a flavor eigenstate $f_{\rm tag}$, i.e., $A_{f_2} = 0$ or $\overline{A}_{f_2} = 0$. In this case, the time dependent decay width is

$$\begin{aligned} d\Gamma/dt_1dt_2 &\propto |A(t_1,t_2)|^2 \\ &= \begin{cases} e^{-\Gamma(t_1+t_2)}|A_{f_+}|^2\Big[|A_{f_1}|^2 + |\overline{A}_{f_1}|^2 & \\ \quad -(|A_{f_1}|^2 - |\overline{A}_{f_1}|^2)\cos(\Delta m_d\Delta t) + 2{\rm Im}\left(\frac{q}{p}\overline{A}_{f_1}A^*_{f_1}\right)\sin(\Delta m_d\Delta t)\Big] & (f_{\rm tag} = f_+) \\ e^{-\Gamma(t_1+t_2)}|\overline{A}_{f_-}|^2\Big[|A_{f_1}|^2 + |\overline{A}_{f_1}|^2 & \\ \quad +(|A_{f_1}|^2 - |\overline{A}_{f_1}|^2)\cos(\Delta m_d\Delta t) - 2{\rm Im}\left(\frac{q}{p}\overline{A}_{f_1}A^*_{f_1}\right)\sin(\Delta m_d\Delta t)\Big] & (f_{\rm tag} = f_-) \end{cases} \end{aligned} \tag{2.85}$$

where f_+ is a final state to which B^0 only can decay, f_- is its CP conjugate, i.e., $f_- = \overline{f}_+$, $\Delta t \equiv t_1 - t_2$, and we assume $|p/q| = 1$. Since no (or very small) direct CP violation is expected in the decay modes like $B^0 \to f_+$[6], we assume

$$|A_{f_+}| = |\overline{A}_{f_-}| . \tag{2.86}$$

Thus, we can simplify equation (2.85) as

$$\begin{aligned} d\Gamma/dt_1dt_2 \propto & e^{-\Gamma(t_1+t_2)}\Big[|A_{f_1}|^2 + |\overline{A}_{f_1}|^2 \\ & - q_{\rm tag}\cdot(|A_{f_1}|^2 - |\overline{A}_{f_1}|^2)\cos(\Delta m_d\Delta t) + q_{\rm tag}\cdot 2{\rm Im}\left(\frac{q}{p}\overline{A}_{f_1}A^*_{f_1}\right)\sin(\Delta m_d\Delta t)\Big] , \end{aligned} \tag{2.87}$$

where $q_{\rm tag}$ denotes the flavor of decaying B_2, or $B_{\rm tag}$ hereafter, and $q_{\rm tag} = +1(-1)$ for $f_{\rm tag} = f_{+(-)}$. At B-factories, we can only measure Δt and cannot measure $t_1 + t_2$. By integrating equation (2.87) over the unmeasurable direction of $t_1 + t_2$, we obtain

$$\begin{aligned} d\Gamma/d\Delta t &\equiv \int_{|\Delta t|}^{+\infty} d(t_1+t_2)\,\frac{d\Gamma}{dt_1dt_2} \\ &\propto e^{-\Gamma|\Delta t|}\Big[|A_{f_1}|^2 + |\overline{A}_{f_1}|^2 \\ &\quad - q_{\rm tag}\cdot(|A_{f_1}|^2 - |\overline{A}_{f_1}|^2)\cos(\Delta m_d\Delta t) + q_{\rm tag}\cdot 2{\rm Im}\left(\frac{q}{p}\overline{A}_{f_1}A^*_{f_1}\right)\sin(\Delta m_d\Delta t)\Big] . \end{aligned} \tag{2.88}$$

[6] Violation of this assumption is part of so-called tag-side interference (TSI), of which we take account in the systematic error.

This is what we observe as the time evolution of the two B meson system at B-factories.

3-3 B Meson Decaying to a CP Eigenstate

When the decay mode of interest f_1 is a CP eigenstate, we call it f_{CP} and it satisfies

$$\left|\overline{f}_{CP}\right\rangle \equiv CP\left|f_{CP}\right\rangle = \eta_{f_{CP}}\left|f_{CP}\right\rangle \ , \tag{2.89}$$

where $\eta_{f_{CP}}$ is the CP eigenvalue of the state f_{CP} and $\eta_{f_{CP}} = \pm 1$. This relation leads to

$$\overline{A}_{\overline{f}_{CP}} \equiv \left\langle \overline{f}_{CP}\right| \hat{H} \left|\overline{B}^0\right\rangle = \eta_{f_{CP}} \left\langle f_{CP}\right| \hat{H} \left|\overline{B}^0\right\rangle = \eta_{f_{CP}} \overline{A}_{f_{CP}} \ , \tag{2.90}$$

and Eq. (2.88) is rewritten in terms of $A_{f_{CP}}$ and $\overline{A}_{\overline{f}_{CP}}$ as

$$\begin{aligned} d\Gamma/d\Delta t \propto e^{-\Gamma|\Delta t|}\Big[& |A_{f_{CP}}|^2 + |\overline{A}_{\overline{f}_{CP}}|^2 \\ & - q_{\rm tag}\cdot(|A_{f_{CP}}|^2 - |\overline{A}_{\overline{f}_{CP}}|^2)\cos(\Delta m_d \Delta t) \\ & + q_{\rm tag}\cdot\eta_{f_{CP}} 2{\rm Im}\left(\frac{q}{p}\overline{A}_{\overline{f}_{CP}} A^*_{f_{CP}}\right)\sin(\Delta m_d\Delta t)\Big] \ . \end{aligned} \tag{2.91}$$

Note that the additional factor $\eta_{f_{CP}}$ is introduced when we transform $\overline{A}_{f_{CP}}$ into $\overline{A}_{\overline{f}_{CP}}$. For convenience, we take $|A_{f_{CP}}|^2 + |\overline{A}_{\overline{f}_{CP}}|^2$ as overall normalization and rewrite the equation (2.91) as

$$d\Gamma/d\Delta t \propto e^{-\Gamma|\Delta t|}\Big[1 + q_{\rm tag}\cdot\mathcal{A}_{f_{CP}}\cos(\Delta m_d\Delta t) + q_{\rm tag}\cdot\mathcal{S}_{f_{CP}}\sin(\Delta m_d\Delta t)\Big] \ , \tag{2.92}$$

with

$$\mathcal{A}_{f_{CP}} \equiv \frac{|\overline{A}_{\overline{f}_{CP}}|^2 - |A_{f_{CP}}|^2}{|\overline{A}_{\overline{f}_{CP}}|^2 + |A_{f_{CP}}|^2} = \frac{|\lambda_{f_{CP}}|^2 - 1}{|\lambda_{f_{CP}}|^2 + 1} \ , \tag{2.93}$$

$$\mathcal{S}_{f_{CP}} \equiv \eta_{f_{CP}} \frac{2\,{\rm Im}\left(\frac{q}{p}\overline{A}_{\overline{f}_{CP}} A^*_{f_{CP}}\right)}{|\overline{A}_{\overline{f}_{CP}}|^2 + |A_{f_{CP}}|^2} = \eta_{f_{CP}} \frac{2\,{\rm Im}\lambda_{f_{CP}}}{|\lambda_{f_{CP}}|^2 + 1} \ . \tag{2.94}$$

Here, $\lambda_{f_{CP}}$ is defined as

$$\lambda_{f_{CP}} \equiv \frac{q}{p}\frac{\overline{A}_{\overline{f}_{CP}}}{A_{f_{CP}}} \ . \tag{2.95}$$

Note that $\eta_{f_{CP}}$ could be outside of $\mathcal{S}_{f_{CP}}$ in some definitions.

When the CP is violated in the process, we see the difference between $B^0 \to f_{CP}$ and $\overline{B}^0 \to f_{CP}$, corresponding to non-zero $\mathcal{A}_{f_{CP}}$ or $\mathcal{S}_{f_{CP}}$. This only happens when $\lambda_{f_{CP}} \neq \pm 1$, which is consistent with the condition of Eq. (2.53). $\mathcal{A}_{f_{CP}}$ is non-zero only when $|\lambda_{f_{CP}}| \neq 1$ (direct CP violation).

3-4 B Meson Decaying to a non-CP Eigenstate

When the decay mode of interest f_1 is not a CP eigenstate, there is no relation corresponding to Eq. (2.90). Thus, the coefficients of $\cos(\Delta m_d\Delta t)$ and $\sin(\Delta m_d\Delta t)$ could be non-zero even if the CP is conserved. We can still obtain information related to the CP violating effect by measuring the time-dependences of both $B^0(\overline{B}^0) \to f_1$ and $B^0(\overline{B}^0) \to \overline{f}_1$ processes. This can

be demonstrated as follows. We rewrite equation (2.88) as

$$d\Gamma/d\Delta t \propto \begin{cases} (1+\mathcal{A}^{CP})\, e^{-\Gamma|\Delta t|}\Big[1 - q_{\rm tag}\cdot(\mathcal{C}+\Delta\mathcal{C})\cos(\Delta m_d \Delta t) & \\ \qquad\qquad +q_{\rm tag}\cdot(\mathcal{S}+\Delta\mathcal{S})\sin(\Delta m_d\Delta t)\Big] & (\text{for } B_1 \to f_1) \\ (1-\mathcal{A}^{CP})\, e^{-\Gamma|\Delta t|}\Big[1 - q_{\rm tag}\cdot(\mathcal{C}-\Delta\mathcal{C})\cos(\Delta m_d \Delta t) & \\ \qquad\qquad +q_{\rm tag}\cdot(\mathcal{S}-\Delta\mathcal{S})\sin(\Delta m_d\Delta t)\Big] & (\text{for } B_1 \to \overline{f}_1) \end{cases} \tag{2.96}$$

where

$$1+\mathcal{A}^{CP} \propto |A_{f_1}|^2 + |\overline{A}_{f_1}|^2 \,, \tag{2.97}$$

$$1-\mathcal{A}^{CP} \propto |A_{\overline{f}_1}|^2 + |\overline{A}_{\overline{f}_1}|^2 \,, \tag{2.98}$$

$$\mathcal{C}+\Delta\mathcal{C} \equiv \frac{|A_{f_1}|^2 - |\overline{A}_{f_1}|^2}{|A_{f_1}|^2 + |\overline{A}_{f_1}|^2} \,, \tag{2.99}$$

$$\mathcal{C}-\Delta\mathcal{C} \equiv \frac{|A_{\overline{f}_1}|^2 - |\overline{A}_{\overline{f}_1}|^2}{|A_{\overline{f}_1}|^2 + |\overline{A}_{\overline{f}_1}|^2} \,, \tag{2.100}$$

$$\mathcal{S}+\Delta\mathcal{S} \equiv \frac{2\,\mathrm{Im}\left(\frac{q}{p}\overline{A}_{f_1}A^*_{f_1}\right)}{|A_{f_1}|^2 + |\overline{A}_{f_1}|^2} \,, \tag{2.101}$$

$$\mathcal{S}-\Delta\mathcal{S} \equiv \frac{2\,\mathrm{Im}\left(\frac{q}{p}\overline{A}_{\overline{f}_1}A^*_{\overline{f}_1}\right)}{|A_{\overline{f}_1}|^2 + |\overline{A}_{\overline{f}_1}|^2} \,, \tag{2.102}$$

and thus

$$\mathcal{A}^{CP} = \frac{\big(1-|\lambda_{f_1}|^2\big) - |\xi|^2\big(1-|\lambda_{\overline{f}_1}|^2\big)}{\big(1+|\lambda_{f_1}|^2\big) + |\xi|^2\big(1+|\lambda_{\overline{f}_1}|^2\big)} \,, \tag{2.103}$$

$$\mathcal{C} = \frac{1-|\lambda_{f_1}|^2|\lambda_{\overline{f}_1}|^2}{\big(1+|\lambda_{\overline{f}_1}|^2|\xi|^2\big)\big(1+|\lambda_{f_1}|^2/|\xi|^2\big)} \,, \tag{2.104}$$

$$\Delta\mathcal{C} = \frac{|\lambda_{f_1}|^2/|\xi|^2 - |\lambda_{\overline{f}_1}|^2|\xi|^2}{\big(1+|\lambda_{\overline{f}_1}|^2|\xi|^2\big)\big(1+|\lambda_{f_1}|^2/|\xi|^2\big)} \,, \tag{2.105}$$

$$\mathcal{S} = \frac{\mathrm{Im}\Big[(1-\lambda^*_{f_1}\lambda^*_{\overline{f}_1})(\lambda_{f_1}/\xi + \lambda_{\overline{f}_1}\xi)\Big]}{\big(1+|\lambda_{\overline{f}_1}|^2|\xi|^2\big)\big(1+|\lambda_{f_1}|^2/|\xi|^2\big)} \,, \tag{2.106}$$

$$\Delta\mathcal{S} = \frac{\mathrm{Im}\Big[(1+\lambda^*_{f_1}\lambda^*_{\overline{f}_1})(\lambda_{f_1}/\xi - \lambda_{\overline{f}_1}\xi)\Big]}{\big(1+|\lambda_{\overline{f}_1}|^2|\xi|^2\big)\big(1+|\lambda_{f_1}|^2/|\xi|^2\big)} \,, \tag{2.107}$$

with

$$\lambda_{f_1} \equiv \frac{q}{p}\frac{\overline{A}_{\overline{f}_1}}{A_{f_1}} \,, \tag{2.108}$$

$$\lambda_{\overline{f}_1} \equiv \frac{q}{p}\frac{\overline{A}_{f_1}}{A_{\overline{f}_1}} \,, \tag{2.109}$$

$$\xi \equiv \frac{A_{\overline{f}_1}}{A_{f_1}} \,. \tag{2.110}$$

When CP is violated, we see the difference between $B^0 \to f_1$ and $\overline{B}^0 \to \overline{f}_1$ (or, $B^0 \to \overline{f}_1$ and $\overline{B}^0 \to f_1$), corresponding to non-zero $\mathcal{A}^{CP}$, $\mathcal{C}$, or $\mathcal{S}$[7]. This only happens when $\lambda_{f_1} \neq \pm 1$ or $\lambda_{\overline{f}_1} \neq \pm 1$, which are consistent with Eq. (2.53). In particular, $\mathcal{A}^{CP}$ and $\mathcal{C}$ can only be non-zero when $|\lambda_{f_1}| \neq 1$ or $|\lambda_{\overline{f}_1}| \neq 1$ (direct CP violation). On the other hand, $\Delta\mathcal{C}$ and $\Delta\mathcal{S}$ can be non-zero even if CP is conserved.

It is convenient for the interpretation of the direct CP violation to define $\mathcal{A}^{\pm\mp}$ as [25]

$$\mathcal{A}^{+-} \equiv \frac{|\lambda_{f_1}|^2 - 1}{|\lambda_{f_1}|^2 + 1} = -\frac{\mathcal{A}^{CP} + \mathcal{C} + \mathcal{A}^{CP}\Delta\mathcal{C}}{1 + \Delta\mathcal{C} + \mathcal{A}^{CP}\,\mathcal{C}}\,, \quad \text{and} \tag{2.114}$$

$$\mathcal{A}^{-+} \equiv \frac{|\lambda_{\overline{f}_1}|^2 - 1}{|\lambda_{\overline{f}_1}|^2 + 1} = +\frac{\mathcal{A}^{CP} - \mathcal{C} - \mathcal{A}^{CP}\Delta\mathcal{C}}{1 - \Delta\mathcal{C} - \mathcal{A}^{CP}\,\mathcal{C}}\,, \tag{2.115}$$

which can be interpreted as

$$\mathcal{A}^{+-} = \frac{\Gamma(\overline{B}^0 \to \overline{f}_1) - \Gamma(B^0 \to f_1)}{\Gamma(\overline{B}^0 \to \overline{f}_1) + \Gamma(B^0 \to f_1)}\,, \quad \text{and} \tag{2.116}$$

$$\mathcal{A}^{-+} = \frac{\Gamma(\overline{B}^0 \to f_1) - \Gamma(B^0 \to \overline{f}_1)}{\Gamma(\overline{B}^0 \to f_1) + \Gamma(B^0 \to \overline{f}_1)}\,. \tag{2.117}$$

4 CP Asymmetry in $B^0 \to \pi^+\pi^-\pi^0$ Decay Process

4-1 Basic Properties

Since the decay process $B^0 \to \pi^+\pi^-\pi^0$ is dominated by $B^0 \to \rho^+\pi^-$ and $\rho^-\pi^+$, the final state is not a CP eigenstate. Although there exists $B^0 \to \rho^0\pi^0$ process, where the final state is a CP eigenstate, the contribution from this process is small as described later. Thus, one has to measure both of $B^0 \to \rho^+\pi^-$ and $B^0 \to \rho^-\pi^+$, at least, to observe the CP violating effect. The processes $B^0 \to \rho^+\pi^-$, $\rho^-\pi^+$, $\rho^0\pi^0$ are described by tree, color suppressed tree, and penguin diagrams, up to $\mathcal{O}(\lambda)$ precision [26]. Figure 2.4 shows the diagrams. Since the process $B^0 \to \rho^0\pi^0$ does not have the contribution from the color allowed tree diagram, a small branching fraction is expected. The amplitudes corresponding to the diagrams are

$$A_T(B^0 \to \rho^+\pi^-) = T^{+-}\,V_{ub}^* V_{ud}\,, \tag{2.118}$$

$$A_P(B^0 \to \rho^+\pi^-) = P^{+-}\,V_{tb}^* V_{td}\,, \tag{2.119}$$

$$A_T(B^0 \to \rho^-\pi^+) = T^{-+}\,V_{ub}^* V_{ud}\,, \tag{2.120}$$

$$A_P(B^0 \to \rho^-\pi^+) = P^{-+}\,V_{tb}^* V_{td}\,, \tag{2.121}$$

$$A_C(B^0 \to \rho^0\pi^0) = C^{00}\,V_{ub}^* V_{ud}\,, \tag{2.122}$$

$$A_P(B^0 \to \rho^0\pi^0) = P^{00}\,V_{tb}^* V_{td}\,, \tag{2.123}$$

[7] Suppose the case to compare the difference between $B^0 \to f_1$ and $\overline{B}^0 \to \overline{f}_1$, for example. In this case, the time-dependent decay widths for them are

$$d\Gamma/d\Delta t \propto (1 + \mathcal{A}^{CP})\,e^{-\Gamma|\Delta t|}\Big[1 - (-1)\cdot(\mathcal{C} + \Delta\mathcal{C})\cos(\Delta m_d \Delta t) + (-1)\cdot(\mathcal{S} + \Delta\mathcal{S})\sin(\Delta m_d \Delta t)\Big]\,, \tag{2.111}$$

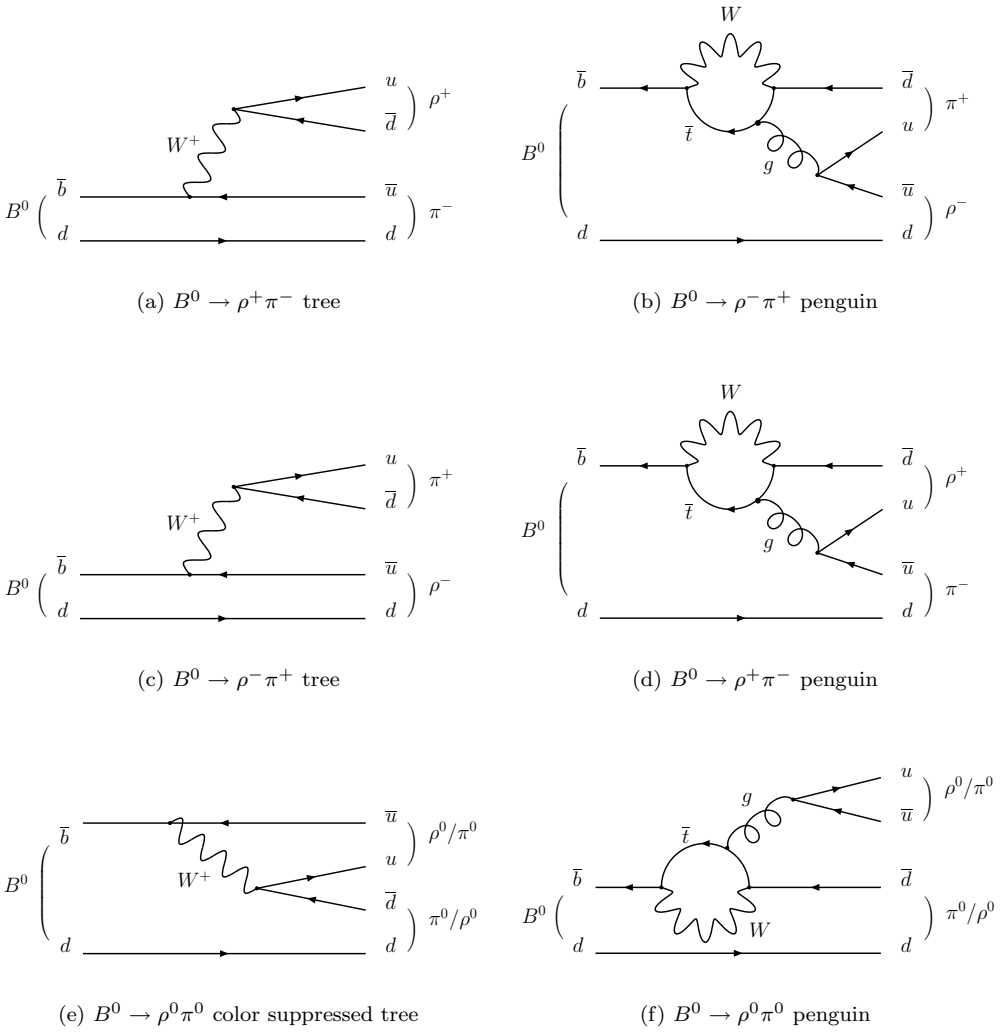
and

$$d\Gamma/d\Delta t \propto (1 - \mathcal{A}^{CP})\,e^{-\Gamma|\Delta t|}\Big[1 - (+1)\cdot(\mathcal{C} - \Delta\mathcal{C})\cos(\Delta m_d \Delta t) + (+1)\cdot(\mathcal{S} - \Delta\mathcal{S})\sin(\Delta m_d \Delta t)\Big]\,, \tag{2.112}$$

respectively. Thus, their difference is proportional to

$$2\mathcal{A}^{CP} e^{-\Gamma|\Delta t|}\Big[1 - \Delta\mathcal{C}\cos(\Delta m_d \Delta t) + \Delta\mathcal{S}\sin(\Delta m_d \Delta t)\Big] + 2e^{-\Gamma|\Delta t|}\Big[\mathcal{C}\cos(\Delta m_d \Delta t) - \mathcal{S}\sin(\Delta m_d \Delta t)\Big]\,, \tag{2.113}$$

and can be non-zero only when either of $\mathcal{A}^{CP}$, $\mathcal{C}$, or $\mathcal{S}$ is non-zero.



(a) $B^0 \to \rho^+\pi^-$ tree

(b) $B^0 \to \rho^-\pi^+$ penguin

(c) $B^0 \to \rho^-\pi^+$ tree

(d) $B^0 \to \rho^+\pi^-$ penguin

(e) $B^0 \to \rho^0\pi^0$ color suppressed tree

(f) $B^0 \to \rho^0\pi^0$ penguin

Figure 2.4: Feynman diagrams related to $B^0 \to (\rho\pi)^0$ processes.

and

$$
\begin{aligned}
A_T(\overline{B}^0 \to \rho^+\pi^-) &= T^{-+}\, V_{ub}V_{ud}^* \,, && (2.124)\\
A_P(\overline{B}^0 \to \rho^+\pi^-) &= P^{-+}\, V_{tb}V_{td}^* \,, && (2.125)\\
A_T(\overline{B}^0 \to \rho^-\pi^+) &= T^{+-}\, V_{ub}V_{ud}^* \,, && (2.126)\\
A_P(\overline{B}^0 \to \rho^-\pi^+) &= P^{+-}\, V_{tb}V_{td}^* \,, && (2.127)\\
A_C(\overline{B}^0 \to \rho^0\pi^0) &= C^{00}\, V_{ub}V_{ud}^* \,, && (2.128)\\
A_P(\overline{B}^0 \to \rho^0\pi^0) &= P^{00}\, V_{tb}V_{td}^* \,, && (2.129)
\end{aligned}
$$

where the subscripts T, C, and P in the left hand sides denote tree, color suppressed tree, and penguin transitions, respectively; and $T^{\pm\mp}$, C^{00}, and $P^{\pm\mp(00)}$, denotes the amplitudes excluding the explicitly written factors coming from CKM matrix. Ignoring the difference of strong interaction between $B^0 \to \rho^+\pi^-$ and $B^0 \to \rho^-\pi^+$, i.e., the factor T^{-+}/T^{+-}, and contributions from penguin transitions, the coefficients of $\cos(\Delta m_d \Delta t)$ and $\sin(\Delta m_d \Delta t)$ of the processes $B^0 \to \rho^\pm\pi^\mp$ calculated from equations (2.99)-(2.102) are

$$
\mathcal{C} \pm \Delta\mathcal{C} = \frac{|\lambda'_{\rho^\pm\pi^\mp}|^2 - 1}{|\lambda'_{\rho^\pm\pi^\mp}|^2 + 1} \sim 0 \,, \tag{2.130}
$$

$$
\mathcal{S} \pm \Delta\mathcal{S} = \frac{2\mathrm{Im}\lambda'_{\rho^\pm\pi^\mp}}{|\lambda'_{\rho^\pm\pi^\mp}|^2 + 1} \sim \sin(2\phi_2) \,, \tag{2.131}
$$

with

$$
\begin{aligned}
\lambda'_{\rho^\pm\pi^\mp} &\equiv \frac{q}{p}\frac{A(\overline{B}^0 \to \rho^\pm\pi^\mp)}{A(B^0 \to \rho^\pm\pi^\mp)} \\
&\sim \frac{V_{tb}^*V_{td}}{V_{tb}V_{td}^*}\frac{V_{ub}V_{ud}^*}{V_{ub}^*V_{ud}} \\
&= e^{2i\phi_2} \,,
\end{aligned} \tag{2.132}
$$

where the last equality comes from the definition of ϕ_2 in Eq. (2.19). Thus, ϕ_2 is the CKM angle that is related to the process $B^0 \to \rho^\pm\pi^\mp$.

However, we cannot measure the ϕ_2 directly from $\mathcal{S} \pm \Delta\mathcal{S}$ in practice. This is because the two effects ignored above are actually not negligible; the effects from 1) the factor T^{-+}/T^{+-} originating from strong interaction difference between $B^0 \to \rho^+\pi^-$ and $B^0 \to \rho^-\pi^+$, and 2) contributions from penguin transitions are to be taken into account.

As for the former, the sizes and phases of the amplitudes $T^{\pm\mp}$, C^{00}, and $P^{\pm\mp(00)}$ are dependent on the decay modes, where the difference originates from the different contribution of strong interaction related to each mode. Due to the CP-conserving property of strong interaction, the amplitudes are common between $B^0 \to \rho^+\pi^-$ and $\overline{B}^0 \to \rho^-\pi^+$, as the equations above show. In the time-dependent analysis, however, what we measure is the interference between $B^0 \to \rho^+\pi^-$ and $\overline{B}^0 \to \rho^+\pi^-$ (or, between $B^0 \to \rho^-\pi^+$ and $\overline{B}^0 \to \rho^-\pi^+$), and the contributions from strong interaction can be different, i.e., the factor T^{-+}/T^{+-} is different from unity in general. With this factor taken into account and assuming tree transitions only, $\lambda'_{\rho^\pm\pi^\mp}$ in equation (2.132) becomes

$$
\lambda'_{\rho^\pm\pi^\mp} \sim \frac{q}{p}\frac{A_T(\overline{B}^0 \to \rho^\pm\pi^\mp)}{A_T(B^0 \to \rho^\pm\pi^\mp)} = r_T^\pm e^{i(2\phi_2 \pm \delta_T)} \,, \tag{2.133}
$$

with

$$r_T^+ \equiv \frac{1}{r_T^-} \equiv \left|\frac{T^{-+}}{T^{+-}}\right| , \quad \text{and} \tag{2.134}$$

$$\delta_T \equiv \arg\left(\frac{T^{-+}}{T^{+-}}\right) . \tag{2.135}$$

Consequently, the coefficients $\mathcal{C} \pm \Delta\mathcal{C}$ and $\mathcal{S} \pm \Delta\mathcal{S}$ are

$$\mathcal{C} \pm \Delta\mathcal{C} \sim \frac{r_T^{\pm\,2} - 1}{r_T^{\pm\,2} + 1} , \quad \text{and} \tag{2.136}$$

$$\mathcal{S} \pm \Delta\mathcal{S} \sim \sqrt{1 - (\mathcal{C} \pm \Delta\mathcal{C})^2}\, \sin(2\phi_2 \pm \delta_T) . \tag{2.137}$$

Thus, it is still possible to measure ϕ_2 using all four observables of $\mathcal{C} \pm \Delta\mathcal{C}$ and $\mathcal{S} \pm \Delta\mathcal{S}$.

The second problem, the contribution from the penguin diagrams, is more serious. This is because the phases from both weak and strong interactions are different for the contributions from tree and penguin diagrams. The parameter corresponding to $\lambda'_{\rho^\pm\pi^\mp}$ for the penguin contribution is

$$\begin{aligned} \lambda'^{P}_{\rho^\pm\pi^\mp} &\equiv \frac{q}{p}\frac{A_P(\overline{B}^0 \to \rho^\pm\pi^\mp)}{A_P(B^0 \to \rho^\pm\pi^\mp)} \\ &= \frac{P^{-+}}{P^{+-}}\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}\frac{V_{tb} V_{td}^*}{V_{tb}^* V_{td}} \\ &= r_P^\pm\, e^{\pm i\delta_P} , \end{aligned} \tag{2.138}$$

with

$$r_P^+ \equiv \frac{1}{r_P^-} \equiv \left|\frac{P^{-+}}{P^{+-}}\right| , \quad \text{and} \tag{2.139}$$

$$\delta_P \equiv \arg\left(\frac{P^{-+}}{P^{+-}}\right) . \tag{2.140}$$

Here, the corresponding CP-violating weak phase is 0 and the strong interaction factor is P^{-+}/P^{+-}, both being different from those of tree diagram contributions. In addition to r_P^+ and δ_P, the complex ratio P^{+-}/T^{+-} is also a newly introduced unknown parameter. With the penguin diagram contribution included, we have model parameters corresponding to seven degrees of freedom: ϕ_2, T^{-+}/T^{+-}, P^{-+}/P^{+-}, and P^{+-}/T^{+-}. The number of model parameters is now larger than that of observables and one cannot constrain ϕ_2 without further assumptions or additional observables.

There are two approaches proposed to overcome this issue: the isospin (pentagon) analysis and the time-dependent Dalitz plot analysis. We describe them in the following subsections. Note that they are not exclusive with each other; to constrain ϕ_2 at the end of this thesis, we use both of them simultaneously to make full use of the information we have.

4-2 Isospin (Pentagon) Analysis

There are four isospin relations between the amplitudes related to $B^0 \to \rho^\pm\pi^\mp$, $\rho^0\pi^0$, $B^+ \to \rho^+\pi^0$, and $\rho^0\pi^+$ [17, 18]:

$$\begin{aligned} A^+ + A^- + 2A^0 &= \widetilde{A}^+ + \widetilde{A}^- + 2\widetilde{A}^0 \\ &= \sqrt{2}(A^{+0} + A^{0+}) = \sqrt{2}(\widetilde{A}^{-0} + \widetilde{A}^{0-}) , \end{aligned} \tag{2.141}$$

$$A^{+0} - A^{0+} - \sqrt{2}(A^+ - A^-) = \widetilde{A}^{-0} - \widetilde{A}^{0-} - \sqrt{2}(\widetilde{A}^- - \widetilde{A}^+) \,, \tag{2.142}$$

where

$$\begin{aligned} A^+ &\equiv A(B^0 \to \rho^+\pi^-) \,, & (2.143)\\ A^- &\equiv A(B^0 \to \rho^-\pi^+) \,, & (2.144)\\ A^0 &\equiv A(B^0 \to \rho^0\pi^0) \,, & (2.145)\\ A^{+0} &\equiv A(B^+ \to \rho^+\pi^0) \,, & (2.146)\\ A^{0+} &\equiv A(B^+ \to \rho^0\pi^+) \,, & (2.147) \end{aligned}$$

$$\begin{aligned} \overline{A}^+ &\equiv \frac{p}{q} A(\overline{B}^0 \to \rho^+\pi^-) \,, & (2.148)\\ \overline{A}^- &\equiv \frac{p}{q} A(\overline{B}^0 \to \rho^-\pi^+) \,, & (2.149)\\ \overline{A}^0 &\equiv \frac{p}{q} A(\overline{B}^0 \to \rho^0\pi^0) \,, & (2.150)\\ A^{-0} &\equiv \frac{p}{q} A(B^- \to \rho^-\pi^0) \,, & (2.151)\\ A^{0-} &\equiv \frac{p}{q} A(B^- \to \rho^0\pi^-) \,, & (2.152) \end{aligned}$$

and

$$\widetilde{A}^\kappa \equiv e^{-2i\phi_2}\overline{A}^\kappa \,, \quad \widetilde{A}^{-0} \equiv e^{-2i\phi_2}A^{-0} \,, \quad \widetilde{A}^{0-} \equiv e^{-2i\phi_2}A^{0-} \,. \tag{2.153}$$

Now the related model parameters correspond to

$$\begin{aligned} &(10 \text{ amplitudes} = 20 \text{ d.o.f.}) + \phi_2 \\ &\quad - (1 \text{ global phase}) - (4 \text{ isospin relations} = 8 \text{ d.o.f.}) = 12 \text{ d.o.f.} \,, \end{aligned} \tag{2.154}$$

while the number of observables is

$$\begin{aligned} &(5 \text{ branching fractions}) + (2 \text{ charge asymmetries of } B^+ \text{ decay modes}) \\ &\quad + (2 \times 3 \text{ time-dependent coefficients of } B^0 \text{ decay modes}) = 13 \,. \end{aligned} \tag{2.155}$$

Thus, with more observables than model parameters, we can solve this problem and constrain ϕ_2.

In practice, however, it is difficult to constrain ϕ_2 only with this method. This is due to the size of branching fraction of $B^0 \to \rho^0\pi^0$, which is not so small that we can ignore the contribution from the process but not so large that we can measure the CP violation parameters of the process with a good precision. Without the good measurements of the CP violation parameters of the process, we only have 11 observables, effectively, and we cannot constrain ϕ_2 very well.

4-3 Time-Dependent Dalitz Plot Analysis

Snyder and Quinn pointed out that the time-dependent Dalitz plot analysis can be a powerful method to measure ϕ_2 using the $B^0 \to (\rho\pi)^0 \to \pi^+\pi^-\pi^0$ decay process [19]. The essence of this method is to measure the time-dependent CP violation parameters of the interference between the three decay modes: $B^0 \to \rho^+\pi^-$, $B^0 \to \rho^-\pi^+$, and $B^0 \to \rho^0\pi^0$. This increases the number of observables and makes it possible to constrain ϕ_2 combined with the isospin relation of the first equality in equation (2.141), even with the presence of the penguin contributions. Further, the information from the interference can solve the discrete ambiguity between ϕ_2 and $\pi - \phi_2$, which cannot be solved by the measurements using other modes like $B \to \pi\pi$ and $B \to \rho\rho$.

In the time-dependent Dalitz plot analysis, the amplitudes A_{f_1} and $\overline{A}_{f_1}$ in equation (2.88) have the Dalitz plot dependence, which are denoted by $A_{3\pi}(s_+, s_-)$ and $\overline{A}_{3\pi}(s_+, s_-)$. Here, we define the Lorentz-invariant Dalitz plot variables as

$$\begin{aligned} s_+ &\equiv (p_+ + p_0)^2 \,, \\ s_- &\equiv (p_- + p_0)^2 \,, \quad \text{and} \\ s_0 &\equiv (p_+ + p_-)^2 \,, \end{aligned} \tag{2.156}$$

where p_+, p_-, and p_0 are the four-momenta of π^+, π^-, and π^0 in the $B^0 \to \pi^+\pi^-\pi^0$ decay, respectively. Among the Dalitz plot variables, the following relation holds

$$s_+ + s_- + s_0 = {m_{B^0}}^2 + 2{m_{\pi^+}}^2 + {m_{\pi^0}}^2 \,. \tag{2.157}$$

Ignoring the $B^0 \to \pi^+\pi^-\pi^0$ contributions from the processes other than $B^0 \to (\rho\pi)^0 \to \pi^+\pi^-\pi^0$, the Dalitz plot amplitudes $\overset{(-)}{A}_{3\pi}(s_+, s_-)$ can be written as

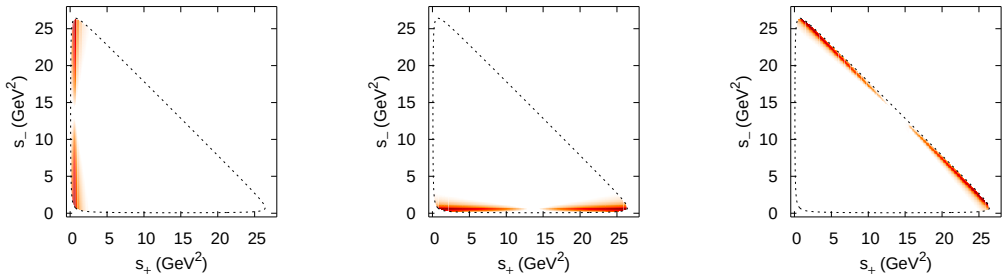
$$A_{3\pi}(s_+, s_-) = f_+(s_+, s_-)A^+ + f_-(s_+, s_-)A^- + f_0(s_+, s_-)A^0 \,, \quad \text{and} \tag{2.158}$$

$$\frac{q}{p}\overline{A}_{3\pi}(s_+, s_-) = \overline{f}_+(s_+, s_-)\overline{A}^+ + \overline{f}_-(s_+, s_-)\overline{A}^- + \overline{f}_0(s_+, s_-)\overline{A}^0 \,, \tag{2.159}$$

where functions $\overset{(-)}{f}_\kappa(s_+, s_-)$ (with ρ charge $\kappa = +, -, 0$) incorporate the kinematic and dynamical properties of B^0 decay into a vector ρ^κ and a pseudoscalar $\pi^{\overline{\kappa}}$, with $(\overline{+}, \overline{-}, \overline{0}) = (-, +, 0)$, corresponding to the mass and helicity distributions of the ρ^κ. Figure 2.5 schematically shows the $f_\kappa(s_+, s_-)$ in the Dalitz plot. We will discuss in Sec. 1 of chapter 6 the detail of the functions $\overset{(-)}{f}_\kappa(s_+, s_-)$. As described there, we assume the relation

$$\overline{f}_\kappa(s_+, s_-) = f_\kappa(s_+, s_-) \tag{2.160}$$

in our nominal fit. The definition of the complex coefficients A^κ and $\overline{A}^\kappa$ here are consistent with equations (2.143)-(2.145) and (2.148)-(2.150).



(a) $|f_+(s_+, s_-)|$ (b) $|f_-(s_+, s_-)|$ (c) $|f_0(s_+, s_-)|$

Figure 2.5: Schematic figures of the $f_\kappa(s_+, s_-)$. Dotted lines show the kinematic boundary.

With equations (2.158), (2.159), and (2.160), we rewrite Eq. (2.88) as

$$\begin{aligned} \frac{d\Gamma}{d\Delta t\, ds_+ ds_-} \propto e^{-\Gamma|\Delta t|}\Big[& |A_{3\pi}(s_+, s_-)|^2 + |\overline{A}_{3\pi}(s_+, s_-)|^2 \\ & - q_{\rm tag} \cdot (|A_{3\pi}(s_+, s_-)|^2 - |\overline{A}_{3\pi}(s_+, s_-)|^2)\cos(\Delta m_d \Delta t) \\ & + q_{\rm tag} \cdot 2{\rm Im}\left(\frac{q}{p}\overline{A}_{3\pi}(s_+, s_-)A_{3\pi}(s_+, s_-)^*\right)\sin(\Delta m_d \Delta t)\Big] \,. \end{aligned} \tag{2.161}$$

where

$$|A_{3\pi}(s_+,s_-)|^2 \pm |\overline{A}_{3\pi}(s_+,s_-)|^2 = \sum_{\kappa\in\{+,-,0\}} |f_\kappa|^2(|A^\kappa|^2 \pm |\overline{A}^\kappa|^2) + 2\sum_{\kappa<\sigma\in\{+,-,0\}} \left(\mathrm{Re}[f_\kappa f_\sigma^*]\,\mathrm{Re}[A^\kappa A^{\sigma *} \pm \overline{A}^\kappa \overline{A}^{\sigma *}] - \mathrm{Im}[f_\kappa f_\sigma^*]\,\mathrm{Im}[A^\kappa A^{\sigma *} \pm \overline{A}^\kappa \overline{A}^{\sigma *}]\right), \quad \text{and} \tag{2.162}$$

$$\mathrm{Im}\left(\frac{q}{p}\overline{A}_{3\pi}(s_+,s_-)A_{3\pi}(s_+,s_-)^*\right) = \sum_{\kappa\in\{+,-,0\}} |f_\kappa|^2 \mathrm{Im}[A^\kappa A^{\kappa *}] + 2\sum_{\kappa<\sigma\in\{+,-,0\}} \left(\mathrm{Re}[f_\kappa f_\sigma^*]\,\mathrm{Im}[\overline{A}^\kappa A^{\sigma *} + \overline{A}^\sigma A^{\kappa *}] + \mathrm{Im}[f_\kappa f_\sigma^*]\,\mathrm{Re}[\overline{A}^\kappa A^{\sigma *} - \overline{A}^\sigma A^{\kappa *}]\right). \tag{2.163}$$

Here, the $|f_\kappa|^2$ and Re(Im)$[f_\kappa f_\sigma^*]$ are nine linear-independent functions in the Dalitz plot. Since there are three types of distribution in Δt direction, $e^{-|\Delta t|/\tau_{B^0}}$, $e^{-|\Delta t|/\tau_{B^0}}\cos(\Delta m_d \Delta t)$, and $e^{-|\Delta t|/\tau_{B^0}}\sin(\Delta m_d \Delta t)$, we have 27 linear-independent functions in Δt-Dalitz plot space in total. Exploiting the information of both Δt and Dalitz plot, therefore, we can measure all of 27 coefficients of the independent functions, which are sufficient to determine all of the amplitudes A^κ and $\overline{A}^\kappa$ except for overall phase and normalization, in principle.

Equations (2.141) and (2.153) are derived from the fact that each combination-sum of the amplitudes connected by equality is written only by tree diagram contributions; we can write the combination-sums as

$$A^+ + A^- + 2A^0 = T^{\mathrm{All}} e^{-i\phi_2}, \quad \text{and} \tag{2.164}$$

$$\overline{A}^+ + \overline{A}^- + 2\overline{A}^0 = T^{\mathrm{All}} e^{+i\phi_2}. \tag{2.165}$$

Thus, with all the amplitudes A^κ and $\overline{A}^\kappa$ determined by the time-dependent Dalitz plot analysis, we can determine the ϕ_2 using the relation of

$$e^{2i\phi_2} = \frac{\overline{A}^+ + \overline{A}^- + 2\overline{A}^0}{A^+ + A^- + 2A^0}, \tag{2.166}$$

where the ratio in the right hand side can be determined without the unmeasured overall phase and normalization. Note that here we have no discrete ambiguity related to ϕ_2, which is an advantage of this method compared to the isospin (pentagon) analysis of $B \to \rho\pi$ and the analysis with other decay processes, $B \to \pi\pi$ and $B \to \rho\rho$.

Another advantage of this method, compared to the isospin (pentagon) analysis described in the previous section, is the large number of observables. The model parameters here are 9, calculated as

$$\begin{aligned}&(6 \text{ complex amplitudes} = 12 \text{ d.o.f.}) + \phi_2 \\ &\quad - (1 \text{ overall phase}) - (1 \text{ overall normalization}) - (1 \text{ isospin relation} = 2 \text{ d.o.f.}) = 9,\end{aligned} \tag{2.167}$$

while the number of observables are 26:

$$(27 \text{ coefficients}) - (1 \text{ overall normalization}) = 26. \tag{2.168}$$

The number of observables are far larger than that of model parameters. This allows us to determine the ϕ_2 even in the situation where some of the observables cannot be measured with good precisions due to the small branching fraction of $B^0 \to \rho^0\pi^0$.

Chapter 3

Experimental Apparatus

In this chapter, we describe the experimental apparatus of the KEK B factory, which consists of the KEKB accelerator and the Belle detector. The experiment is located at the High Energy Accelerator Research Organization (KEK) in Japan.

1 KEKB Accelerator

KEKB [20] is a two-ring energy-asymmetric e^+e^- collider and aims to produce copious B and anti-B mesons as in a factory. Figure 3.1 shows a schematic layout of KEKB. It consists of two 3 km-long storage rings, an 8 GeV electron ring (HER) and a 3.5 GeV positron ring (LER), and an injection linear accelerator. The two rings cross at one point, called the interaction point (IP), where electrons and positrons collide with a finite crossing angle of ± 11 mrad. The Belle detector surrounds IP to catch particles produced by the collisions. The center-of-mass energy is 10.58 GeV, which corresponds to the mass of the $\Upsilon(4S)$ resonance. Due to the energy asymmetry, the $\Upsilon(4S)$ are produced with a Lorentz boost of $(\beta\gamma)_\Upsilon = 0.425$. On average, the separation of the decay vertices of two B mesons is approximately $\langle \Delta z \rangle = c\tau_B(\beta\gamma)_\Upsilon \sim 200\,\mu$m.

The design luminosity of KEKB is $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Now the accelerator operates routinely with a peak luminosity of $1.5 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which is the world record as of Oct. 2006. In early 2004, a new method of operation of KEKB was successfully introduced. It is called "continuous injection mode" and removes the dead time of the ordinary injection method. Without the continuous injection, data taking has to stop every hour to replenish the beams. Now the KEKB can produce more than $1\,\mathrm{fb}^{-1}$ per day. The best records up to Oct. 2006 are $1.6517 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for the peak luminosity and $1.2315\,\mathrm{fb}^{-1}$ per day.

2 Belle Detector

Belle detector [21] is a general-purpose 4-π detector surrounding IP. It consists of a barrel, forward, and backward components. Figure 3.2 shows the configuration of the Belle detector.

Precision tracking and vertex measurements are provided by a central drift chamber (CDC) and a silicon vertex detector (SVD). The identifications of charged pions and kaons are based on the information from three subdetectors: the dE/dx measurement by CDC, a set of time-of-flight counters (TOF), and a set of aerogel Čerenkov counters (ACC). Electromagnetic particles are detected in an array of CsI(Tl) crystal calorimeters (ECL). The electron identification is based on a combination of the dE/dx measurements by CDC, the response of ACC, and the information of position, shape, and energy of the electromagnetic shower in ECL. The above detectors are located inside a superconducting solenoid of 1.7 m radius that maintains 1.5 T magnetic field. The outermost detector subsystem is a K_L and muon detector (KLM). A pair

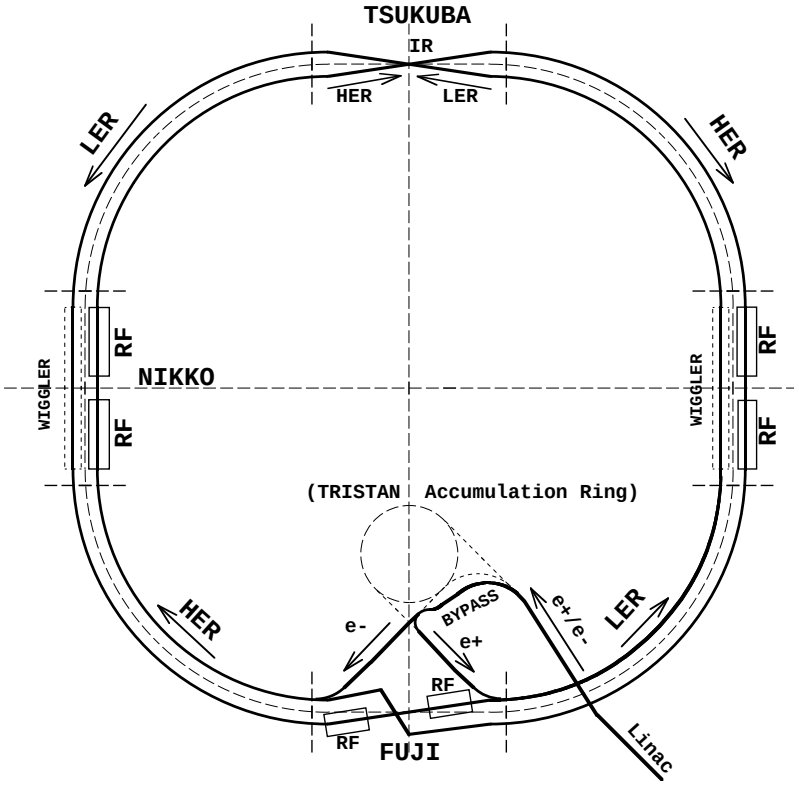


Figure 3.1: Schematic view of the layout of KEKB.

of BGO crystal array called extreme forward calorimeter (EFC), which is placed on the surface of the QCS cryostats, provides coverage at small angle uncovered by the other detectors.

Two inner detector configurations are used. A 3-layer SVD with a 2 cm radius beam-pipe is used until the summer of 2003. A data sample corresponding to the integrated luminosity of 140 fb^{-1} (DS-I) is collected with this configuration. In the summer of 2003, a 4-layer SVD, a 1.5 cm radius beam-pipe, and a small-cell inner drift chamber are installed. A data sample corresponding to the integrated luminosity of 274 fb^{-1} (DS-II) is collected with this configuration. Performance parameters of the detectors are summarized in Table 3.1.

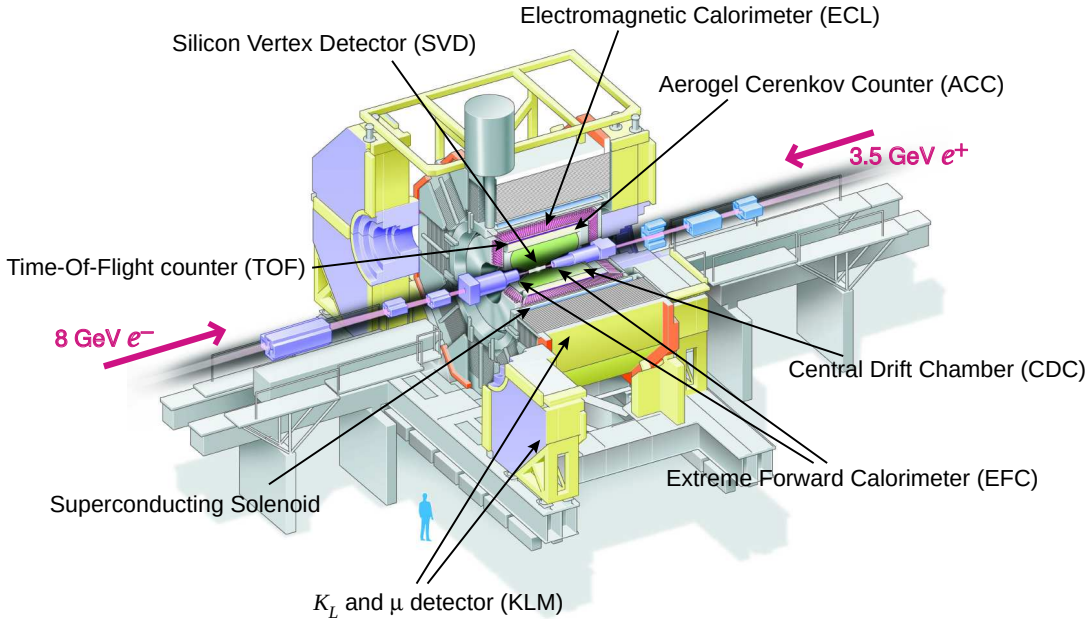


Figure 3.2: Overview of the Belle detector.

2-1 Silicon Vertex Detector (SVD)

It is crucially important for the time-evolution study to measure the difference between the flight lengths of the two B mesons in the z direction, where z is defined as the opposite of the positron beam direction. SVD [27] provides the essential information for the precise reconstruction of the decay vertices close to IP. Since the average separation of two B-decay vertices is $\sim 200\,\mu$m, the required z resolution is $\sim 200\,\mu$m. In addition, the vertex detector can be useful for identifying and measuring the decay vertices of D and τ particles.

Since most particles of interest in Belle have momenta of 1 GeV/c or less, the vertex resolution is dominated by the multiple-Coulomb scattering. This imposes strict constraints on the design of the detector. In particular, the innermost layer of the vertex detector must be placed as close to IP as possible, the support structure must be light in weight but rigid, and the readout electronics are to be placed outside the tracking volume. The design must also withstand large beam backgrounds. With the high luminosity operation of KEKB, the radiation dose to the detector is measured to be 1 kRad/day as of October 2006. Radiation doses of this level could both degrade the noise performance of the electronics and induce leakage currents in the silicon detectors.

Figure 3.3 shows the side and end views of the SVD for DS-I (SVD1). SVD1 consists of three concentric cylindrical layers arranged in a barrel and covers the angle range $23^\circ < \theta < 139^\circ$ (θ

Table 3.1: Performance parameters for the Belle detector.

Detector	Type	Configuration	Readout	Performance
Beam-pipe				
DS-I	Beryllium double wall	Cylindrical, $r = 20$ mm, 0.5/2.5/0.5 mm = Be/He/Be, He gas cooled		
DS-II	Beryllium double wall	Cylindrical, $r = 15$ mm, 0.5/2.5/0.5 mm = Be/Paraffin/Be, Paraffin liquid cooled		
SVD				
DS-I	Double-sided Si strip	3-layers: 3/8/14 ladders, Strip pitch: $25(\phi)/42(z)$ μm	ϕ : 40.96 k, z : 40.96 k	$\sigma(\Delta z) \sim 80\,\mu$m for $B^0 \to (\rho\pi)^0$
DS-II	Double-sided Si strip	4-layers: 6/12/18/18 ladders, Strip pitch: $50(\phi)/75(z)$ μm(layer 1-3), $65(\phi)/73(z)$ μm(layer 4)	ϕ : 55.296 k, z : 55.296 k	$\sigma(\Delta z) \sim 70\,\mu$m for $B^0 \to (\rho\pi)^0$
CDC				
DS-I	Small cell drift chamber	Anode: 50 layers, Cathode: 3 layers, $r = 8.3$-87.4 cm, $-78.72 \le z \le 158.77$ cm	A: 8.4 k, C: 1.8 k	$\sigma_{r\phi} = 130\,\mu m$ $\sigma_z = 200 \sim 1400\,\mu m$ $\sigma_{p_t}/p_t = 0.3\%\sqrt{{p_t}^2 + 1}$ $\sigma_{dE/dx} = 8\%$
DS-II	Small cell drift chamber	Anode: 49 layers, No cathode layer, $r = 10.4$-87.4 cm, $-78.72 \le z \le 158.77$ cm	A: 8.5 k	
ACC	Silica aerogel	960 barrel / 228 end-cap FM-PMT readout		$N_{p.e.} \ge 6$, K/π separation: $1.2 < p < 3.5$ GeV
TOF/TSC	Scintillator	128/64 ϕ segmentation, $r = 120$ cm, 3-m long	128×2 / 64	$\sigma_t = 100$ ps K/π separation: up to 1.2 GeV/c
ECL	CsI (Towered-structure)	Barrel: $r = 125$-162 cm, Endcap: $z = -102$ cm and $+196$ cm	6624, 1152(FW), 960(BW)	$\sigma_E/E = 1.3\%/\sqrt{E}$ $\sigma_{\rm pos} = 0.5\,{\rm cm}/\sqrt{E}$ (E in GeV)
KLM	Resistive plate counters	14 layers (5 cm Fe + 4 cm gap) 2 RPCs in each gap	θ: 16 k, ϕ: 16 k	$\Delta\phi = \Delta\theta = 30$ mr for $K_L \sim 1\%$ hadron fake
EFC	BGO	Photodiode readout, Segmentation: 32 in ϕ, 5 in θ	100×2	Energy resolution (rms): 7.3% at 8 GeV, 5.8% at 8 GeV
Magnet	Super-conducting	Inner radius = 170 cm		$B = 1.5$ T

being the polar angle from the z axis), which corresponds to 86% of the full solid angle. The three layers at radii of 30.0 mm, 45.5 mm, and 60.5 mm surround the beam pipe, a double-wall beryllium cylinder of 2.3 cm radius and 1 mm thickness. There are 8/10/14 ladders along ϕ in layers 1/2/3, respectively, where ϕ is the azimuthal angle around the z axis. Each ladder consists of double-sided silicon strip detectors (DSSDs) reinforced by boron-nitride support ribs.

The DSSD's fabricated by Hamamatsu Photonics (HPK)[1], originally designed for the DELPHI microvertex detector, are used for the SVD1. Each DSSD consists of 1280 sense strips and 640 readout pads on each side, where the pitch size of the strips is 42 μm (25 μm) in z-side (ϕ-side). The overall DSSD size is $57.5 \times 33.5\,\mathrm{mm}^2$ with 300 μm thickness. In total 102 DSSD's are used and the number of readout channels is 81,920. For the z-coordinate measurement, the n-side strips are used and a double-metal structure running parallel to z is employed to route the signals from orthogonal z-sense strips to the ends of the detector. Adjacent strips are connected to a single readout trace on the second metal layer which gives an effective strip pitch of 84 μm. A p-stop structure is employed to isolate the z-sense strips. A relatively large thermal noise ($\sim 600e^-$) is observed due to the common-p-stop design. On the ϕ side, every other sense-strip is only connected to a readout channel. Charge collected by the floating strips, the strips unconnected to readout channels, in between is read from adjacent strips by means of capacitive charge division.

The readout chain for DSSD's is based on the VA1 integrated circuit [28, 29]. The VA1 chip is a 128-channel CMOS integrated circuit produced by IDEAS[2]. It is specially designed for the readout of silicon vertex detectors and other small-signal devices that require low-noise preamplifier. VA1 has excellent noise performance and reasonably good radiation tolerance of 200 krad (1 Mrad) for VA1 fabricated in the Austrian Micro Systems (AMS) 1.2 μm (0.8 μm) process [30], where the VA1 with 0.8 μm process is used from the summer of 2000.

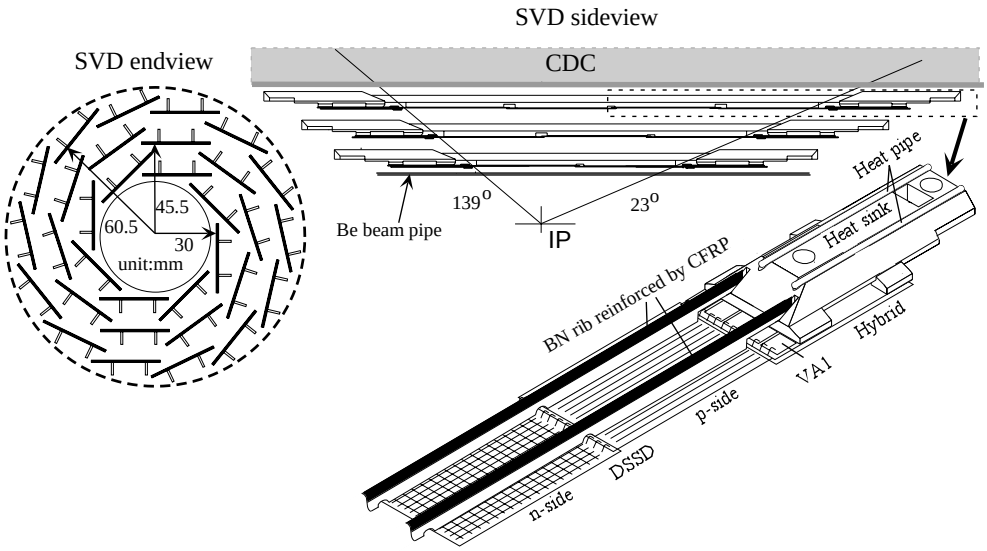


Figure 3.3: Detector configuration of SVD1.

In the summer of 2003, a new vertex detector, SVD2, was installed [31]. Figure 3.4 schematically shows the configuration of SVD2. It has four detector layers; there are 6/12/18/18 ladders

[1] http://www.hamamatsu.com/
[2] http://www.ideas.no/

at the radii of 20.0/43.5/70.0/88.0 mm for the 1/2/3/4 layers, respectively. The polar angle acceptance is expanded to $17^\circ < \theta < 150^\circ$, which is the same as CDC and corresponds to the 92% of the full solid angle. The beam-pipe, surrounded by the ladders, has 1.5 cm radius and double-wall structure, where inner and outer layers have 0.6 mm and 0.35 mm thickness, respectively, with cooling liquid[3] circulated through the space of 0.5 mm between them.

The DSSD's are fabricated by HPK. The size of the DSSD for the layer 1-3 (4) is $28.4 \times 79.2\,\mathrm{mm}^2$ ($34.9 \times 76.4\,\mathrm{mm}^2$) with $300\,\mu$m thickness. Each DSSD has 1024 and 512 sense strips in $z(p)$-side and $\phi(n)$-side[4], respectively. For the inner three layers, the z-strip (ϕ-strip) pitch is $75\,\mu$m ($50\,\mu$m). For the fourth layer, the z-strip (ϕ-strip) pitch is $73\,\mu$m ($65\,\mu$m). Every ϕ-strip for each DSSD is read out, while every other strip is read out in z-side. In total, 246 DSSD's are used and the number of readout channels is 110,592. Flex circuits are used instead of double-metal structure to read out the z-strips, which reduces the capacitance due to the double-metal layer. Typical noise of a ladder is $\sim$ 500-1100 e^-.

The readout chain for DSSD's is based on the VA1TA integrated circuit. The VA1TA chip is a 128-channel CMOS integrated circuit, produced by IDEAS, having a trigger capability in addition to the preamplifier function. It is fabricated with the AMS $0.35\,\mu$m process and has an excellent radiation tolerance of over 20 Mrad [30].

The impact parameter resolution for reconstructed tracks is measured as a function of the track momentum p (measured in GeV/c) and the polar angle θ. It can be fitted with a function form of

$$\sigma = \sqrt{{\sigma_1}^2 + \left(\frac{\sigma_2}{\widetilde{p}}\right)^2}\,, \tag{3.1}$$

and is symbolically written as

$$\sigma = \sigma_1 \oplus \sigma_2/\widetilde{p}\,. \tag{3.2}$$

Here, $\widetilde{p}$ is the pseudo-momentum defined as

$$\widetilde{p} \equiv \begin{cases} p\beta \sin^{3/2}\theta & \text{for } r\text{-}\phi \text{ side}\,, \\ p\beta \sin^{5/2}\theta & \text{for } z \text{ side}\,. \end{cases} \tag{3.3}$$

As shown in Fig. 3.5, the impact parameter resolutions of SVD2 is better than those of SVD1, mainly owing to the smaller radius of the first layer. The impact parameter resolutions measured with the cosmic ray events are

$$\sigma_{r\phi}\ (\mu\mathrm{m}) = 19.2 \oplus 54.0/\widetilde{p}\,, \quad \sigma_z\ (\mu\mathrm{m}) = 42.2 \oplus 44.3/\widetilde{p}\,, \tag{3.4}$$

for SVD1, and

$$\sigma_{r\phi}\ (\mu\mathrm{m}) = 21.9 \oplus 35.5/\widetilde{p}\,, \quad \sigma_z\ (\mu\mathrm{m}) = 27.8 \oplus 31.9/\widetilde{p}\,, \tag{3.5}$$

for SVD2.

2-2 Central Drift Chamber (CDC)

The efficient reconstruction of charged particle tracks and precise determination of their momenta are the prerequisite to almost all of the measurements in the Belle experiment. The resolution of a transverse momentum p_t, which is the momentum component transverse to the z axis, is required to be $\sigma_{p_t}/p_t \sim 0.5\%\sqrt{1+{p_t}^2}$ (p_t in GeV/c) for all charged particles with $p_t \geq 100\,\mathrm{MeV}/c$ in the polar angle region of $17^\circ \leq \theta \leq 150^\circ$. In addition, the charged particle tracking system is expected to provide important information for the trigger system and particle identification information by the precise measurement of dE/dx.

[3]Normal paraffin grade L, Nippon Oil Corp. (`http://www.eneos.co.jp/english/`)

[4]Note that $z(\phi)$-side corresponds to the $p(n)$-side in the SVD2, while $\phi(z)$-side corresponds to the $p(n)$-side in the SVD1.

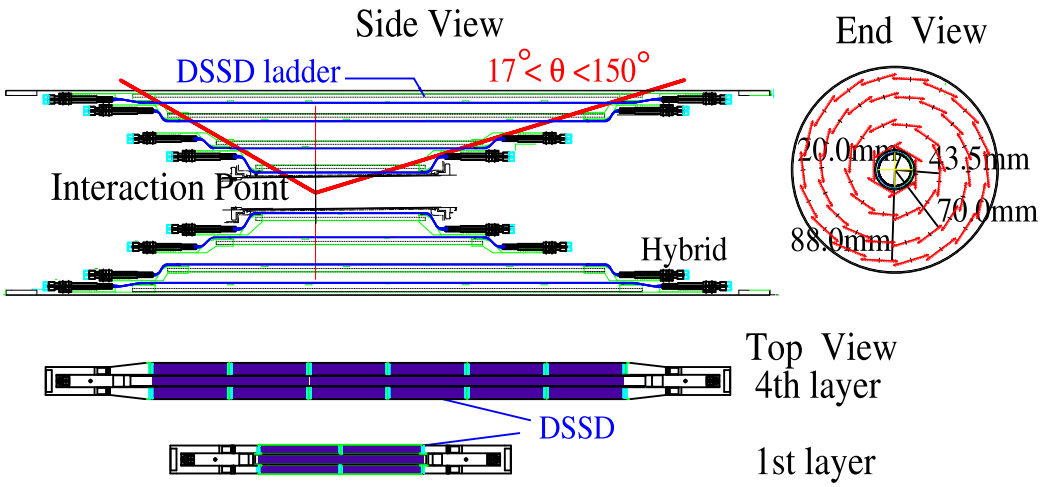


Figure 3.4: Detector configuration of SVD2.

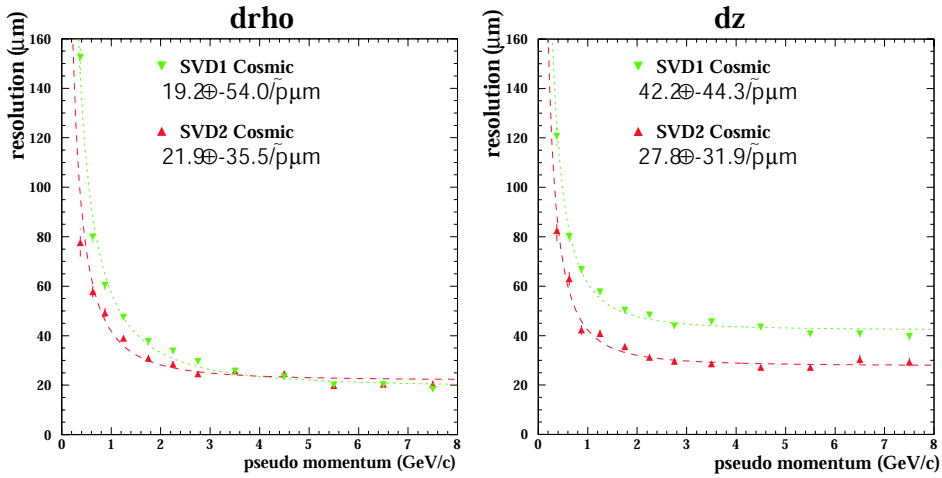


Figure 3.5: Comparison of the impact parameter resolutions in the directions of r-ϕ (left) and z (right) measured with cosmic ray data.

The central drift chamber (CDC) [32, 33] has been designed and constructed to meet above requirements for the central tracking system. Since the majority of the B decay daughters have momenta lower than 1 GeV/c, minimization of multiple scattering is important to achieve the required momentum resolution. A low-Z gas is selected to reduce the multiple scattering.

The structure of CDC used to collect DS-I is shown in Fig. 3.6. It is asymmetric in the z, providing an angular coverage of $17^\circ \leq \theta \leq 150^\circ$, which corresponds to 92% of the full solid angle. The longest wires are 2400 mm long. The outer radius is 874 mm and the inner one is extended down to 83 mm. CDC is a small-cell cylindrical drift chamber with 50 layers of anode wires, which consist of 32 axial- and 18 stereo-wire layers, and three cathode strip layers. Axial wires are parallel to the z axis, while stereo wires slant to the z axis to provide z position information. Stereo wires also provide a highly efficient fast z-trigger combined with the cathode strips. CDC has a total of 8400 drift cells. At the inner layers of CDC, three cathode-strip layers are made for higher precision z measurement at the position where the particles enter CDC, which is especially beneficial for the purpose of trigger.

In the summer of 2003, the inner part structure of CDC has been modified jointly with the upgrade of SVD. The three inner layers with cathode strips were removed to make the space for the upgraded SVD with larger radius. Instead, we have installed two layers of smaller cells, which we call small-cell CDC (sCDC). The inner radius after the modification is 104 mm, while the other geometry is unchanged. The sCDC maintains the performance of the Level-1 trigger by keeping the number of inner layers used for the trigger to be five, which was six before the modification. In addition, we exploit the small drift time due to the smaller cell to provide additional information for the Level-0 trigger logic required by SVD, which was provided by the information from TOF alone before the upgrade.

A low-Z gas mixture, consisting of 50% He and 50% ethane (C_2H_6), is used to minimize multiple Coulomb scattering to achieve a good momentum resolution, especially for low momentum particles. 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 radiation. Even though the gas mixture has a low-Z, a good dE/dx resolution is obtained by the large ethane component.

The measured spatial resolution in the r-ϕ direction[5] is $\sim$ 120-150 μm with a dependence on the incident angles and layers. The p_t resolution obtained by the study using cosmic ray is

$$\frac{\sigma_{p_t}}{p_t}\,(\%) = \sqrt{(0.28p_t)^2 + (0.35/\beta)^2} \quad (p_t \text{ in GeV}/c) \tag{3.6}$$

without the SVD information, and

$$\frac{\sigma_{p_t}}{p_t}\,(\%) = \sqrt{(0.19p_t)^2 + (0.30/\beta)^2} \quad (p_t \text{ in GeV}/c) \tag{3.7}$$

with the SVD information. (Fig 3.7)

The dE/dx measurement in CDC can distinguish particle species, since the mean energy loss $\langle dE/dx \rangle$ for a charged particle is given as a function of the velocity, by Bethe-Bloch formula. A scatter plot of the measured dE/dx and particle momentum is shown in Fig. 3.8, together with the expected mean energy losses for different particle species. Populations of pions, kaons, protons, and electrons can be clearly seen. The dE/dx resolution is measured to be 7.8% in the momentum range from 0.4 to 0.6 GeV/c.

2-3 Aerogel Čerenkov Counter System (ACC)

Particle identification, particularly the identification of $\pi^\pm$ against $K^\pm$, plays an important role in the many analyses of B decays. An array of silica-aerogel threshold Čerenkov counters

[5] We define the "r-ϕ direction" as the axis that is perpendicular to the z direction on the plane of each DSSD.

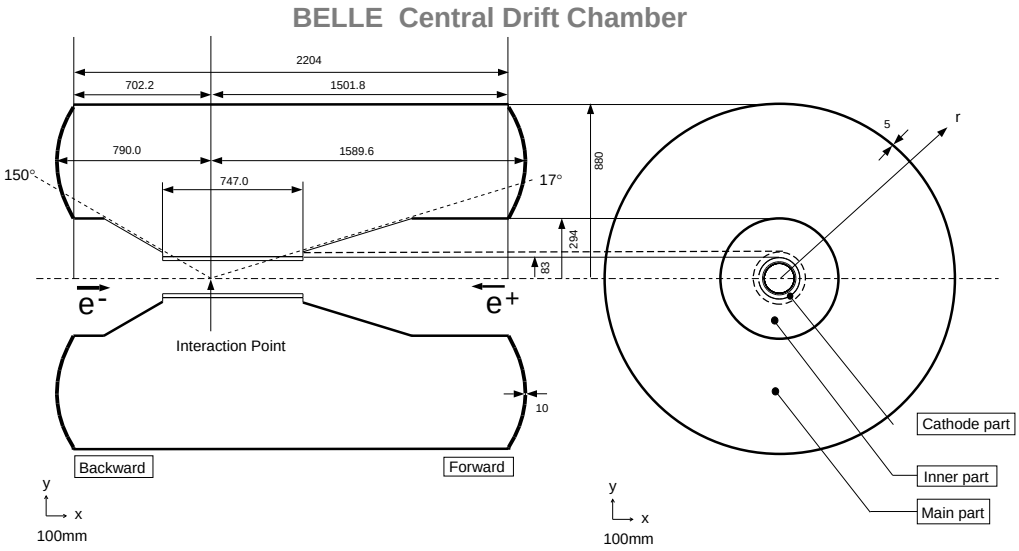


Figure 3.6: Overview of the CDC structure

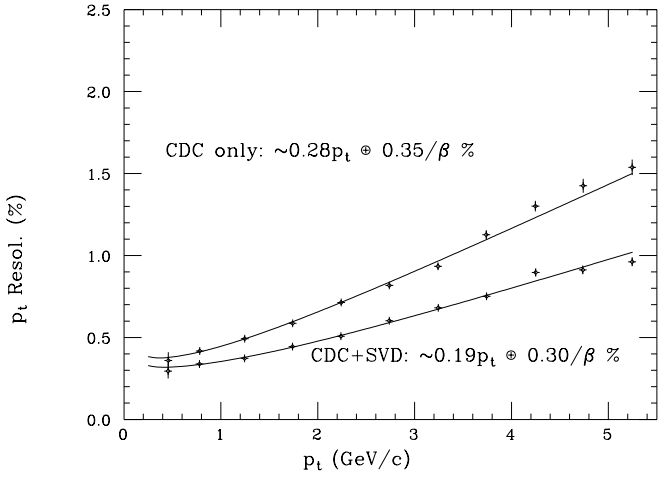


Figure 3.7: p_t resolution studied using cosmic rays.

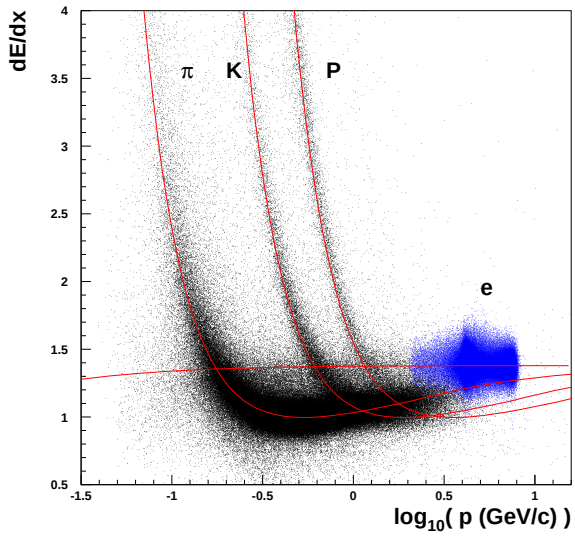


Figure 3.8: Charged track momenta vs. dE/dx observed in collision data.

is selected as a part of the Belle particle identification system. It covers the momentum range between 1.5 GeV and 3.5 GeV with respect to the $K^{\pm}/\pi^{\pm}$ separation, extending the coverage beyond the reach of dE/dx measurements by CDC and time-of-flight measurements by TOF.

The Čerenkov radiations are emitted in case of

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

where β, m, and p are the velocity, mass, and momentum of the charged particle, respectively; and n is the refractive index of the matter through which the particle is passing. Since $m_{K^{\pm}} > m_{\pi^{\pm}}$, there is a momentum region where the pions emit Čerenkov light, while kaons and heavier particles do not. Thus, one can identify pions against kaons by choosing proper refractive index n for the momentum region of interest.

The configuration of ACC [34] is shown in Fig. 3.9. ACC consists of 960 counter modules segmented into 60 cells in the ϕ direction for the barrel part and 228 modules arranged in five concentric layers for the forward end-cap part of the detector. All the modules are arranged in a semi-tower geometry, pointing to IP. A typical ACC module consists of five aerogel tiles stacked in a thin (0.2 mm thick) aluminum box with an approximate size of $12 \times 12 \times 12\mathrm{cm}^3$.

To detect the Čerenkov lights, two (one) fine-mesh type photomultiplier tubes (FM-PMTs) are attached to each module in the barrel (end-cap) part. This FM-PMTs are designed to operate in strong magnetic field of 1.5 T [35].

In order to obtain a good $K^{\pm}/\pi^{\pm}$ separation for the required kinematic range, the refractive indices of aerogels are selected to be between 1.01 and 1.03, depending on their polar angle region. In barrel part, they are optimized for the momentum corresponding to the daughter particles of B meson two-body decays. In end-cap part, they are optimized for the momentum of $K^{\pm}$ from B cascade decays, which is advantageous in B flavor tagging.

The performance of ACC is confirmed using the decay process of $D^{*+} \to \pi^+ D^0 (\to K^- \pi^+)$,

where the identification of the charged particles from the D^0 decay can be determined without using the ACC information, by the charge of π from the D^{*+} decay. Figure 3.10 shows the number of photo-electron distribution of $\pi^{\pm}$ and $K^{\pm}$ in this decay process, where $\pi^{\pm}$ is well separated from $K^{\pm}$, being consistent with MC.

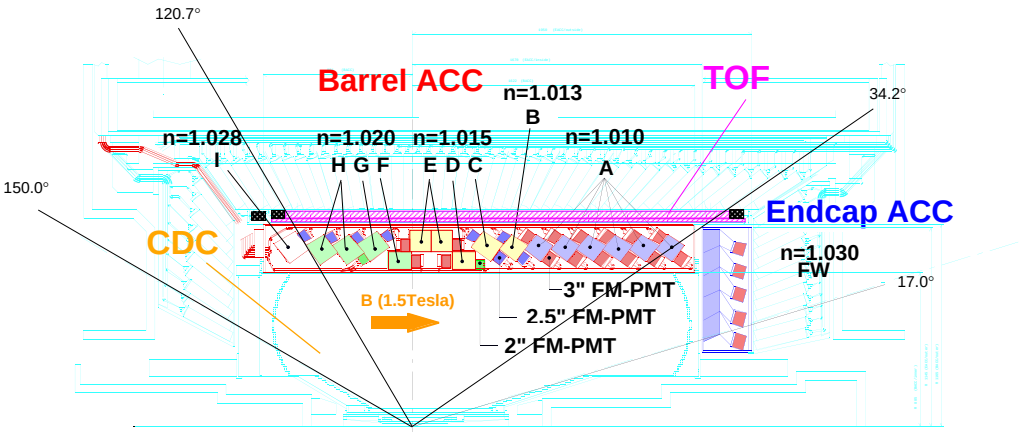


Figure 3.9: Arrangement of ACC in Belle detector.

2-4 Time-of-Flight Counter (TOF)

A time-of-flight (TOF) detector system using plastic scintillation counters is very powerful for particle identification in e^+e^- collider detectors. For a 1.2 m flight path, the TOF system with 100 ps time resolution is effective for particle momenta below about 1.2 GeV/c. Roughly 90% of the particles produced in $\Upsilon(4S)$ decays are in this momentum region. It can provide clean and efficient b-flavor tagging. In addition to particle identification, the TOF counters provide fast timing signals for the trigger system. To avoid pile-up in the trigger queue, the rate of the TOF trigger signals must be kept below 70 kHz. Simulation studies indicate that to keep the fast trigger rate below 70 kHz in any beam background conditions, the TOF counters should be supplemented by thin trigger scintillation counters (TSC).

The following relation is satisfied between the time-of-flight T measured with TOF and the momentum p measured momentum with CDC:

$$T = \frac{L}{c\beta} = \frac{L}{c}\sqrt{1+\left(\frac{m}{p}\right)^2}\,, \tag{3.9}$$

where L is a length of the flight. For example, when $L = 120$ cm and $p = 1.2$ GeV/c, $T = 4.0$ ns for a pion ($m_{\pi^\pm} = 140$ MeV/c^2), while $T = 4.3$ for a kaon ($m_{K^\pm} = 494$ MeV/c^2). Thus, the difference of T between pions and kaons is ~ 300 ps and $K^\pm/\pi^\pm$ separation with 3σ significance is obtained with the time resolution of 100 ps.

The Belle TOF system [36] consists of 128 TOF counters and 64 TSC counters. Two trapezoidally shaped TOF counters and one TSC counter, with a 1.5 cm intervening radial gap, form one module. In total 64 TOF/TSC modules located at a radius of 1.2 m from IP cover a polar angle range from 34° to 120°. The minimum transverse momentum to reach the TOF counters is about 0.28 GeV/c. The dimensions of a module are given in Fig. 3.11. The modules are individually mounted on the inner wall of the barrel ECL container. The

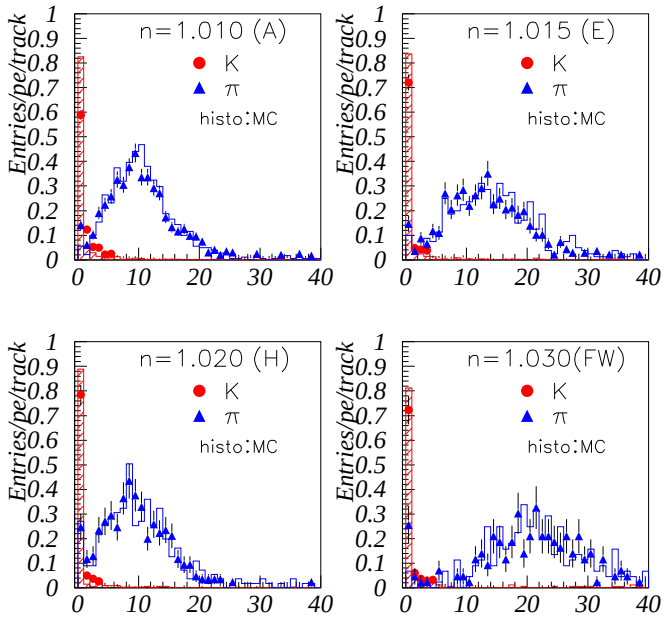


Figure 3.10: ACC number of photo-electron distribution for $\pi^{\pm}$ and $K^{\pm}$ from $D^{*\pm}$ decays. Each plot corresponds to the different set of modules with a different refractive index.

1.5 cm gaps between the TOF counters and TSC counters are introduced to isolate TOF from photon conversion backgrounds by taking the coincidence between the TOF and TSC counters. Electrons and positrons created in the TSC layer are impeded from reaching the TOF counters due to this gap in a 1.5 T field. Fine-mesh photomultiplier tubes (FM-PMTs) are attached to both ends of the TOF counter with air gaps of 0.1 mm. The air gaps for the TOF counter selectively pass earlier arrival photons with small incident angle and reduce a gain saturation effect of FM-PMTs due to large pulses at a very high rate. Since the time resolution is determined by the rising edge of the time profile of arrival photons at PMT, the air gaps hardly affect the time resolution. As for the TSC counters, the tubes are glued to the light guides at the backward ends.

Figure 3.12 shows time resolutions as a function of z for forward and backward PMTs and for the weighted average. The resolution for the weighted average is about 100 ps with a small z dependence. This satisfies the requirement. Figure 3.13 shows the mass distribution for each track in hadron events, calculated using Eq. (3.9) using the momentum of the particle determined from the CDC track fit assuming muon mass. Clear peaks corresponding to pions, kaons, and protons are seen. The data points are in good agreement with an MC expectation (histogram) obtained assuming the time resolution of TOF $\sigma_{\rm TOF} = 100$ ps.

2-5 Electromagnetic Calorimeter (ECL)

The main purpose of the electromagnetic calorimeter is the detection of photons 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, since important two-body decay modes, such as $B \to K^*\gamma$ and $B^0 \to \pi^0\pi^0$, produce photons energies up to 4 GeV, good resolution for high momentum region is also needed to reduce backgrounds for these modes. Electron identification in Belle relies primarily on a comparison of the charged particle

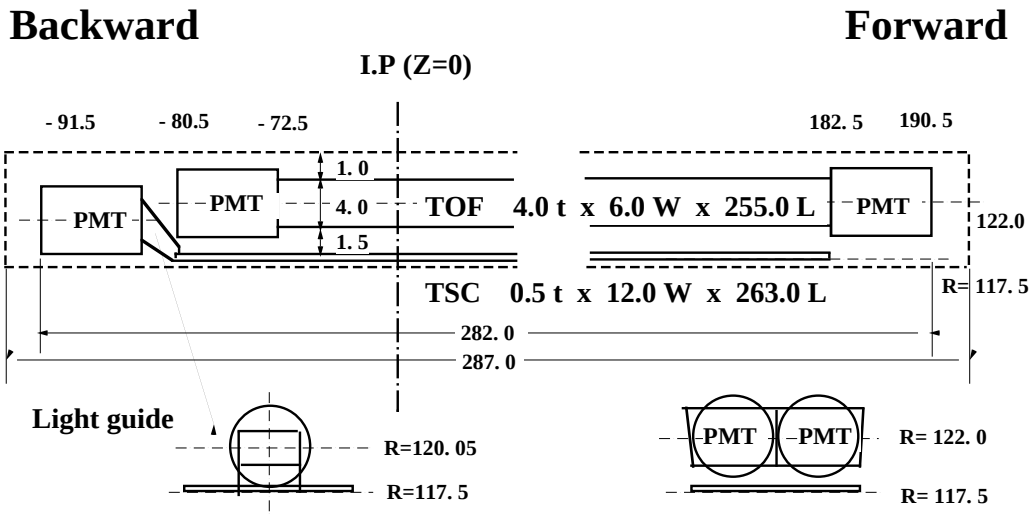


Figure 3.11: Dimensions of a TOF/TSC module.

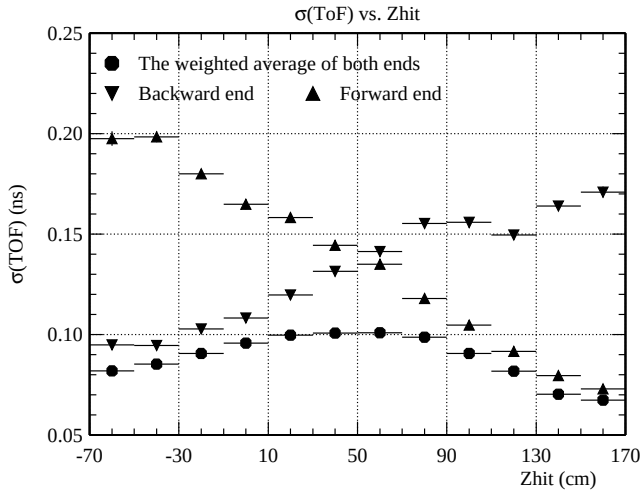


Figure 3.12: Time resolution of TOF for $e^+e^- \to \mu^+\mu^-$ events.

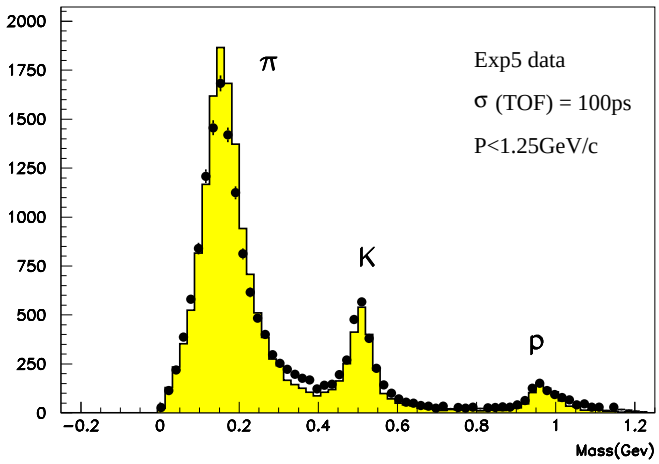


Figure 3.13: Mass distribution from TOF measurements for particles with momenta below 1.2 GeV/c. The histogram corresponds to MC distribution.

momentum and the energy deposits in the electromagnetic calorimeter. Good electromagnetic energy resolution results in better hadron rejection against electron. High momentum π^0 detection requires the separation of two nearby photons and a precise determination of their opening angle. This requires a fine-grained segmentation in the calorimeter.

In order to satisfy the above requirements, we use a highly segmented array of CsI(Tl) crystals with silicon photodiode readout installed in a magnetic field of 1.5 T inside a superconducting solenoid magnet. CsI(Tl) crystals have desirable features of a large photon yield, weak hygroscopicity, mechanical stability, and moderate price.

The overall configuration of the Belle calorimeter system, ECL [37], is shown in Fig. 3.14. ECL consists of the barrel section of 3.0 m in length with the inner radius of 1.25 m and the annular end-caps at $z = +2.0$ m and -1.0 m from IP. Each crystal has a tower-like shape and is arranged to point almost to IP. There are small tilt angles from the direction exactly pointing to the IP to avoid photons to escape through the gap of the crystals. In the barrel section the tilt is $\sim 1.3^\circ$ in the θ and ϕ directions. Forward (backward) end-cap crystals are tilted by $\sim 1.5^\circ$ ($\sim 4.0^\circ$) in the θ direction. The calorimeter covers the polar angle region of $17.0^\circ < \theta < 150.0^\circ$, corresponding to a total solid angle coverage of 91% of 4π sr. Small gaps between the barrel and end-cap crystals provide a pathway for cables and room for supporting structures of the inner detectors. The loss of solid angle associated with these gaps is approximately 3% of the total acceptance. The entire system contains 8736 CsI(Tl) counters and weighs 43 tons.

The size of a crystal in the θ-ϕ direction is determined so that a crystal contains approximately 80% of the total energy deposit by a photon injected at the center of its front face. The typical dimension of a crystal is 55 mm × 55 mm at front face and 65 mm × 65 mm at rear face for the barrel part. The length (in r direction) is 30 cm, which corresponds to 16.2 X_0 (radiation length). This length is long enough to avoid deterioration of the energy resolution at high energies due to the shower leakage from rear of the counter.

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

$$\sigma\,(\mathrm{mm}) = 0.27 + \frac{3.4}{\sqrt{E}} + \frac{1.8}{\sqrt[4]{E}} \quad (E \text{ in GeV}) , \tag{3.10}$$

which is shown in Fig. 3.15. As can be seen in the figure, the estimation is well consistent with the result of the beam test [37] in the measured energy region. The energy resolution given by

the beam test is

$$\frac{\sigma_E}{E}\,(\%) = \sqrt{\left(\frac{0.066}{E}\right)^2 + \left(\frac{0.81}{\sqrt[4]{E}}\right)^2 + 1.34^2} \quad (E \text{ in GeV}) \,. \tag{3.11}$$

This is consistent with the collision data calibrated by $e^+e^- \to e^+e^-$ (Bhabha) events, where the energy resolutions are 1.5%, 1.9%, and 2.5% for the barrel, forward, and backward ECL, respectively. (Fig. 3.16)

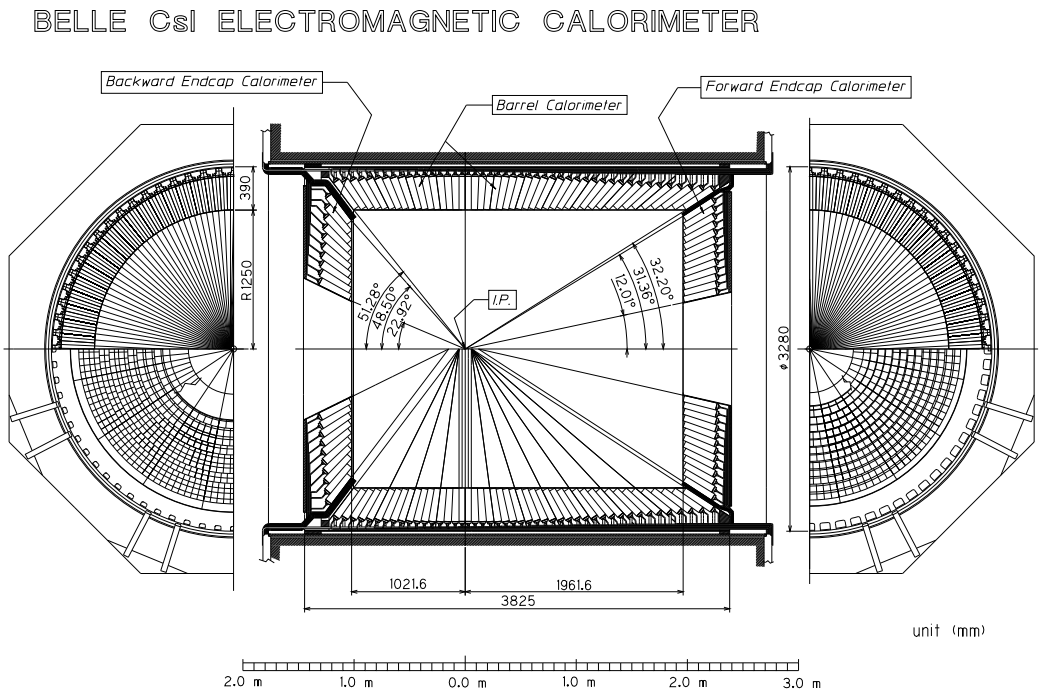


Figure 3.14: Configuration of ECL.

2-6 K_L^0 and Muon Detection System (KLM)

KLM [38] is designed to identify K_L^0 and muon with high efficiency over a broad momentum range greater than 600 MeV/c.

KLM consists of alternating layers of charged particle detectors and 4.7 cm-thick iron plates. The barrel-shaped region around IP 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°. 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 iron plates provide a total of 3.9 interaction lengths of material for a particle traveling normal to the detector planes. In addition, ECL provides another 0.8 interaction length of material to convert K_L^0. K_L^0 that interacts in the iron plates of KLM or ECL produces a shower of ionizing particles. The position information of this shower with respect to the IP determines the flight direction of K_L^0, while the fluctuations in the size of the shower is so large that it is impossible to measure the energy of K_L^0 in useful resolution. The multiple layers of charged particle detectors and iron allow the discrimination between muons and charged hadrons ($\pi^\pm$ or $K^\pm$) based on their range and transverse scattering. Muons travel much farther with smaller deflections on average than strongly interacting hadrons.

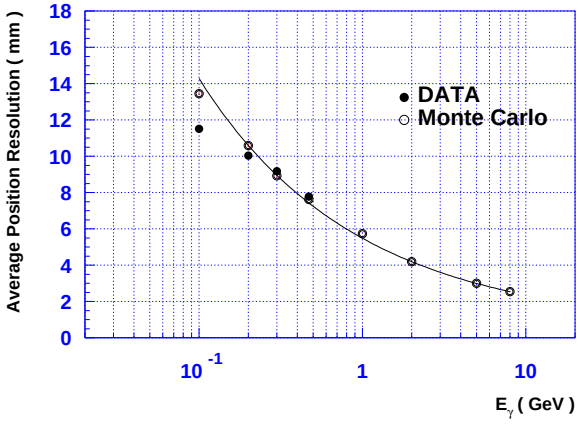


Figure 3.15: Energy dependence of the average position resolution. The solid curve is the result of the fit to the MC.

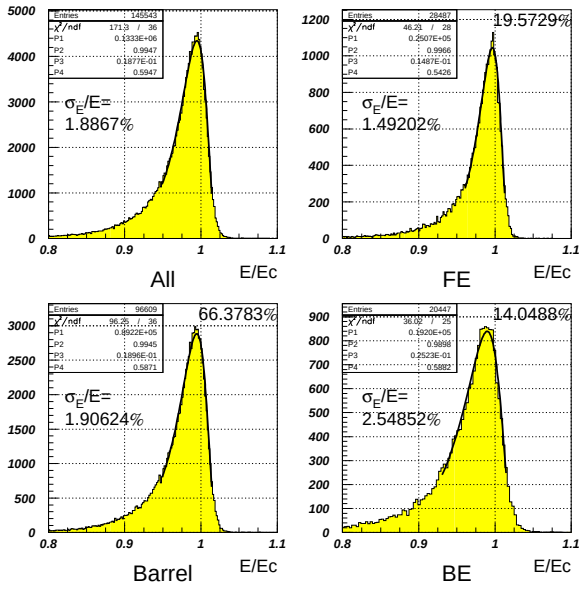


Figure 3.16: Energy resolutions calibrated and measured with $e^+e^- \to e^+e^-$ (Bhabha) events. The plots correspond the overall average (top left) and each of the barrel (top right), forward end-cap (bottom left), and backward end-cap (bottom right) sections.

The detection of charged particles is provided by glass electrode resistive-plate counters (RPCs). Resistive plate counters have two parallel plate electrodes with high bulk resistivity ($\geq 10^{10}\,\Omega$cm) separated by a gas-filled gap. We choose a incombustible mixture of 62% HFC-134a, 30% argon, and 8% butane silver. Butane silver is a mixture of approximately 70% n-butane and 30% iso-butane. In a 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.

Figure 3.17 shows the cross-section of a superlayer for the barrel region, in which two RPCs are sandwiched between the orthogonal θ and ϕ pickup strips with the ground planes for signal reference and proper impedance. This unit structure of two RPCs and two readout planes is enclosed in an aluminum box and is less than 3.7 cm thick. Each RPC is electrically insulated with a double layer of 0.125 mm thick mylar. Signals from both RPCs are picked up by copper strips above and below the pair of RPCs, providing a three-dimensional space point information for particle tracking. Each barrel module has two rectangular RPCs with 48 z pickup strips perpendicular to the beam direction. The smaller seven superlayers closest to IP have 36 ϕ strips and the outer eight superlayers have 48 ϕ strips orthogonal to the z strips. Each end-cap superlayer module contains 10 π-shaped RPCs and have the 96 ϕ and 46 θ pickup strips.

Figure 3.18 shows a histogram of the difference between the direction of the neutral cluster (K_L^0 candidate) detected by KLM and the missing momentum direction in data. The missing momentum vector is calculated using all the other measured particles in the event. The histogram shows a clear peak where the direction of the neutral cluster measured in KLM is consistent with the missing momentum in the event, indicating correct detection of K_L^0. The non-peaking flat-distributed component in the histogram is mainly due to undetected neutrinos and particles escaping the detector acceptance.

2-7 Extreme Forward Calorimeter (EFC)

EFC[6] [39] extends the polar angle coverage by ECL, which is $17^\circ < \theta < 150^\circ$. 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. EFC is also required to function as a beam mask to reduce backgrounds for CDC. In addition, EFC is used for a beam monitor for the KEKB control and a luminosity monitor for the Belle experiment. It can also be used as a tagging device for two-photon physics. Since EFC is placed in the very high radiation level area around the beam pipe near IP, it is required to be radiation hard. Thus, a radiation-hard BGO (Bismuth Germanate, $Bi_4Ge_3O_{12}$) crystal calorimeter is used for EFC. The detector is segmented into 32 in ϕ and 5 in θ for both the forward and backward detectors. The radiation lengths of the forward and backward crystals are 12 and 11, respectively.

The energy sum spectra for $e^+e^- \to e^+e^-$ (Bhabha) events show a correlation between the forward and backward EFC detectors as expected. A clear peak at 8 GeV (3.5 GeV) with a resolution of 7.3% (5.8%) in rms is seen for the forward (backward) EFC.

2-8 Solenoid Magnet

A 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 [40]. The coil is surrounded by a multilayer structure consisting of iron plates and calorimeters, which is integrated into a magnetic return circuit. The iron structure of the Belle detector serves as the return path of magnetic flux and an absorber material for KLM. It also provides the overall support for all of the detector components.

[6] EFC is not explicitly used in this analysis.

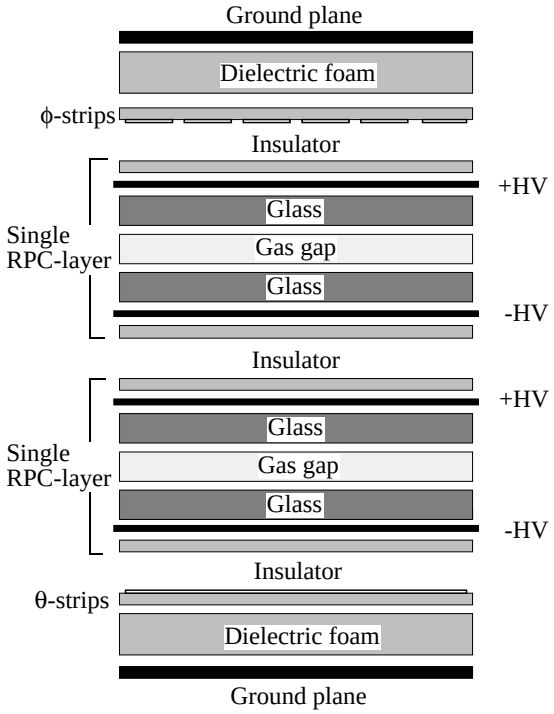


Figure 3.17: Cross section of a KLM superlayer.

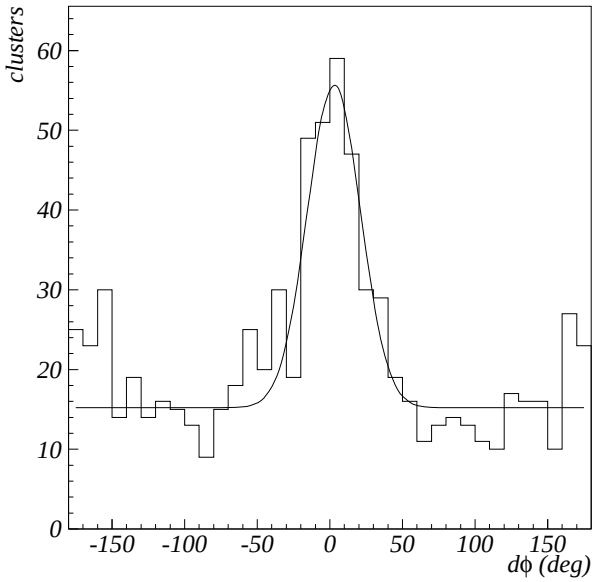


Figure 3.18: Distribution of the difference between directions of the neutral cluster detected by KLM and the missing momentum.

2-9 Trigger System (TRG)

The cross section for physics events of interest, like $e^+e^- \to BB$ event, is smaller than those of the background processes, like $e^+e^- \to u\overline{u}$. Thus, they have to be triggered by appropriately restrictive conditions. In addition, high beam backgrounds are also expected due to the high beam current. Since the rates are very sensitive to actual accelerator conditions, it is difficult to make a reliable estimate. Therefore, 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. It is important to have redundant triggers to keep the efficiency high even for varying conditions. The Belle trigger system is designed to satisfy these requirements.

The Belle trigger system consists of the Level-1 hardware trigger and the Level-3 software trigger. The latter is designed to be implemented in the online computer farm. Figure 3.19 shows the schematic view of the Belle Level-1 trigger system [41]. It consists of the sub-detector trigger systems and the central trigger system called the Global Decision Logic (GDL). The sub-detector trigger systems are classified into two categories: track triggers and energy triggers. CDC and TOF are used to yield trigger signals for charged particles. CDC provides r-ϕ and r-z track trigger signals. The ECL trigger system provides triggers based on the total energy deposit and the cluster counting of crystal hits. These two categories allow sufficient redundancy. The KLM trigger gives additional information on muons and the EFC triggers are used for tagging two photon events as well as Bhabha events. The sub-detectors process event signals in parallel and provide trigger information to GDL, where all information is combined to characterize an event type.

Accompanied with the upgrade of SVD in the summer of 2003, we have implemented trigger capability using the track information from SVD. Though we do not use it for the data taking currently, it has proved to effectively reduce the trigger rate with slight loss of the events of interest and to be useful in future when the accelerator is improved to have higher beam current and the beam background gets larger.

The trigger system provides the trigger signal with the fixed time of 2.2 μs after the event occurrence. The Belle trigger system, including most of the sub-detector trigger systems, is operated in a pipelined manner with clocks synchronized to the KEKB accelerator RF signal. The typical Level-1 trigger rate as of October 2006 is $\sim$ 400-600 Hz, which is well below the requirement, $\lesssim$ 800 Hz, from the current data acquisition system (DAQ). The trigger rate is dominated by the beam background. The trigger efficiency is monitored from the data using the redundant triggers. Each of the multitrack, total energy, and isolated cluster counting triggers provides more than 96% efficiency for multi-hadronic data samples. The combined efficiency is more than 99.5%.

2-10 Data Acquisition System (DAQ)

In order to satisfy the data acquisition requirements so that it works at $\sim$ 400 kHz with a deadtime fraction of less than 10%, the distributed-parallel system is devised. The global scheme of the original system is shown in Fig. 3.20. The entire system is segmented into seven subsystems running in parallel, each handling the data from a sub-detector. Data from each subsystem are combined into a single event record by an event builder, which converts "detector-by-detector" parallel data streams to an "event-by-event" data river. The event builder output is transferred to an online computer farm, where another level of event filtering is done after the fast event reconstruction. The data are then sent to a mass storage system located at the computer center via optical fibers. With this system, a deadtime fraction of $\sim$ 8% for the Level-1 trigger rate of 400 Hz is achieved.

To date, the DAQ system has undergone several improvements to keep up with the in-

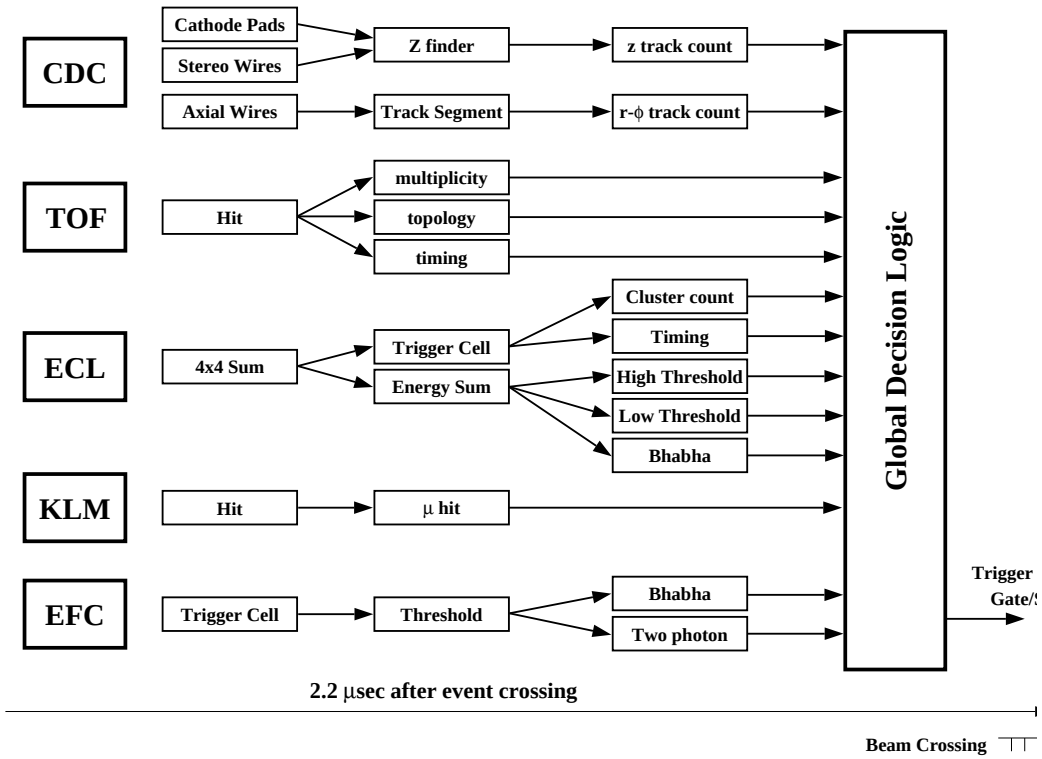


Figure 3.19: Schematic figure of the Level-1 trigger system for Belle detector.

creasing luminosity and beam background. The upgrade of SVD in the summer of 2003 also involved the entire replacement of the DAQ part of the SVD, enhancing the performance of the system such that the deadtime of the SVD DAQ is not a bottleneck for the whole DAQ system anymore. At the same time, some other parts of the DAQ system has also been modified. The DAQ subsystem for some sub-detectors with large numbers of channels are subdivided and the degree of parallelism is enhanced. These modifications greatly improved the performance of the DAQ system as a whole and now it achieves $\sim 2\%$ deadtime fraction for the Level-1 trigger rate of 400 Hz. This is well acceptable for the current trigger rate of $\sim$ 400-600 Hz. There is also an upgrade in the lower stream of the system. A PC farm to perform the event reconstruction (RFARM) is introduced at the lowest stream, which makes it possible to reconstruct the events simultaneously with the data taking. This system has been in full operation since the summer of 2004 and now the raw data is not stored.

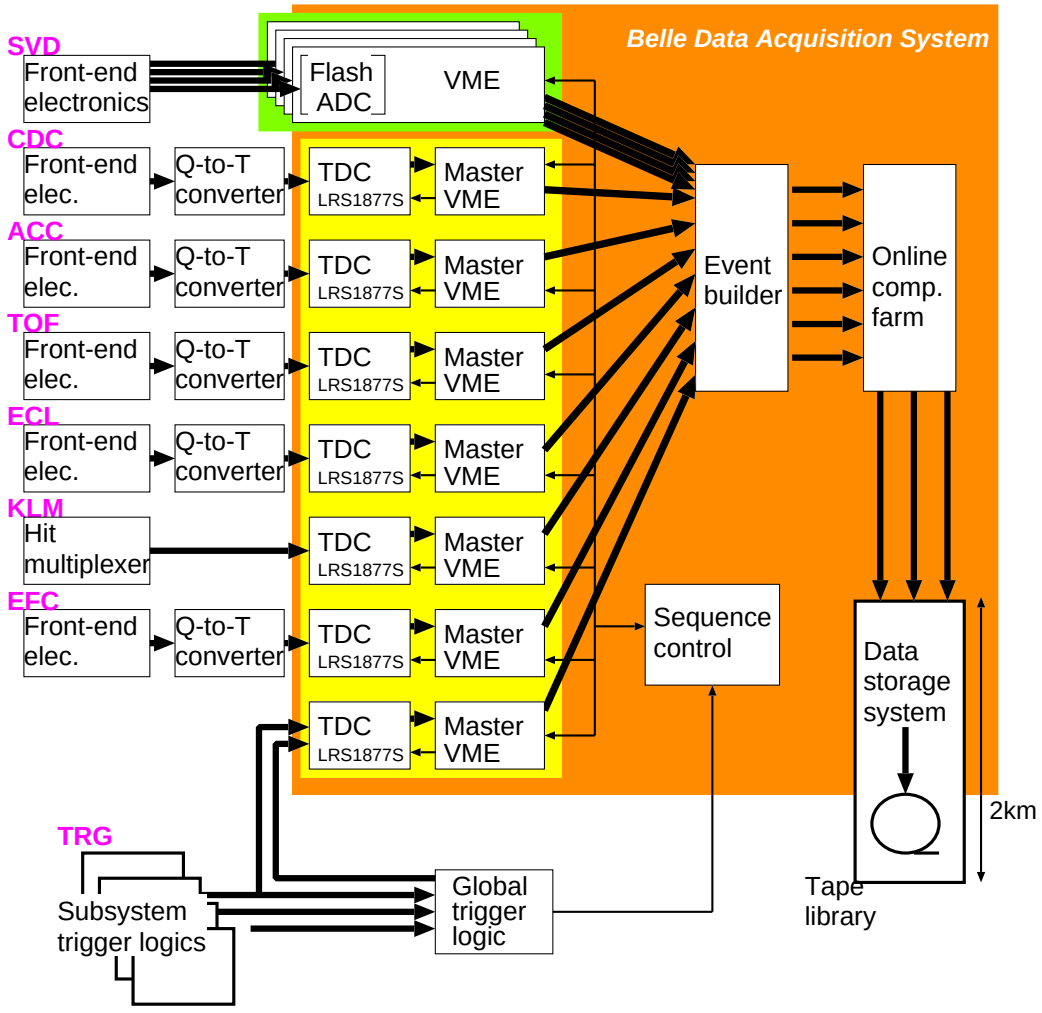


Figure 3.20: An overview of the original Belle DAQ system.

Chapter 4

Analysis Tools and Techniques

1 Particle Identification

In this section, we describe our strategy of the particle identification. Particle identification plays important roles in 1) the suppression of the $b \to c$ decay backgrounds such as $B^0 \to \overline{D}^0\pi^0(\overline{D}^0 \to K^+\pi^-)$, which has large branching fraction and can mimic $B^0 \to \pi^+\pi^-\pi^0$ without K^+ identification, and 2) the flavor tagging, where the charges of the identified leptons and strange particles are very important.

1-1 $K^\pm/\pi^\pm$ Identification

We discriminate between $K^\pm$ and $\pi^\pm$ by combining following three measurements [42]:

- CDC measurement of dE/dx,
- TOF measurement, and
- ACC measurement of the number of photoelectrons ($N_{\rm pe}$).

The three measurements covers different momentum regions; Fig. 4.1 shows the momentum coverage of each measurement for the $K^\pm/\pi^\pm$ separation. We model the probability density functions (PDF's) from the responses of the detectors for each of $K^\pm$ and $\pi^\pm$; the likelihood functions for each detectors are calculated based on the PDF's. The product of the three likelihoods for the three measurements is the overall likelihood probability for being a kaon ($\mathcal{L}_K$) or a pion ($\mathcal{L}_\pi$). A particle is then identified as a kaon or a pion by a selection criterion based on the likelihood ratio $\mathcal{R}_{K/\pi}$:

$$\mathcal{R}_{K/\pi} \equiv \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi} \, . \tag{4.1}$$

The validity of the $K^\pm/\pi^\pm$ identification is demonstrated using the data of the charm decay chain of $D^{*+} \to D^0\pi^+(D^0 \to K^-\pi^+)$. In the decay chain, we can use the charge of the π from the first D^{*+} decay to determine the $K^\pm/\pi^\pm$ identification of the charged tracks from the subsequent D^0 decays. Note that the mass difference between D^{*+} and D^0 is 145 MeV, which is only ~ 6 MeV above the $\pi^\pm$ mass, and thus the π from the first decay of D^{*+} has characteristic low momentum. This makes it possible to obtain a very pure sample ($S/N > 30$) without relying on the information of particle information. Figure 4.2 shows a two-dimensional distribution of the likelihood ratio $\mathcal{R}_{K/\pi}$ and measured momenta for the kaon and pion tracks. The figure demonstrates the clear separation between kaons and pions up to around 4 GeV/c. The measured K efficiency and π fake rate in the barrel region are plotted as functions of the track momentum in Fig. 4.3, where a selection criterion of $\mathrm{R}_{K/\pi} > 0.6$ is applied.

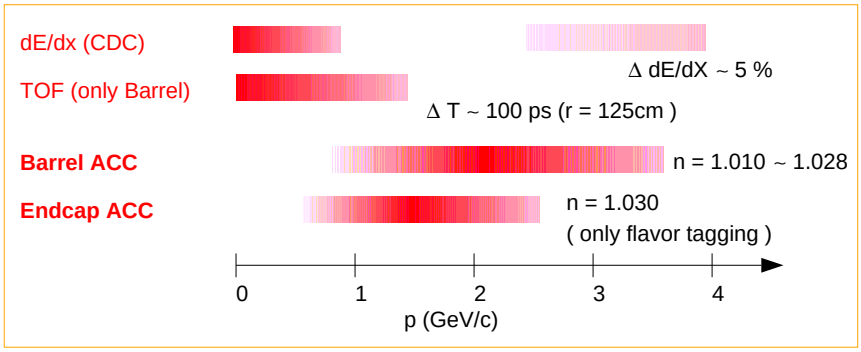


Figure 4.1: Momentum coverage of each detector used for $K^{\pm}/\pi^{\pm}$ separation.

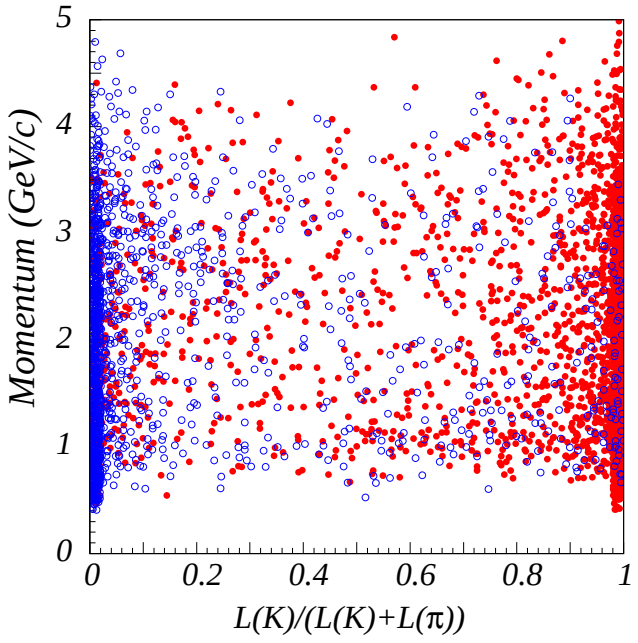


Figure 4.2: A scatter plot of the track momentum (vertical axis) and the likelihood ratio $\mathcal{R}_{K/\pi}$ (horizontal axis) for $K^{\pm}$ (closed circle) and $\pi^{\pm}$ (open circle) obtained from the data of $D^{*+} \to D^0\pi^+(D^0 \to K^-\pi^+)$ decays. Strong concentration in the region of $\mathcal{R}_{K/\pi} \sim 1$ (~ 0) is observed for $K^{\pm}$ ($\pi^{\pm}$) over a wide momentum region up to $\sim 4\,\mathrm{GeV}/c$.

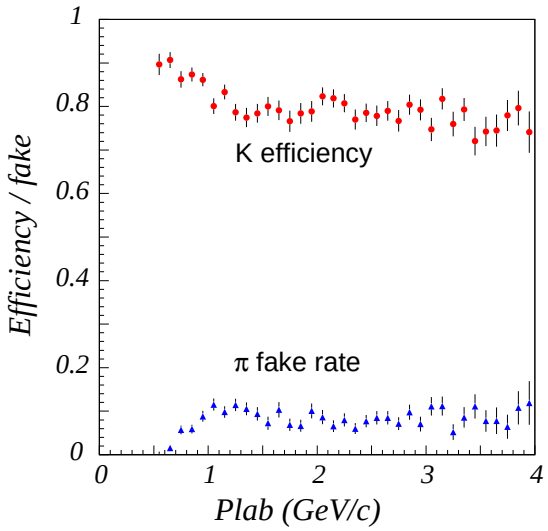


Figure 4.3: K efficiency and π fake rate in the barrel region measured using the data of $D^{*+} \to D^0\pi^+(D^0 \to K^-\pi^+)$ decays. The selection criterion of $\mathrm{R}_{K/\pi} > 0.6$ is applied here.

1-2 Electron Identification

We use the following five discriminants to distinguish electrons against hadrons and muons [43]:

Matching χ^2 Matching between the position of the charged track extrapolated to the ECL and the energy cluster position measured by the ECL. An electron show a good matching, i.e., small χ^2, since the electromagnetic shower in the ECL develops along the electron track. On the other hand, the matching is worse for the hadrons, since the energy deposit by a hadron comes from interactions such as $\pi^+ n \to \pi^0 p$ ($\pi^0 \to \gamma\gamma \to$ EM shower) where the flight directions of the secondary π^0 and the subsequent γ are not always well correlated with that of the primary π^+. (Fig. 4.4, left)

***E/p* ratio** The ratio of the energy measured by the ECL, E, and the charged track momentum, p, measured by the CDC. An electron yields E/p ratio of ~ 1 since it deposits almost all of its energy in ECL, while a hadron does not. (Fig. 4.4, middle)

***E*9/*E*25 ratio** Transverse shower shape at the ECL, defined as the ratio between the energy deposit in the 3×3 array of ECL crystals around the cluster center ($E9$) and that in the 5×5 array ($E25$). Electrons have a peak at around $E9/E25 \sim 1$ with small tail in the small $E9/E25$ region, while hadrons have larger tail in the small $E9/E25$ region. The reason is the same as the case of matching χ^2; secondary π^0 and γ from the hadron interaction tend to have large transverse momentum and energy deposit can spread over a wide region. (Fig. 4.4, right)

***dE/dx* in the CDC** Electrons and hadrons with the same momenta have different velocity and thus exhibit different dE/dx.

Light yield (N_{pe}) in ACC Electrons and hadrons with the same momenta have different velocity and thus yield different amount of light in ACC.

Corresponding to all the five discriminants, likelihood functions are calculated for electrons and non-electrons. Here the non-electrons are the mixture of the particles other than electrons

(hadrons and muons), whose composition is obtained relying on the MC of generic B decays. The total likelihood functions for electrons ($\mathcal{L}_e$) and non-electrons ($\mathcal{L}_{\overline{e}}$) are defined as the products of the five likelihood functions. With the total likelihood, we define the likelihood ratio $\mathcal{R}_{e/\overline{e}}$:

$$\mathcal{R}_{e/\overline{e}} \equiv \frac{\mathcal{L}_e}{\mathcal{L}_e + \mathcal{L}_{\overline{e}}} , \tag{4.2}$$

which we use for track selection criteria. Figure 4.5 shows the likelihood distributions for electrons and pions, where electrons are identified. The performance of the electron identifications, i.e., the efficiency and the fake rate, is well calibrated using various processes of $e^+e^- \to e^+e^-\gamma$, $e^+e^- \to e^+e^-e^+e^-$, hadronic events with a single $e^\pm$, hadronic events with photon conversion, and $J/\psi \to e^+e^-$, where the difference between data and MC is studied and well understood.

1-3 Muon Identification

Muon identification [44] is based on the difference of interaction in material between muons and hadrons. Since a muon are a massive lepton, it deposits its energy only through the multiple-Coulomb scattering, while an electron, the almost massless lepton, deposits its energy by the creation of an electromagnetic shower and a hadron deposit their energy through hadronic interactions. Electrons fully deposit their energy in the ECL and rarely reach KLM and thus can be easily distinguished from muons.

Muons are identified against hadrons as follows. A track is extrapolated from the CDC to the KLM and associated KLM hits are searched; a track is re-fitted with those associated KLM hits, assuming that a track deposit its energy only by multiple scattering; we use the following two information obtained in this procedure for the muon identification:

- Range of the associated KLM hits. The difference between measured and expected ranges is used as the discriminant, and
- Goodness of the matching between the position of the associated KLM hits and that obtained by extrapolating the CDC track.

The likelihood functions are calculated for the two discriminants; we calculate the total likelihood functions for muons ($\mathcal{L}_\mu^{\mu\text{-ID}}$), pions ($\mathcal{L}_\pi^{\mu\text{-ID}}$), and kaons ($\mathcal{L}_K^{\mu\text{-ID}}$), as the products of the likelihood functions of the discriminants. Based on the likelihood functions, the muon likelihood ratio ($\mathcal{R}_{\mu/\pi,K}$) is calculated:

$$\mathcal{R}_{\mu/\pi,K} \equiv \frac{\mathcal{L}_\mu^{\mu\text{-ID}}}{\mathcal{L}_\mu^{\mu\text{-ID}} + \mathcal{L}_\pi^{\mu\text{-ID}} + \mathcal{L}_K^{\mu\text{-ID}}} . \tag{4.3}$$

Figure 4.6 shows the efficiency for muons and the fake rate for pions estimated using the data of $e^+e^- \to e^+e^-\mu^+\mu^-$ and $K_S \to \pi^+\pi^-$. For momentum above 1 GeV, the efficiency is above $\sim 90\%$ while the fake rate is below $\sim 2\%$.

2 Flavor Tagging

In the measurement of CP violation, we need to know the flavor of the B that decays into $f_{\rm tag}$, which we call $B_{\rm tag}$, in the decay chain of $\Upsilon(4S) \to B^0\overline{B}^0 \to f_{CP}f_{\rm tag}$, where $f_{CP} = \pi^+\pi^-\pi^0$ in this analysis. We use the charged track information inclusively, except for that of the tracks in the $\pi^+\pi^-\pi^0$. In this section, we describe the algorithm to determine the flavor of $B_{\rm tag}$ from the inclusive track information.

The flavor tagging algorithm used at Belle is called as multi-dimensional likelihood (MDLH) method [45]. The flavor of the $B_{\rm tag}$ is determined based on the charge information of the following characteristic final state particles (Fig. 4.7):

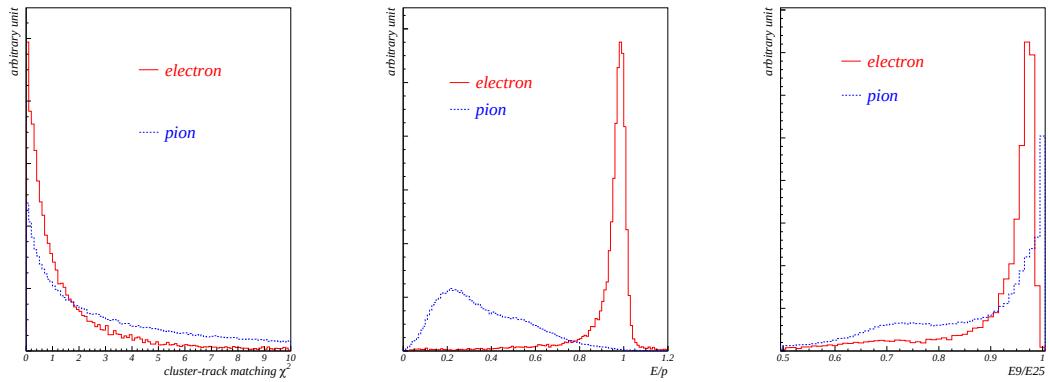


Figure 4.4: Cluster-track matching χ^2 (left), E/p (middle), and $E9/E25$ (right) distributions for electrons (solid line) and charged pions (broken line).

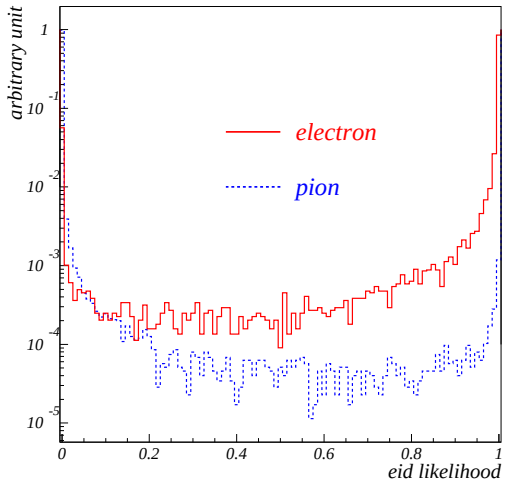


Figure 4.5: Likelihood ratio for the electron identification ($\mathcal{R}_{e/\overline{e}}$). Solid and broken histograms correspond to the electrons and charged pions, respectively.

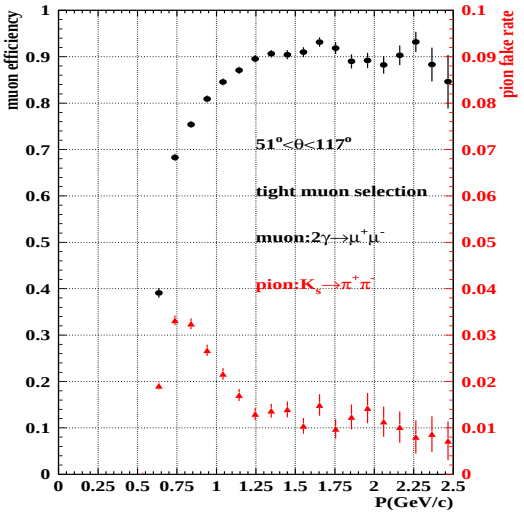


Figure 4.6: Muon efficiency and pion fake rate depending on the track momentum. Here, the criterion of $\mathcal{R}_{\mu/\pi,K} > 0.9$ is applied.

- high momentum leptons from $B \to Xl^{\pm}\nu$,
- intermediate momentum leptons from the cascade decay $B \to DX$, $D \to Kl^{\pm}\nu$,
- kaons from the cascade decay $B \to DX$, $D \to K^{\pm}Y$,
- high momentum pions from $B \to D^{(*)}\pi^{\pm}$ decays,
- slow pions from $B \to D^{*\pm}X$, $D^{*\pm} \to D\pi^{\pm}$, and
- Λ from the $b \to c \to s$ cascade decay.

The flavor tagging proceeds in two steps: track-level and event-level. In the track-level flavor tagging, the information of the tracks, such as the charge, momentum, and particle ID, are examined and the likelihood of the mother $B_{\rm tag}$ being B^0 or $\overline{B}^0$ is calculated for each of the track categories. We describe the flavor tagging information by $q_{\rm tag} \cdot r$. Here, $q_{\rm tag} = +1(-1)$ when $B_{\rm tag}$ is likely to be $B^0(\overline{B}^0)$ and r describes the confidence on the $q_{\rm tag}$ decision, where $r = 1$ when the decision is 100% confident and $r = 0$ for 0% confidence, i.e., $q_{\rm tag} = +1$ and $q_{\rm tag} = -1$ will be given randomly for the case of $r = 0$. In the event-level flavor tagging, we combine all the track-level likelihood and calculate event-level likelihood. The obtained event-level likelihood is calibrated using the data of the control sample and then used in the physics analysis.

2-1 Track-level Flavor Tagging

The track-level consists of four categories (slow pion, lambda, kaon, and lepton) of flavor tagging lookup tables, which are used to calculate the likelihood. Charged tracks which do not belong to $f_{CP} = \pi^+\pi^-\pi^0$ are used in the track-level tagging algorithms. These tracks are required to be associated with the interaction point (IP) except the one used in K^0_S or Λ recontruction. The minimum distance between the track and IP is required to be less than 2 cm x-y plane and 10 cm in z axis.

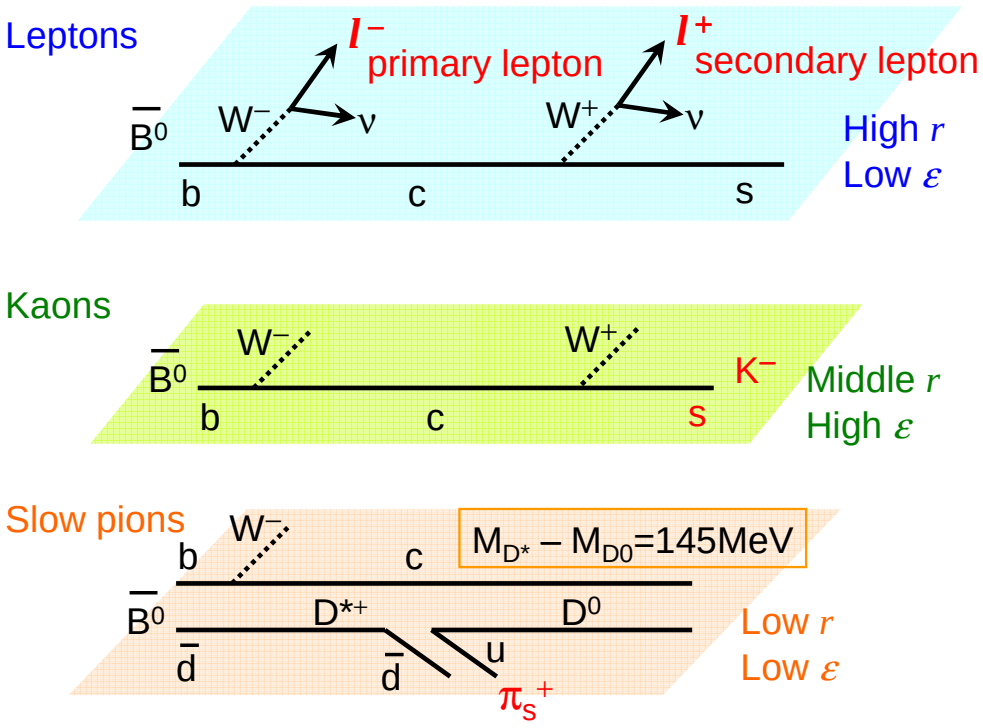


Figure 4.7: Conceptial figure of the characteristic charged tracks used for the flavor tagging. r and ϵ correspond to the quality and efficiency of each category, respectively.

Slow Pion Category A charged track which has a momentum $< 0.25\,\mathrm{GeV}/c$ in the center mass system (cms) is used for slow pion tags. These tracks cannot be kaon like, according to their particle identcation likelihood ratio, $\mathcal{R}_{K/\pi}$. The following variables are exploited to calculate the likelihood: the track's 1) charge, 2) momentum and 3) polar angle in the laboratory frame, 4) the angle between the slow pion and the thrust axis of rest of the tag side particles in the cms (α_{thr}), and 5) a pion/electron identcation ratio $\mathcal{R}_{\pi/e}$ from dE/dx measurements. The angle α_{thr} is useful to suppress the background from non-D^* decays. The ratio $\mathcal{R}_{\pi/e}$ provides additional power for removing background electrons from photon conversion.

Lambda Category A pair of oppositely charged tracks is reconstructed as a Λ candidate. One of the tracks should be identified as a proton. The Λ candidate is required to have an invariant mass, $M_{p\pi}$, in the range of 1.1108-1.1208 GeV/c^2 The angle between the candidate Λ momentum and the Λ flight direction (estimated by the IP and the candidate Λ vertex), θ_{defl}, should be less than 30°. The distance between the tracks in z axis at the Λ vertex position, Δz, should be smaller than 4 cm. The flight distance in x-y plane of the candidate Λ is required to be larger than 0.5 cm. The lookup table includes the flavor of Λ; $M_{p\pi}$, θ_{defl}, and Δz described above; and the presence of K^0_S candidates as the discriminant.

Kaon Category A charged track which is not positively idented as a lepton or a proton is included in this category. The flavor information from the kaon in the $b \to c \to s$ decay is the main concern. The fast pions from $B \to D^{(*)}\pi^{\pm}$ decays are also included[1]. The lookup table includes the following variables as the discriminant: the charge of the target track, the presence of K^0_S candidates, the momentum of the track in the cms, the polar angle of the track in the laboratory frame, and the $K^{\pm}/\pi^{\pm}$ identcation likelihood ratio ($\mathcal{R}_{K/\pi}$).

Lepton Category The lepton tag is the most powerful tagging method. High momentum electrons and muons from the $B \to X l^{\pm}\nu$ decays and the intermediate momentum leptons in the decay chain of $B \to DX$, $D \to K l^{\pm}\nu$ are considered. A charged track with a momentum greater than $0.4\,\mathrm{GeV}/c$ and electron likelihood larger than 0.8 is included as an electron candidate. A muon candidate is required to have a momentum greater than $0.8\,\mathrm{GeV}/c$ and muon likelihood larger than 0.95. The lepton momentum in the cms frame and the polar angle of the track in the laboratory frame are used in the discriminant as basic information. The lepton identcation likelihood is included for the selection of higher purity leptons. The lepton momentum can distinguish between the leptons from the primary $B \to X l^{\pm}\nu$ decays and those from the secondary $D \to K l^{\pm}\nu$ decays, which is important since the charges of the leptons are opposite for the primary and secondary from the same flavor of B_{tag}. The variables of recoil mass and missing momentum also provide information about these two types of decays; the recoil mass may indicate the presence of D mesons, and the missing momentum indicates the momenta of neutrinos.

All the output are calculated based on the lookup tables prepared with MC. The MC events distribute over the lookup tables, where the number of B^0 events ($N(B^0)$) and that of $\overline{B}^0$ events ($N(\overline{B}^0)$) are defined for each bin. Based on the prepared lookup table, the likelihood for a track being from B^0 ($\mathcal{L}(B^0)$) and $\overline{B}^0$ ($\mathcal{L}(\overline{B}^0)$) are calculated by the $N(B^0)$ and $N(\overline{B}^0)$ of the lookup-table bin where the track is located, as

$$\mathcal{L}(B^0) = \frac{N(B^0)}{N(B^0) + N(\overline{B}^0)}\,, \tag{4.4}$$

[1]This is achieved automatically, since we do not apply a cut criterion on the particle ID information, $\mathcal{R}_{K/\pi}$, but use it as a discriminant in the lookup table. In this manner, kaons and pions are treated together without discrete distinction.

$$\mathcal{L}(\overline{B}^0) = 1 - \mathcal{L}(B^0) \ . \tag{4.5}$$

The $(q_{\rm tag} \cdot r)$ as the output of the track-level flavor tagging is calculated from the likelihood as

$$q_{\rm tag} \cdot r = \frac{\mathcal{L}(B^0) - \mathcal{L}(\overline{B}^0)}{\mathcal{L}(B^0) + \mathcal{L}(\overline{B}^0)} \ . \tag{4.6}$$

2-2 Event-level Flavor Tagging

The flavor tagging information at the track-level are combined into a single $q_{\rm tag}$ and r for each event, as shown in Fig. 4.8. In the lepton and slow pion categories, the tracks with the highest r value is chosen as an input for the event-level flavor tagging. On the other hand, kaon and Λ categories are combined by taking a product of the likelihood for all the tracks and Λ candidates in these categories and the product is used to calculate $(q_{\rm tag} \cdot r)$ as the input for the event-level flavor tagging. This strategy gives a better result than choosing the one track or candidate with the highest r value. A three dimensional event-level lookup table is prepared with those three $(q_{\rm tag} \cdot r)$ values as the input. The likelihood and the resultant event-level $(q_{\rm tag} \cdot r)$ are calculated from the lookup table in the same manner as the track-level flavor tagging.

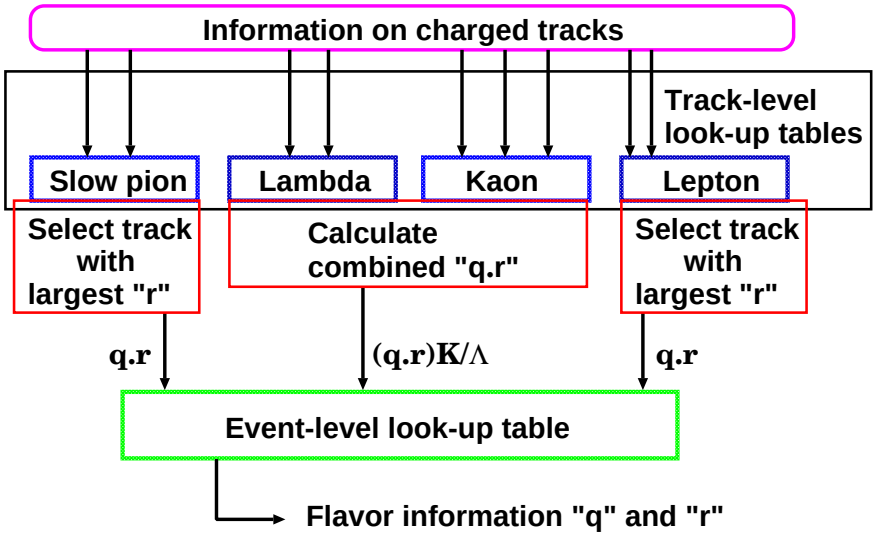


Figure 4.8: A schematic diagram of the flavor tagging algorithm.

2-3 Calibration and the Resultant Performance

Since the lookup tables used above are all based on MC, the performance, or the fraction of wrong-tag, has to be calibrated with data. We describe our method to calibrate it with data in the followings.

The wrong-tag effect enters in our observation as follows. Provided that we have a true Δt-$q_{\rm tag}$ PDF in general, $P_t(\Delta t, q_{\rm tag})$, the distribution we observe, $P(\Delta t, q_{\rm tag})$, is diluted by the wrong-tag effect as

$$P(\Delta t, q_{\rm tag}) = (1-w)P_t(\Delta t, q_{\rm tag}) + wP_t(\Delta t, \overline{q}_{\rm tag}) \ , \tag{4.7}$$

where w is the wrong-tag fraction[2] and $\overline{q}_{\rm tag} = -1(+1)$ for $q_{\rm tag} = +1(-1)$. For example, in the case of CP-violation measurement, the $P_t(\Delta t, q_{\rm tag})$ is given by Eq. (2.92) of chapter 2 as

$$P_t(\Delta t, q_{\rm tag}) = \frac{1}{4\tau_{B^0}} e^{-|\Delta t|/\tau_{B^0}} \left\{ 1 + q_{\rm tag} \cdot \left[\mathcal{A}_{f_{CP}} \cos(\Delta m_d \Delta t) + \mathcal{S}_{f_{CP}} \sin(\Delta m_d \Delta t) \right] \right\} , \quad (4.8)$$

where $\tau_{B^0} = 1/\Gamma$ and overall constant factor is determined to make the PDF normalized to be unity. Then, the distribution with the wrong-tag effect taken into account is

$$\begin{aligned} &P(\Delta t, q_{\rm tag}) \\ &= \frac{1}{4\tau_{B^0}} e^{-|\Delta t|/\tau_{B^0}} \left\{ 1 + q_{\rm tag}(1-2w) \left[\mathcal{A}_{f_{CP}} \cos(\Delta m_d \Delta t) + \mathcal{S}_{f_{CP}} \sin(\Delta m_d \Delta t) \right] \right\} . \end{aligned} \quad (4.9)$$

In practice, the situation is more complex. First, in order to make full use of the statistical power of the output from the flavor tagging, we introduce six regions of the tagging quality r: $0 < r \leq 0.25$, $0.25 < r \leq 0.5$, $0.5 < r \leq 0.625$, $0.625 < r \leq 0.75$, $0.75 < r \leq 0.875$, and $0.875 < r \leq 1.0$. We treat the events in the different regions separately. The wrong-tag fraction has to be defined for each region as w_l, where $l = 1, 2, \cdots, 6$ is the index over the r regions. Secondly, the wrong-tag fraction can be different for B^0 and $\overline{B}^0$[3]. With the wrong-tag probability for $B^0(\overline{B}^0)$ decay defined as $w_l^+(w_l^-)$, Eq. (4.7) is rewritten as

$$\begin{aligned} P(\Delta t, q_{\rm tag} = +1) &= (1 - w_l^+) P_t(\Delta t, +1) + w_l^- P_t(\Delta t, -1) , \\ P(\Delta t, q_{\rm tag} = -1) &= (1 - w_l^-) P_t(\Delta t, -1) + w_l^+ P_t(\Delta t, +1) , \end{aligned} \quad (4.10)$$

and Eq. (4.9) is

$$\begin{aligned} &P(\Delta t, q_{\rm tag}) = P(\Delta t, q_{\rm tag}, l) \\ &= \frac{1}{4\tau_{B^0}} e^{-|\Delta t|/\tau_{B^0}} \left\{ 1 - q_{\rm tag} \Delta w_l + q_{\rm tag}(1-2w_l) \left[\mathcal{A}_{f_{CP}} \cos(\Delta m_d \Delta t) + \mathcal{S}_{f_{CP}} \sin(\Delta m_d \Delta t) \right] \right\} , \end{aligned} \quad (4.11)$$

where $\Delta w_l \equiv w_l^+ - w_l^-$ and $w_l \equiv (w_l^+ + w_l^-)/2$. Here, Δw_l is called wrong-tag fraction difference and its absolute value is much smaller than w_l. The 12 values in total, w_l and Δw_l, are the parameters to be calibrated.

Another useful example is the case to observe B^0-$\overline{B}^0$ mixing. In this case, instead of f_{CP} we measure the decay chain of $\Upsilon(4S) \to B^0\overline{B}^0 \to f_{\rm flv} f_{\rm tag}$, where $f_{\rm flv}$ is a final state of flavor eigenstate for B^0 or $\overline{B}^0$. In this case, the $P_t(\Delta t, q_{\rm tag})$ without the wrong-tag effect is

$$P_t(\Delta t, q_{\rm tag}) = P_t(\Delta t, q_{\rm flv}, q_{\rm tag}) = \frac{1}{8\tau_{B^0}} e^{-|\Delta t|/\tau_{B^0}} \left\{ 1 - q_{\rm flv} q_{\rm tag} \cos(\Delta m_d \Delta t) \right\} , \quad (4.12)$$

where $q_{\rm flv} = +1(-1)$ when $f_{\rm flv}$ is the flavor eigenstate of $B^0(\overline{B}^0)$. The distribution we observe in this case is

$$\begin{aligned} P(\Delta t, q_{\rm tag}) &= P(\Delta t, q_{\rm rec}, q_{\rm tag}, l) \\ &= \frac{1}{8\tau_{B^0}} e^{-|\Delta t|/\tau_{B^0}} \left\{ 1 - q_{\rm tag} \Delta w_l - q_{\rm flv} q_{\rm tag}(1-2w_l) \cos(\Delta m_d \Delta t) \right\} . \end{aligned} \quad (4.13)$$

The key observation here is that the amplitudes of the terms proportional to $e^{-|\Delta t|/\tau_{B^0}}$ and $e^{-|\Delta t|/\tau_{B^0}} \cos(\Delta m_d \Delta t)$ are only dependent on w_l and Δw_l. We exploit this fact to calibrate

[2] Note that there is a relation of $r = 1 - 2w$ by construction of r, if the MC were perfect.

[3] This effect arises from the fact that our detector is made of matter, not antimatter; the detection efficiency can be different for a particle and an antiparticle, such as K^+ and K^-. This potentially makes the asymmetry of the wrong-tag fraction between B^0 and $\overline{B}^0$.

w_l and Δw_l with data. Choosing the control sample events with final states of $D^{(*)-}\pi^+$, $D^{*-}\rho^+$, and $D^{*-}l^+\nu$ as the $f_{\rm flv}$, we perform a time-dependent fit and determine the w_l and Δw_l. Figure 4.9 shows the fit result. The vertical axis is $(N_{\rm OF} - N_{\rm SF})/(N_{\rm OF} + N_{\rm SF})$, where $N_{\rm OF(SF)}$ is the number of events with $q_{\rm flv}q_{\rm tag} = -1(+1)$ in each $|\Delta t|$ bin. As can be seen in the figure, the mixing amplitudes are larger for the bins with large r than those with small r, reflecting the fact that the events with large r have small wrong-tag fraction. Table 4.1 lists the parameters obtained by the fit, which we use in the time-dependent Dalitz plot analysis. Effective tagging efficiency is defined as

$$\epsilon_{\rm tag} \equiv \sum \mathcal{F}^l w_l \ , \tag{4.14}$$

where $\mathcal{F}^l$ is the event fraction for each tagging quality region l; we achieve $\epsilon_{\rm tag} \sim 30\%$.

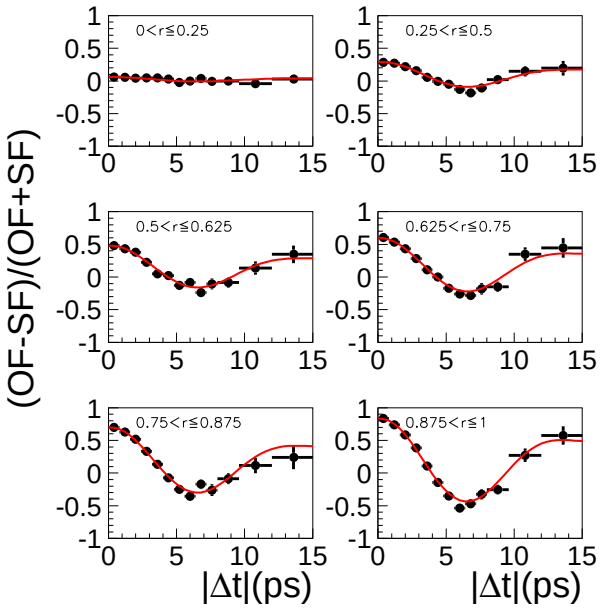


Figure 4.9: Fit result of the time-dependent fit to the B^0-$\overline{B}^0$ mixing. The plots from top-left to bottom-right correspond to the events with $l = 1, 2, \cdots, 6$, respectively. The amplitude of the oscillation is large in the region with large l, since the dilution effect is small there.

3 Proper-Time Difference Reconstruction

We need to measure the proper-time difference of two B meson decays, Δt, to observe the time-dependent CP asymmetry. Since the $\Upsilon(4S)$ is produced in a boosted system, where the boost factor is $(\beta\gamma)_\Upsilon = 0.425$, and the momenta of B's in the $\Upsilon(4S)$ rest frame are small, Δt is related to the position difference between the decay vertices of two B mesons, Δz, as

$$\Delta t = \frac{\Delta z}{c(\beta\gamma)_\Upsilon} \equiv \frac{z_{CP} - z_{\rm tag}}{c(\beta\gamma)_\Upsilon} \ . \tag{4.15}$$

Here z_{CP} and $z_{\rm tag}$ are the decay vertex positions of the B_{CP} (B decaying to $\pi^+\pi^-\pi^0$) and the $B_{\rm tag}$ (tag-side B), respectively. Described in this section is the method to reconstruct the vertex positions from the information of charged tracks and the interaction point (IP).

Table 4.1: Wrong tag fractions w_l and wrong tag fraction difference Δw_l obtained from the time-dependent fit to the B^0-$\overline{B}^0$ mixing. Most of the wrong tag fraction differences are consistent with zero.

DS-I

l	w_l	Δw_l
1	0.464 ± 0.006	-0.011 ± 0.006
2	0.331 ± 0.008	$+0.004 \pm 0.010$
3	0.231 ± 0.009	-0.011 ± 0.010
4	0.163 ± 0.008	-0.007 ± 0.009
5	0.109 ± 0.007	$+0.016 \pm 0.009$
6	0.020 ± 0.005	$+0.003 \pm 0.006$

DS-II

l	w_l	Δw_l
1	0.467 ± 0.006	$+0.005 \pm 0.007$
2	0.324 ± 0.007	-0.029 ± 0.009
3	0.223 ± 0.010	$+0.019 \pm 0.011$
4	0.160 ± 0.011	$+0.008 \pm 0.011$
5	0.101 ± 0.009	-0.022 ± 0.010
6	0.020 ± 0.006	$+0.003 \pm 0.006$

3-1 Interaction Point (IP) Profile

We use the constraint of interaction point, which is the collision point of the electron and positron beams, to improve the efficiency and position resolution of the vertex reconstruction. With this information, we can reconstruct the vertex position even from a single charged track.

We treat the IP as profile, the distribution accumulated and averaged over a certain range of events. The IP profile is calculated from the following information:

Fill-by-fill (run-by-run) information The IP distribution measured by Belle detector accumulated over each beam fill (run)[4]. This information is used to determine the rotation of IP with respect to the detector coordinate, $(\theta_x, \theta_y, \theta_z)$, and the detector-measured size of IP in the rotated coordinate, $(\sigma_{x'}^{\rm mes}, \sigma_{y'}^{\rm mes}, \sigma_{z'}^{\rm mes})$. The typical scale of the size is $\sim (100\,\mu{\rm m}, 70\,\mu{\rm m}, 3\,{\rm mm})$.

Event-by-event information The IP distribution measured by Belle detector accumulated over each 10,000 events, which is usually smaller than the number of events for each run. This information is used to obtain the mean position of the IP: (μ_x, μ_y, μ_z).

KEKB accelerator information The run-by-run beam size information offered by KEKB accelerator, from which we calculate the KEKB-measured size of IP in (x, y) direction only, $(\sigma_x^{\rm acc}, \sigma_y^{\rm acc})$, whose typical size is $(70\,\mu{\rm m}, 5\,\mu{\rm m})$. Note that the cross-section of the beams are designed to be elliptical and thus $\sigma_x^{\rm acc} \gg \sigma_y^{\rm acc}$.

For the first two items measured with Belle detector, we use hadronic events, which consists of $e^+e^- \to q\overline{q}$ ($q = u, d, s, c$) continuum events and $e^+e^- \to B\overline{B}$ events. Since the continuum events have dominant contribution, the tracks are considered to be coming from the primary vertices of the e^+e^- collision. The detector-measured size is affected and smeared by detector resolution.

From the above measurements, we calculate the profile of IP as follows. For the mean position and the rotation, we use the measured (μ_x, μ_y, μ_z) and $(\theta_x, \theta_y, \theta_z)$ as they are. For the size, on the other hand, we use following values calculated based on the measured information:

$$\begin{aligned} \sigma_{x'} &= \sqrt{(\sigma_{x'}^{\rm mes})^2 - (\sigma_{y'}^{\rm mes})^2}\ , \\ \sigma_{y'} &= \sigma_y^{\rm acc}\ , \\ \sigma_{z'} &= \sigma_{z'}^{\rm mes}\ . \end{aligned} \tag{4.16}$$

[4]Until 2002, this information of IP profile was calculated for each beam fill. Since 2003, having large enough luminosity, it has been calculated for each run, which is smaller unit for the region of events than the fill. (A fill consists of several runs.) Note that now we do not have the concept of fill, since KEKB is operated in the continuous injection mode.

The reason we obtain $\sigma_{x'}$ by this relation is that $\sigma_{y'}^{\rm mes}$ can be considered to be detector resolution since $\sigma_{y'}^{\rm mes} \gg \sigma_{y}^{\rm acc}$. Note that the detector resolution in z direction ($\sim 0.1\,$mm) is much better than $\sigma_{z'}^{\rm mes} \sim 3\,$mm and thus we can use $\sigma_{z'}^{\rm mes}$ for $\sigma_{z'}$. The typical size of the $(\sigma_{x'}, \sigma_{y'}, \sigma_{z'})$ is $(70\,\mu$m$, 5\,\mu$m$, 3\,$mm$)$. In the physics analysis, we convolve the obtained $(\sigma_{x'}, \sigma_{y'}, \sigma_{y'})$ with 1) the uncertainty of (μ_x, μ_y, μ_z) measurement and 2) B flight effect, which we estimate to be $21\,\mu$m in x and y directions; and use it for the vertex reconstruction.

3-2 Vertex Reconstruction of B_{CP}

The vertex position of $B_{CP} \to \pi^+\pi^-\pi^0$ decay is reconstructed using two charged tracks of the π^+ and π^-, and IP. We require the charged tracks to have sufficient number of SVD-hits associated; a 2-D hit and another z hit is the least requirement in DS-I, while two 2-D hits is at least required in DS-II. Here, a 2-D hit is a set of r-ϕ hit and z hit in a single SVD layer associated with the same charged track. Only the tracks satisfying this requirement are used for the vertex reconstruction; the vertex reconstruction fails when both of the two charged tracks in an event does not satisfy the requirement.

3-3 Vertex Reconstruction of $B_{\rm tag}$

The decay vertex of the $B_{\rm tag}$ is determined inclusively from the tracks that are not assigned to B_{CP} and satisfy the requirement of SVD-hits, where the SVD-hit requirement is the same as that for B_{CP} vertex reconstruction. Further selection criteria are carefully chosen and required to minimize the effect of poorly-reconstructed tracks and long-lived particles, such as D mesons and K_S^0. The effect of the secondary charm decay moves the decay vertex position of the $B_{\rm tag}$ toward charm flight direction. It also significantly degrades the vertex resolution. For this reason, the resolution of $\Delta z \equiv z_{CP} - z_{\rm tag}$ measurement is dominated by that of tag-side vertex, $z_{\rm tag}$. The followings are the requirements to the tracks:

- The track must be associated with sufficient number of SVD-hits (the same requirement as CP-side).
- The estimated track error in z direction, σ_z, must be less than $500\,\mu$m to remove poorly reconstructed tracks.
- Tracks from K_S^0 candidates are rejected; a pair of tracks is rejected when the invariant mass of the two tracks, $m_{\pi\pi}$, satisfies $|m_{\pi\pi} - m_{K_S^0}| < 15\,{\rm MeV}/c^2$, where $m_{K_S^0}$ is K_S^0 mass.
- Tracks with impact parameter to the CP-side vertex position in r-ϕ plane (δr) greater than $500\,\mu$m are rejected for the further reduction of the K_S^0 daughter.

With the tracks that satisfy the requirements and IP constraint, we reconstruct the $B_{\rm tag}$ vertex. After the vertex reconstruction, we examine the reduced χ^2 defined as χ^2/n, where n is the degree of freedom. If the reduced χ^2 of the vertex fit is larger than 20, the track giving the largest contribution to the reduced χ^2 is removed. This rule has an exception; when the track with the largest χ^2 contribution is a lepton with a momentum greater than 1.1 GeV in the cms, the track is kept and the track with second largest χ^2 contribution is removed. This is because a lepton with high momentum is likely to come from a primary semileptonic B decays[5]. We perform this procedure iteratively until the reduced χ^2 gets less than 20 or a single track is left.

In the DS-II, the $B_{\rm tag}$ vertex reconstruction has another step. In the case where the resultant vertex is reconstructed from a single track and IP constraint after the above procedure,

[5]Note that we intend to remove the tracks from secondary decays by the removal of the tracks with large χ^2 contributions. The high momentum leptons are likely to originate from the primary vertices of the processes such as $B^0 \to D^{(*)-}l^+\nu$.

we go back to the beginning of the procedure and redo the same procedure with modified requirements to the tracks. The only modification is the requirement of SVD-hits; here, we require the tracks to have two 2-D hits in 2nd, 3rd, or 4th layer, i.e., we ignore the 1st-layer hit in the examination of the SVD-hit requirement. There is a fact that the single track used for the vertex reconstruction tends to be poorly reconstructed when the track is associated with only two SVD-hits and one of them is in the 1st layer. This is because the 1st layer has smaller radius in DS-II than DS-I and thus suffers from more fake hits due to beam background. In this condition, the two-hits requirement to reject the poorly reconstructed tracks wrongly associated with SVD-hits does not work as intended. The redoing procedure described above works well to remove such poorly reconstructed tracks with little reduction of the vertex-reconstruction efficiency.

3-4 Requirement to the Quality and Performance

We examine the qualities of the reconstructed vertices with following variable

$$\xi \equiv \frac{1}{2n_{\rm trk}} \sum_{i}^{n_{\rm trk}} \frac{(z_{\rm fit} - z^i_{\rm trk})^2}{{\sigma^i_z}^2} \ , \tag{4.17}$$

where $n_{\rm trk}$, i, $z_{\rm fit}$, $z^i_{\rm trk}$, and σ^i_z are the number of tracks used for the vertex reconstruction, the index over the tracks, the z position of the fitted vertex, the z position of i-th track at the closest approach to the vertex, and the estimated error of the i-th track in z direction, respectively. This is similar to usual χ^2 but only z direction is used. This is because the usual χ^2 is correlated with B flight length and thus with Δz, since IP constraint is strong in x-y direction. Consequently, use of the usual χ^2 can cause a possible bias in Δz measurement and thus we use the ξ defined above, which has no correlation with Δz. We require $\xi < 100$ for both of the vertices of B_{CP} and $B_{\rm tag}$.

The efficiencies of the vertex reconstruction in the $B^0 \to \pi^+\pi^-\pi^0$ decay process are 91%, 88%, and 86% for B_{CP}, $B_{\rm tag}$, and both of them, respectively, which are estimated using MC. Note that the efficiency for the both vertex reconstructions to success is not a product of the efficiencies for those of B_{CP} and $B_{\rm tag}$, indicating that when one of the two vertices in an event is poorly reconstructed, another also tends to be poorly reconstructed. To further remove the events with poorly reconstructed vertices, we apply a criterion of $|\Delta t| < 70\,{\rm ps}$. The efficiency of this cut is 99.8%.

We evaluate the performance of the vertex reconstruction using MC as shown in Fig. 4.10, where the reconstructed position and the true position are compared. In DS-I (DS-II), typical resolutions in rms are $80\,\mu$m, $150\,\mu$m, and $170\,\mu$m ($70\,\mu$m, $140\,\mu$m, and $150\,\mu$m) for z_{CP}, $z_{\rm tag}$, and Δz, respectively. Corresponding Δt resolution is 1.3 ps (1.2 ps).

Since the detector resolution is comparable to the lifetime of B^0 and the period of B^0-$\overline{B}^0$ oscillation, the smearing effect due to the resolution has to be taken into account in the time-dependent CP-violation measurement. We treat this effect as resolution function, $R(\delta\Delta t)$, where $\delta\Delta t$ is the difference between the measured Δt and true value of it. We perform a detailed study of the resolution function using a control sample data of the decay modes $B^0 \to D^{*-}l^+\nu$, $D^{(*)-}\pi^+$, $D^{*-}\rho^+$, $B^+ \to D^0\pi^+$, and $J/\psi K^+$ [46]. As the PDF for the maximum likelihood fit, $P_{\rm fit}(\Delta t)$, we use a convolution of the resolution function and the ideal distribution without the resolution effect, $P_{\rm ideal}(\Delta t)$:

$$P_{\rm fit}(\Delta t) = [R \otimes P_{\rm ideal}](\Delta t) \ . \tag{4.18}$$

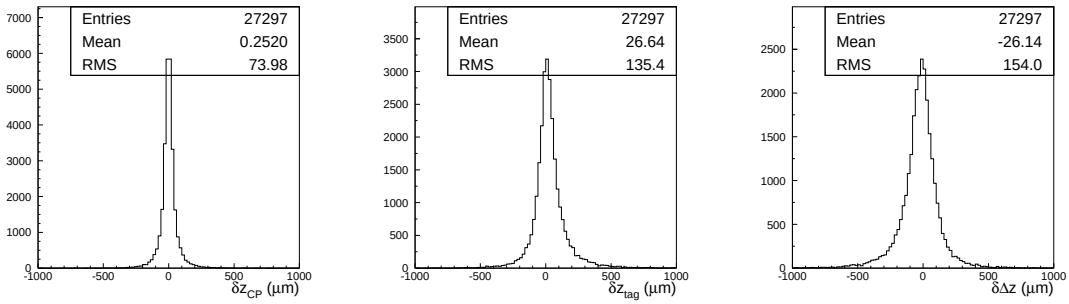


Figure 4.10: The distributions of δz_{CP} (left), $\delta z_{\rm tag}$ (middle), and $\delta\Delta z$ (right) in DS-II, where δz_{CP}, $\delta z_{\rm tag}$, and $\delta\Delta z$ are the difference between the reconstructed and generated (reconstructed − generated) values for z_{CP}, $z_{\rm tag}$, and Δz, respectively. In the $\delta z_{\rm tag}$ distribution, there is a bias of $+25\,\mu$m, which is due to the effect of the secondary vertex of D mesons. This leads to the bias of $-25\,\mu$m in the $\delta\Delta z$ distribution.

4 Techniques Dedicated to $B^0 \to \pi^+\pi^-\pi^0$ Dalitz Analysis

4-1 Square Dalitz plot (SDP)

The signal and the continuum background $e^+e^- \to q\overline{q}(q = u, d, s, c)$, which is the dominant background in this analysis, populate the kinematic boundaries of the usual Dalitz plot as shown in Figs. 4.11 and 4.12. Since we model part of the Dalitz plot PDF's with binned histograms, the distribution concentrated in a narrow region is not easy to treat. We therefore apply the transformation

$$ds_+ ds_- \to |\det \boldsymbol{J}| dm' d\theta' \ , \tag{4.19}$$

which defines the square Dalitz plot (SDP) [47, 48]. The new coordinates are

$$m' \equiv \frac{1}{\pi} \arccos\left(2\frac{m_0 - m_0^{\rm min}}{m_0^{\rm max} - m_0^{\rm min}} - 1 \right) \ , \tag{4.20}$$

$$\theta' \equiv \frac{1}{\pi}\theta_0 \quad \left(= \frac{1}{\pi}\theta^{-0} \right) \ . \tag{4.21}$$

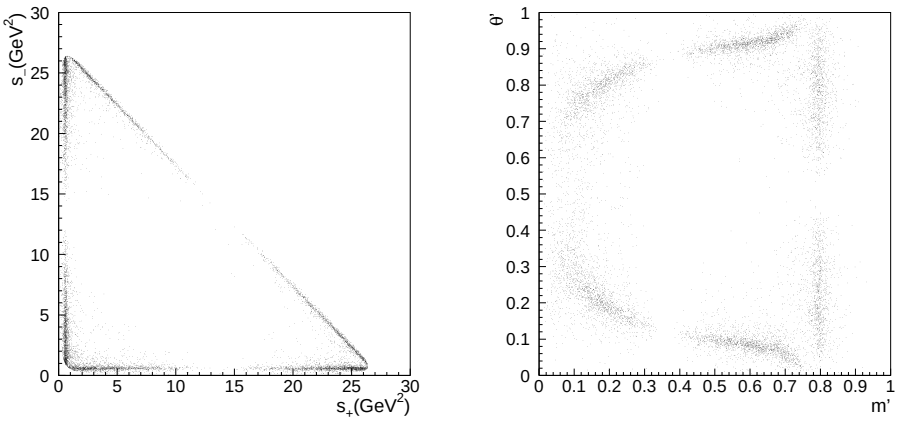
Here, $m_0 = \sqrt{s_0}$ and θ_0 are the mass and the helicity angle of ρ^0 (or $\pi^+\pi^-$), respectively; $m_0^{\rm max} \equiv m_{B^0} - m_{\pi^0}$ and $m_0^{\rm min} \equiv 2m_{\pi^+}$ are the kinematic limits of m_0, and $\boldsymbol{J}$ is the Jacobian of the transformation. The determinant of the Jacobian is given by

$$|\det \boldsymbol{J}| = 4|\vec{p}_+||\vec{p}_0| m_0 \cdot \frac{m_0^{\rm max} - m_0^{\rm min}}{2} \pi \sin(\pi m') \cdot \pi \sin(\pi\theta') \ , \tag{4.22}$$

where $\vec{p}_+$ and $\vec{p}_0$ are the three momenta of π^+ and π^0 in the $\pi^+\pi^-$ rest frame. The detail of the parameter transformation and some useful relations can be found in appendix C.

4-2 Parameterization

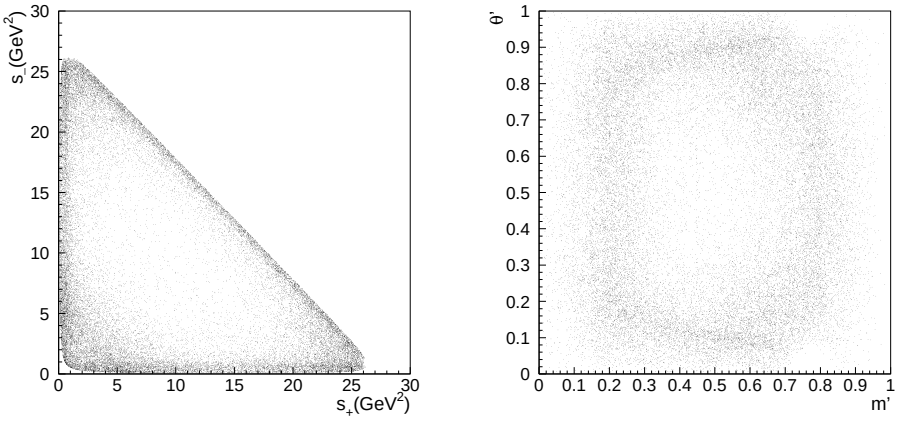
We parameterize the coefficients of equations (2.162) and (2.163), which are the parameters to be determined by this analysis, not by the complex amplitudes $\overline{A}^k$ directly but by the



(a) Usual Dalitz plot

(b) Square Dalitz plot

Figure 4.11: Distribution of signal Monte Carlo (without detector efficiency and smearing) in the Dalitz plot.



(a) Usual Dalitz plot

(b) Square Dalitz plot

Figure 4.12: Distribution of $q\overline{q}$ background (from the data $M_{\rm bc}$ sideband) in the Dalitz plot.

coefficients corresponding to their bilinear products as follows

$$\begin{aligned}
&|A_{3\pi}(s_+,s_-)|^2 \pm |\overline{A}_{3\pi}(s_+,s_-)|^2 \\
&\quad\propto \sum_{\kappa\in\{+,-,0\}} |f_\kappa|^2 U_\kappa^{\pm} + 2 \sum_{\kappa<\sigma\in\{+,-,0\}} \left(\mathrm{Re}[f_\kappa f_\sigma^*]\, U_{\kappa\sigma}^{\pm,\mathrm{Re}} - \mathrm{Im}[f_\kappa f_\sigma^*]\, U_{\kappa\sigma}^{\pm,\mathrm{Im}} \right) ,
\end{aligned} \tag{4.23}$$

$$\begin{aligned}
&\mathrm{Im}\left(\frac{q}{p} \overline{A}_{3\pi}(s_+,s_-) A_{3\pi}(s_+,s_-)^* \right) \\
&\quad\propto \sum_{\kappa\in\{+,-,0\}} |f_\kappa|^2 I_\kappa + 2 \sum_{\kappa<\sigma\in\{+,-,0\}} \left(\mathrm{Re}[f_\kappa f_\sigma^*]\, I_{\kappa\sigma}^{\mathrm{Im}} + \mathrm{Im}[f_\kappa f_\sigma^*]\, I_{\kappa\sigma}^{\mathrm{Re}} \right) ,
\end{aligned} \tag{4.24}$$

with the coefficients being related to the amplitudes as

$$U_\kappa^{\pm} = \left(|A^\kappa|^2 \pm |\overline{A}^\kappa|^2 \right) / N , \tag{4.25}$$

$$I_\kappa = \mathrm{Im}\left[\overline{A}^\kappa A^{\kappa *} \right] / N , \tag{4.26}$$

$$U_{\kappa\sigma}^{\pm,\mathrm{Re(Im)}} = \mathrm{Re(Im)}\left[A^\kappa A^{\sigma *} \pm \overline{A}^\kappa \overline{A}^{\sigma *} \right] / N , \tag{4.27}$$

$$I_{\kappa\sigma}^{\mathrm{Re(Im)}} = \mathrm{Re(Im)}\left[\overline{A}^\kappa A^{\sigma *} - (+)\, \overline{A}^\sigma A^{\kappa *} \right] / N , \tag{4.28}$$

where N is an overall normalization factor. The 27 coefficients (4.25)-(4.28) are the parameters determined by the fit [49]. The parameters (4.25)-(4.26) and (4.27)-(4.28) are called non-interfering and interfering parameters, respectively.

This parameterization makes the fit well behaved, with the fit parameters being Gaussian distributed and having no local minimum. The Dalitz-Δt distribution of signal is spanned by 27 linearly independent basis functions: there are nine linearly independent basis functions in Dalitz plot dimension as shown in the Fig. 4.13 and three linearly independent basis functions in Δt-q_{tag} dimension, which are

$$e^{-|\Delta t|/\tau_{B^0}} , \quad q_{\mathrm{tag}} \cdot e^{-|\Delta t|/\tau_{B^0}} \cos(\Delta m_d \Delta t) , \quad \text{and} \quad q_{\mathrm{tag}} \cdot e^{-|\Delta t|/\tau_{B^0}} \sin(\Delta m_d \Delta t) . \tag{4.29}$$

All the combination products of the Dalitz plot and Δt basis functions are the basis functions for Dalitz-Δt, whose number is 27. The above parameterization is to parameterize the signal distribution by the linear combination of the 27 basis functions with their coefficients as fit parameters, and thus the fit is well behaved. Since the overall normalization is arbitrary, we fix it by requiring $U_+^+ = 1$, i.e., we take $N = |A^+|^2 + |\overline{A}^+|^2$ as the normalization.

All these are analogous to the case of usual time-dependent CP violation measurement. In a usual analysis without the Dalitz plot dependence, the time-dependent decay width of equation (2.88) is reparameterized as

$$d\Gamma / d\Delta t \propto e^{-\Gamma|\Delta t|} \left[\mathcal{D} + q_{\mathrm{tag}} \cdot \mathcal{A} \cos(\Delta m_d \Delta t) + q_{\mathrm{tag}} \cdot \mathcal{S} \sin(\Delta m_d \Delta t) \right] , \tag{4.30}$$

with

$$\mathcal{D} = \left(|\overline{A}_{f_1}|^2 + |A_{f_1}|^2 \right) / N , \tag{4.31}$$

$$\mathcal{A} = \left(|\overline{A}_{f_1}|^2 - |A_{f_1}|^2 \right) / N , \tag{4.32}$$

$$\mathcal{S} = 2\mathrm{Im}\left(\frac{q}{p} \overline{A}_{f_1} A_{f_1}^* \right) \Big/ N , \tag{4.33}$$

where N is an overall normalization. Then, to fix overall normalization we require $\mathcal{D} = 1$, i.e., take $N = |\overline{A}_{f_1}|^2 + |A_{f_1}|^2$, and obtain the usually used formalism of

$$d\Gamma / d\Delta t \propto e^{-\Gamma|\Delta t|} \left[1 + q_{\mathrm{tag}} \cdot \mathcal{A} \cos(\Delta m_d \Delta t) + q_{\mathrm{tag}} \cdot \mathcal{S} \sin(\Delta m_d \Delta t) \right] . \tag{4.34}$$

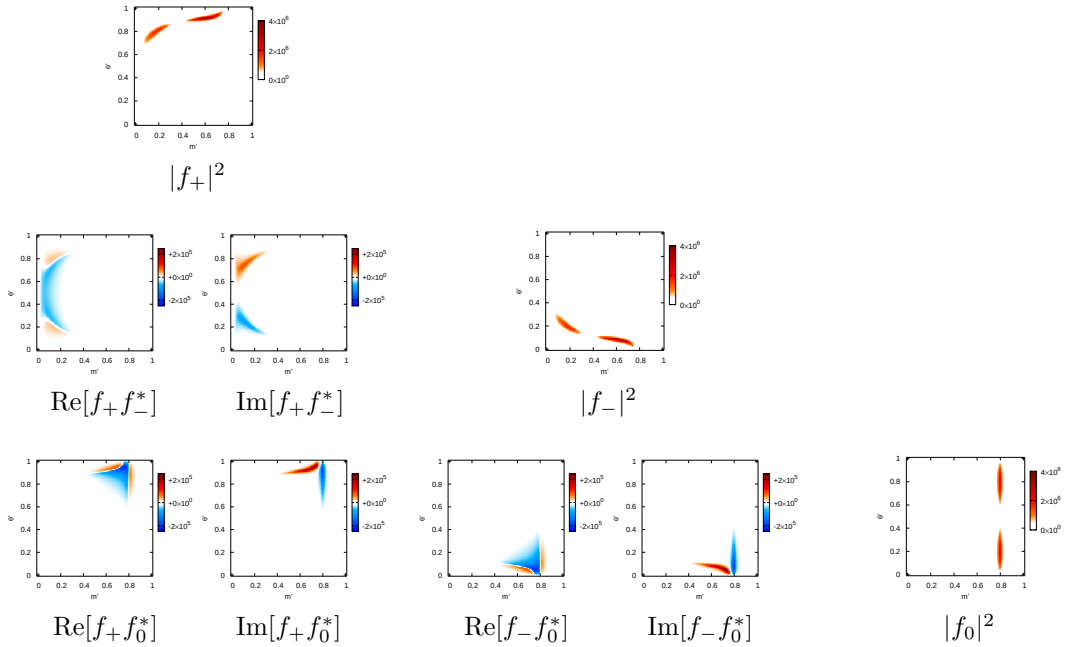


Figure 4.13: The nine basis functions of the signal Dalitz plot distribution in the square Dalitz plot. Only the contribution from $\rho(770)$ is assumed here.

Here, to use $\mathcal{D}$, $\mathcal{A}$, and $\mathcal{S}$ instead of A_{f_1} and $\frac{q}{p}\overline{A}_{f_1}$ corresponds to using the 27 coefficients instead of A^κ and $\overline{A}^\kappa$; and the requirement of $\mathcal{D} = 1$ corresponds to the requirement of $U_+^+ = 1$.

Chapter 5

Event Selection and Signal Extraction

1 Data Set

The analysis presented in this thesis is based on the data sample taken from January 2000 to June 2005. The integrated luminosity for each day and the history of the total accumulated luminosity are shown in Fig. 5.1. The total integrated luminosity of $461\,\mathrm{fb}^{-1}$ has been accumulated in this period, among which $414\,\mathrm{fb}^{-1}$ and $47\,\mathrm{fb}^{-1}$ are taken on and 50 MeV-below the $\Upsilon(4S)$ resonance, respectively. The data sample contains $449 \times 10^6 B\overline{B}$ pairs.

2 Event Selection

2-1 Reconstruction of $B^0 \to \pi^+\pi^-\pi^0$ Candidates

To reconstruct candidate $B^0 \to \pi^+\pi^-\pi^0$ decays, charged tracks reconstructed with the CDC and SVD are required to originate from the interaction point (IP):

$$|dr| < 0.1\,\mathrm{cm}\,, \quad \text{and} \quad |dz| < 4\,\mathrm{cm}\,, \tag{5.1}$$

where dr and dz are transverse and longitudinal components of the tracks' impact parameters with respect to the IP, respectively. They are also required to have transverse momenta greater than 0.1 GeV/c. We distinguish charged kaons from pions based on the likelihood ratio $\mathcal{R}_{K/\pi}$ and require the tracks to be pion like:

$$\mathcal{R}_{K/\pi} < 0.4\,. \tag{5.2}$$

We reject tracks that are positively identified as electrons; in terms of the likelihood ratio $\mathcal{R}_{e/\overline{e}}$, tracks satisfying

$$\mathcal{R}_{e/\overline{e}} < 0.95 \tag{5.3}$$

are selected.

Photons are identified as isolated ECL clusters that are not matched to any charged track. We reconstruct π^0 candidates from pairs of photons detected in the barrel (endcap) ECL with $E_\gamma > 0.05$ (0.1) GeV, where E_γ is the photon energy measured with the ECL. Photon pairs with momenta greater than 0.1 GeV/c in the laboratory frame and with an invariant mass between 0.1178 GeV/c^2 to 0.1502 GeV/c^2, roughly corresponding to $\pm 3\sigma$ of the mass resolution, are used as π^0 candidates.

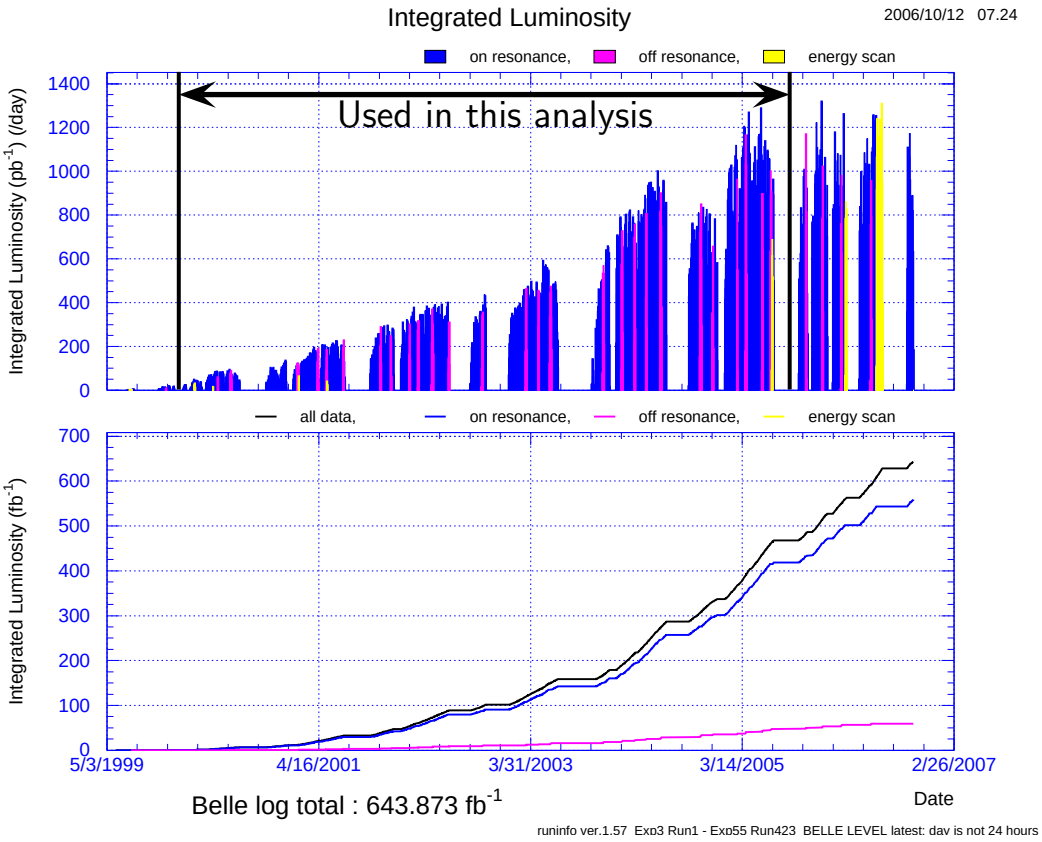


Figure 5.1: History of KEKB luminosity accumulated at Belle.

We identify B meson decays using the energy difference

$$\Delta E \equiv E_B^{\rm cms} - E_{\rm beam}^{\rm cms} \ , \tag{5.4}$$

and the beam-energy constrained mass

$$M_{\rm bc} \equiv \sqrt{(E_{\rm beam}^{\rm cms})^2 - (p_B^{\rm cms})^2} \ , \tag{5.5}$$

where $E_{\rm beam}^{\rm cms}$ is the beam energy in the center mass system (cms), and $E_B^{\rm cms}$ and $p_B^{\rm cms}$ are the cms energy and momentum of the reconstructed B candidate, respectively.

We select candidates in the large fitting region that is defined as $-0.2\,{\rm GeV} < \Delta E < 0.2\,{\rm GeV}$ and $5.2\,{\rm GeV}/c^2 < M_{\rm bc} < 5.3\,{\rm GeV}/c^2$, The large fitting region consists of the signal region defined as $-0.1\,{\rm GeV} < \Delta E < 0.08\,{\rm GeV}$ and $M_{\rm bc} > 5.27\,{\rm GeV}/c^2$, and the complement, called the sideband region, which is dominated by background events.

2-2 Vertexing and Flavor Tagging

We then apply the vertexing and the flavor tagging, detail of which are described in Sec. 2 and Sec. 3 of chapter 4.

2-3 Continuum Suppression Cut

The continuum events $e^+e^- \to q\overline{q}$ ($q = u, d, s, c$) is the dominant background in the $B \to \pi^+\pi^-\pi^0$ candidates. We reduce the background events by exploiting the event topology; we make use the fact that a continuum event is jet-like in the cms while a $B\overline{B}$ event has a spherical topology due to the large mass of the B meson. To characterize the event topology, we employ a modified version of Super Fox-Wolfram moment, so-called KSFW. The KSFW is an algorithm to form a Fisher discriminant from the Super Fox-Wolfram moments and missing mass. The distributions of the discriminants are shown in Fig. 5.3. More detailed description on the KSFW can be found eleswhere [50].

In addition to the KSFW, we also use the B flight direction with respect to the beam axis in cms, $\cos\theta_B^*$, to separate signal and continuum events. The B flight direction has the distribution with a dependence of $1-\cos^2\theta_B^*$, since the property of helicity conservation in the weak interaction vertex requires the helicities of the interacting e^+ and e^- to be opposite and thus the angular momentum of $B^0\overline{B}^0$ satisfies $L_z = \pm 1$. On the other hand, the continuum events have a uniform angular distribution[1]. Thus, the $\cos\theta_B^*$ works as a good discriminant. Figure 5.2 shows the distributions of signal and continuum components.

We use signal MC (sideband data) to determine the likelihood distribution of KSFW and $\cos\theta_B^*$ for the signal (continuum background) component. We approximate the likelihood function of $\cos\theta_B^*$ of the continuum component by a flat distribution, and that of signal component by fitting the MC-generated distribution with 2-nd order polynomial, the result of which is

$$\mathcal{L}^{\rm sig}_{\cos\theta_B^*} = 1.48 - 0.02 \cdot |\cos\theta_B^*| - 1.41 \cdot |\cos\theta_B^*|^2 \ . \tag{5.6}$$

We compose a likelihood ratio, KLR, using the likelihood distributions of KSFW and $\cos\theta_B^*$:

$$KLR \equiv \frac{\mathcal{L}^{\rm sig}_{\cos\theta_B^*}\mathcal{L}^{\rm sig}_{KSFW}}{\mathcal{L}^{\rm sig}_{\cos\theta_B^*}\mathcal{L}^{\rm sig}_{KSFW} + \mathcal{L}^{qq}_{\cos\theta_B^*}\mathcal{L}^{qq}_{KSFW}} \ . \tag{5.7}$$

[1]Note that $q\overline{q}$ from the $e^+e^- \to q\overline{q}$ themselves have angular dependence. However, fake B's reconstructed from the $q\overline{q}$ have uniform distribution, since they are random combinations of the particles from both of the q and $\overline{q}$.

Figure 5.4 shows the distributions of the KLR in each bins of the flavor tagging quality r. As can be seen there, the distribution is dependent on the r-bin.

We use the KLR to define a selection criterion to reduce the continuum background events; we reject the events with KLR smaller than cut values. To determine the criterion, we define the Figure of Merit (F.o.M.) as a function of the cut value of KLR:

$$\text{F.o.M.} \equiv \frac{N_{\text{sig}}}{\sqrt{N_{\text{sig}} + N_{q\overline{q}}}} , \tag{5.8}$$

where N_{sig} and $N_{q\overline{q}}$ are the expected numbers of events of signal and continuum components, respectively, for the given cut value. The cut criterion that gives the maximum F.o.M. is expected to be optimum, i.e., yields the largest statistical power. We calculate N_{sig} and $N_{q\overline{q}}$ using the likelihood distributions obtained above and assuming the branching fraction of 2.4×10^{-5} for the signal, and obtain the F.o.M. curve shown in the Fig. 5.5. Here we treat the different flavor tagging quality regions separately, since the distributions and signal to noise ratios are dependent on the r[2]. Using the F.o.M. curves, we determine the KLR cut value for each r-bin as listed in Table 5.1.

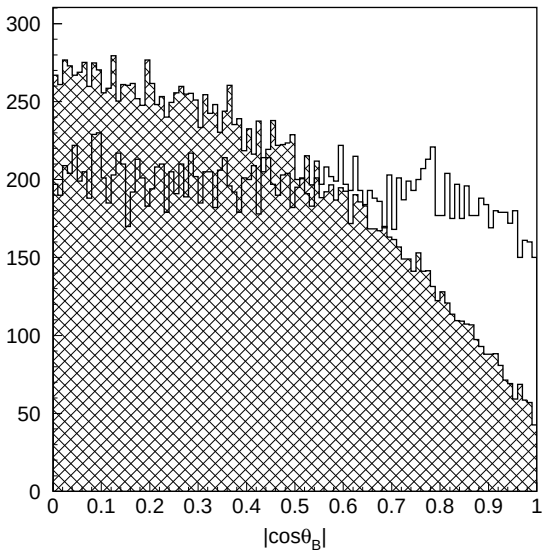


Figure 5.2: The distribution of $|cos\theta^*_B|$. Hatched and outlined histograms correspond to the signal MC and sideband data, respectively.

2-4 Best Candidate Selection

The reconstructed signal events contain substantial fraction of incorrectly reconstructed candidates, what we call self cross feed (SCF). Consequently, each event can have multiple $B \to \pi^+\pi^-\pi^0$ candidates. To select one candidate for each event, effectively suppressing the SCF, we apply best candidate selection.

Since the distribution of fitted π^0 mass and KLR are different for the correctly reconstructed signal and SCF as shown in Fig. 5.6, we use the two variables for the best candidate selection.

[2]In general, the signal to noise ratio is larger for the regions with large r than those with small r.

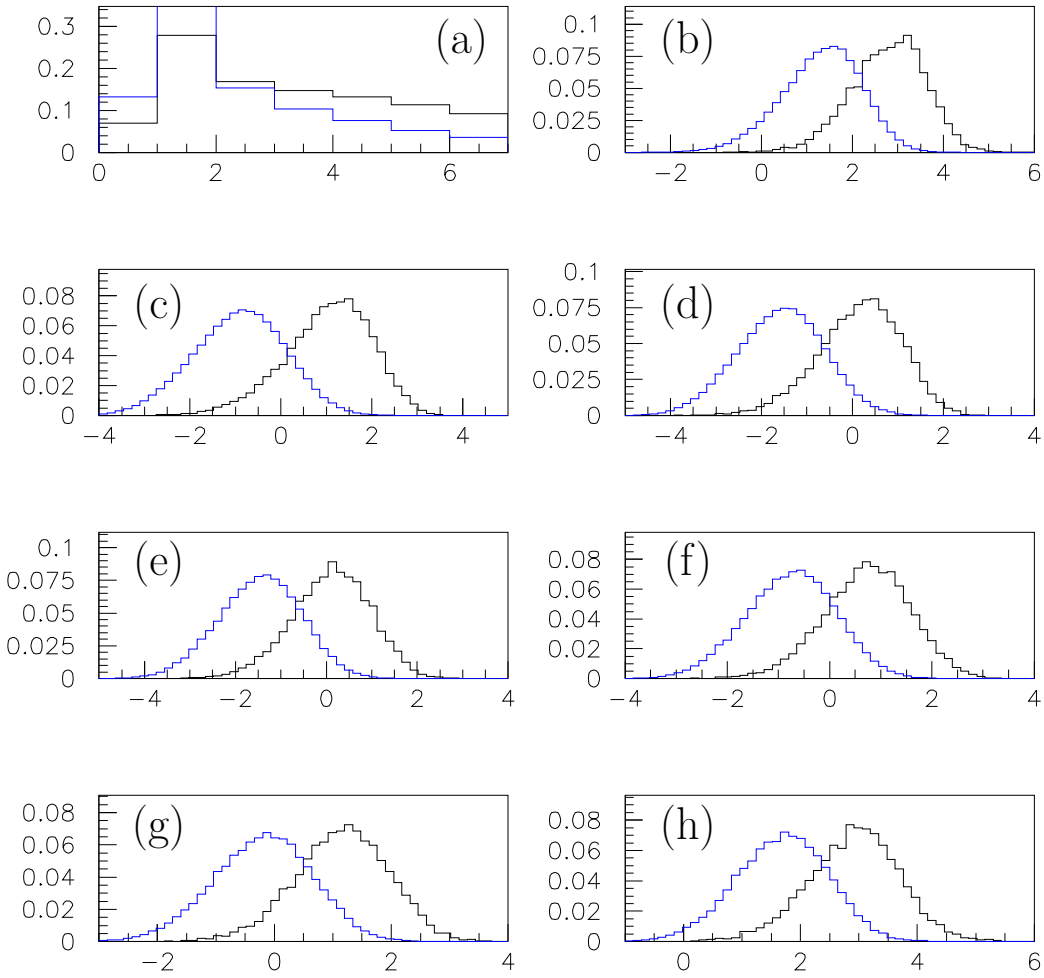


Figure 5.3: The distributions of the missing mass (a) and the Fisher discriminant (b)-(h) obtained from the KSFW algorithm for the signal (black) and continuum background (blue) events. Here, (b)-(h) correspond to the Fisher distributions for the different missing mass regions. Since the distribution depends on the missing mass, we treat the events in the different missing mass region separately.

Table 5.1: KLR cut value of each r-bin.

l (r-bin)	KLR cut value
1	0.95
2	0.92
3	0.90
4	0.90
5	0.87
6	0.30

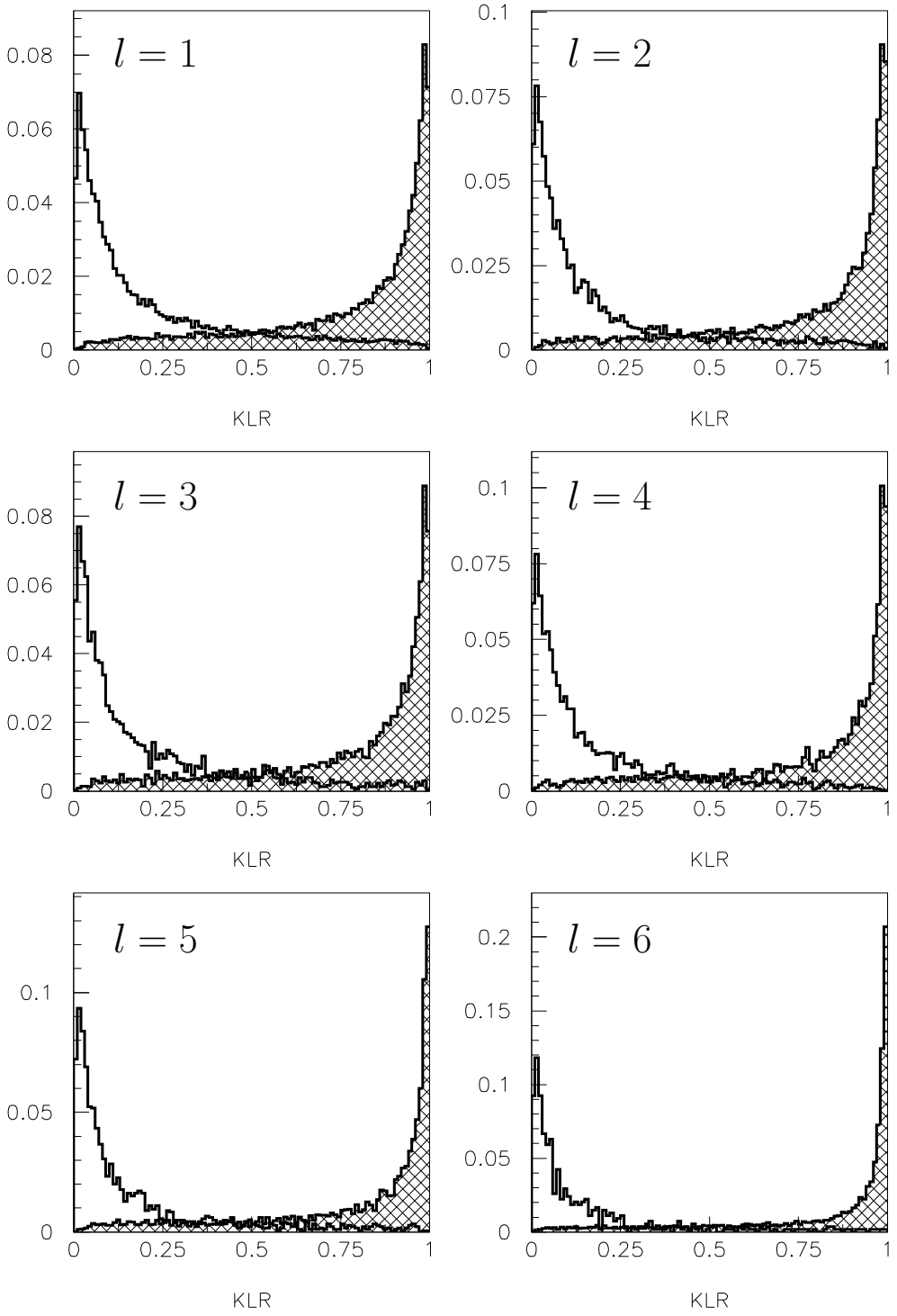


Figure 5.4: The KLR distribution for each r-bin. Hatched and outlined histograms correspond to the signal MC and sideband data, respectively.

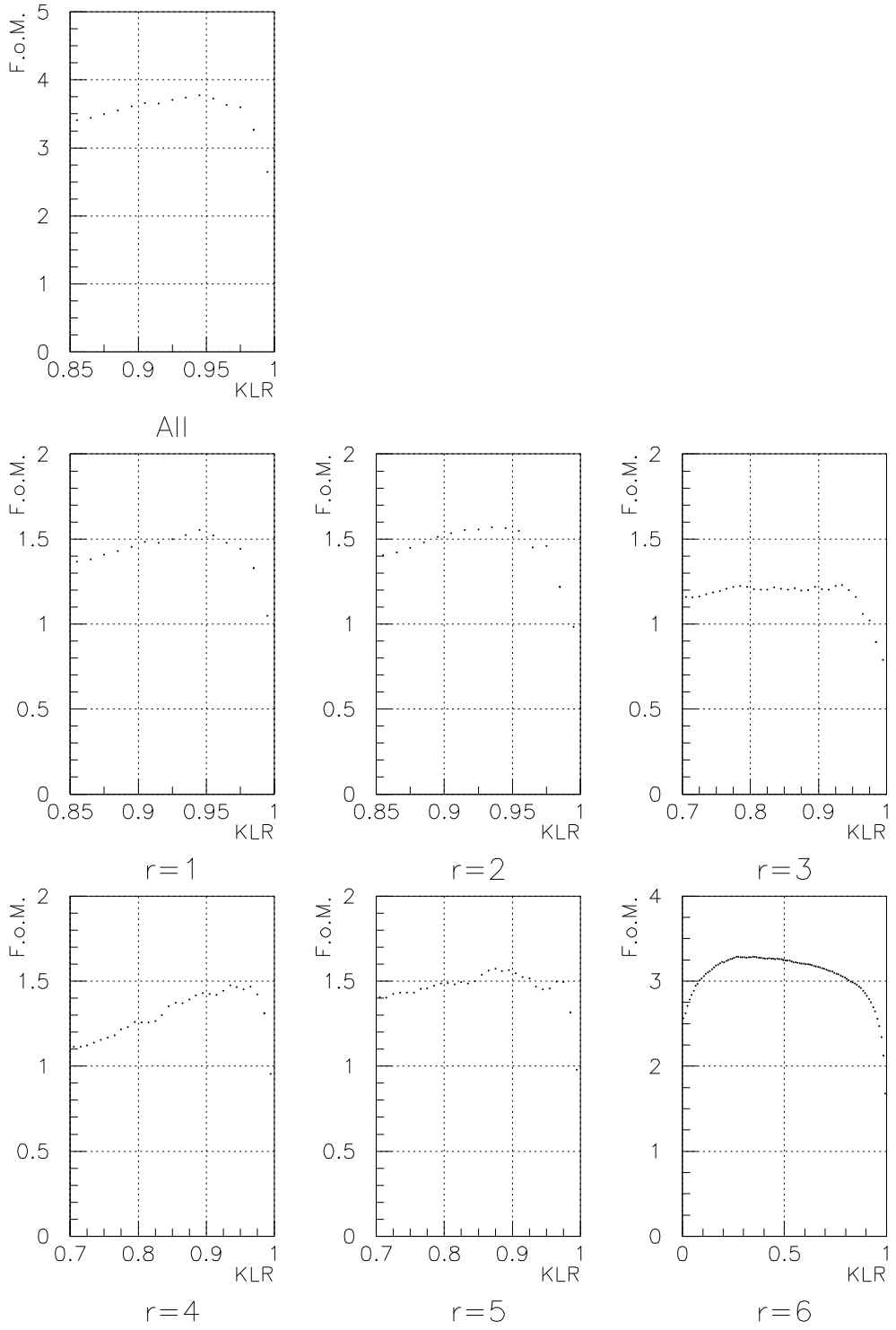
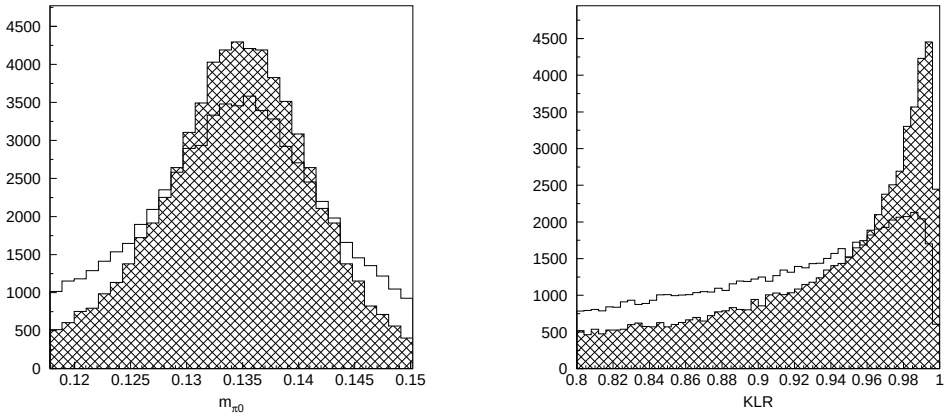


Figure 5.5: Figure of Merit (F.o.M.) for each r-bin. The cut values of KLR that maximize the F.o.M. are large in the region with large r, reflecting the fact that the signal to noise ratio is larger in the region with large r than that with small r.

The distribution depends on the r-bin and thus we treat the different r-bins separately. We use 2-dimensional PDF's of correctly reconstructed signal and SCF, denoted by $P^l_{\rm true}(m_{\pi^0}, KLR)$ and $P^l_{\rm SCF}(m_{\pi^0}, KLR)$, respectively. Here, l represents the dependence on the tagging quality r. Using the PDF's, we calculate a likelihood ratio, $\mathcal{R}_{BC}$:

$$\mathcal{R}_{BC} \equiv \frac{P^l_{\rm true}(m_{\pi^0}, KLR)}{P^l_{\rm true}(m_{\pi^0}, KLR) + P^l_{\rm SCF}(m_{\pi^0}, KLR)} \ . \tag{5.9}$$

Figure 5.7 shows the distribution of the $\mathcal{R}_{BC}$ for both correctly reconstructed signal and SCF. When a event has multiple candidates reconstructed, we calculate the $\mathcal{R}_{BC}$ for each candidate and select the candidate with the largest $\mathcal{R}_{BC}$.



(a) m_{π^0} distribution.

(b) KLR distribution.

Figure 5.6: The distributions of m_{π^0} and KLR for 4th r-bin (as an example). Hatched and outlined histograms correspond to the correctly reconstructed signal and SCF, respectively. In both plots, the correctly reconstructed signal distributes sharply, as expected.

2-5 Veto in Dalitz Plot

The contributions from $\rho(1450)$ and $\rho(1700)$ are considered to be contamination in this analysis. Thus, we apply a veto in the Dalitz plot to minimize the contaminating effect. Here, each event are required to satisfy one of following conditions:

$$\begin{aligned} 0.55\,{\rm GeV/c^2} < \sqrt{s_+} < 1.0\,{\rm GeV/c^2}\ , \\ 0.55\,{\rm GeV/c^2} < \sqrt{s_-} < 1.0\,{\rm GeV/c^2}\ , \\ 0.55\,{\rm GeV/c^2} < \sqrt{s_0} < 0.95\,{\rm GeV/c^2}\ . \end{aligned} \tag{5.10}$$

Figure 5.8 show the vetoed region schematically. We use a different condition for s_0 to minimize possible contribution from $B^0 \to f_0(980)\,\pi^0$ background.

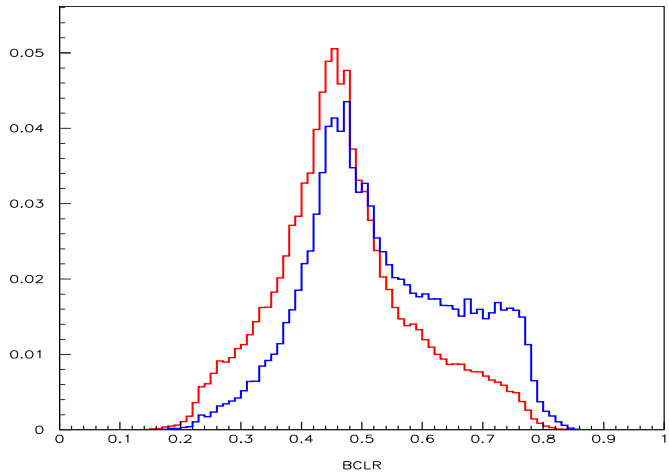


Figure 5.7: The distribution of LR_{BC} for truly reconstructed signal (blue) and SCF (red).

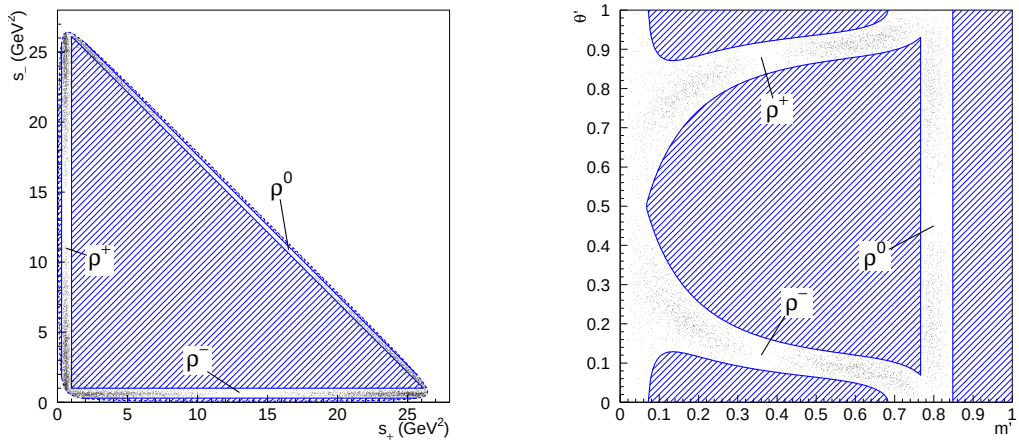


Figure 5.8: The mass window in usual Dalitz plot (left) and square Dalitz plot (right), overlayed with MC-generated signal distribution. Hatched region corresponds to the vetoed region.

3 Signal Extraction

3-1 PDF for the Fit

The event-by-event PDF for signal extraction is basically the same as that for time-dependent Dalitz plot analysis but integrated over the Δt direction and summed over $q_{\rm tag}$

$$P_{\not\Delta t}(\Delta E, M_{\rm bc}; m', \theta'; l) = \sum_{q_{\rm tag}=\pm 1} \int d\Delta t \, P(\Delta E, M_{\rm bc}; m', \theta'; \Delta t, q_{\rm tag}; l) \ , \tag{5.11}$$

where $P_{\not\Delta t}(\Delta E, M_{\rm bc}; m', \theta'; l)$ and $P(\Delta E, M_{\rm bc}; m', \theta'; \Delta t, q_{\rm tag}; l)$ are the PDF's for the fit here and that for the time-dependent Dalitz plot analysis described later in Sec. 2 of chapter 6.

We use the Dalitz plot information only for the events inside the signal region, i.e., the PDF's inside the signal region ($P^{\rm SR}_{\not\Delta t}$) and the sideband region ($P^{\rm SB}_{\not\Delta t}$) are

$$P^{\rm SR}_{\not\Delta t}(\Delta E, M_{\rm bc}; m', \theta'; l) = P_{\not\Delta t}(\Delta E, M_{\rm bc}; m', \theta'; l) \ , \tag{5.12}$$

and

$$P^{\rm SB}_{\not\Delta t}(\Delta E, M_{\rm bc}; l) = \iint dm' \, d\theta' \, P_{\not\Delta t}(\Delta E, M_{\rm bc}; m', \theta'; l) \ , \tag{5.13}$$

respectively. There are two reasons to treat the events in the signal region and those in the sideband region separately. One reason is that the correlation between ΔE-$M_{\rm bc}$ and Dalitz plot distributions is too large to be properly treated if we use the Dalitz plot information in the sideband region[3]. By limiting the events to use Dalitz plot information only to those in the ΔE-$M_{\rm bc}$ signal region, we can make the correlation smaller and under control. Another reason is that there is no benefit to use the Dalitz plot information for the events in the sideband region. The Dalitz plot information is a good discriminant to separate signal events and continuum background events. Since few signal events are in the sideband region, the Dalitz plot is not useful there.

From the event-by-event PDF, we compose an extended likelihood function, $\mathcal{L}_{\rm ext}$, for the fit. The detail of the construction of $\mathcal{L}_{\rm ext}$ can be found in appendix F. The number of free parameters, which are listed and described in table 5.2, is 36 in total. By maximizing the likelihood $\mathcal{L}_{\rm ext}$, we obtain the fit results for the parameters.

3-2 Fit Result

Figure 5.9 shows the $M_{\rm bc}$ (ΔE) distribution for the reconstructed $B^0 \to \pi^+\pi^-\pi^0$ candidates within the ΔE ($M_{\rm bc}$) signal region, together with the histograms of fitted PDF's. The fit result is listed in Table 5.2, yielding

$$N_{\rm sig} = \sum_{\text{DS-I, DS-II}} \nu_{\rm SR} f_{\rm sig} = 971 \pm 42 \ , \tag{5.14}$$

$$\left(\quad f_{\rm sig} = 1 - \sum_l f^l_{q\overline{q}} - f_{b\to c} - f_{b\to u} \quad \right)$$

where $N_{\rm sig}$ is the total number of $B^0 \to \pi^+\pi^-\pi^0$ events in the signal region. The error is statistical only and correlations are taken into account in its calculation.

[3]Since both ΔE-$M_{\rm bc}$ and Dalitz plot are kinematic variables, it is natural for them to have a correlation with each other.

Table 5.2: The parameters determined in the ΔE-$M_{\rm bc}$ and Dalitz fit for the signal yield extraction. The errors shown here are statistical only. We find significant dependence of the ΔE slope parameters p_1^l on the flavor tagging quality bin l, which is a common tendency to the analyzes of various decay modes in Belle.

Name	DS-I	DS-II	Description
$\nu_{\rm SR}$	$+913 \pm 23$	$+1848 \pm 33$	Fitted number of events in the signal region.
$f^1_{q\overline{q}}$	$+0.1969^{+0.0070}_{-0.0068}$	$+0.1877^{+0.0048}_{-0.0047}$	The products of continuum component fraction in the signal region, $f_{q\overline{q}}$, and event fraction in each tagging quality region, $\mathcal{F}^l_{q\overline{q}}$, defined as $f^l_{q\overline{q}} \equiv f_{q\overline{q}}\mathcal{F}^l_{q\overline{q}}$.
$f^2_{q\overline{q}}$	$+0.0888^{+0.0037}_{-0.0036}$	$+0.1019 \pm 0.0029$	
$f^3_{q\overline{q}}$	$+0.0677^{+0.0030}_{-0.0029}$	$+0.0662 \pm 0.0021$	
$f^4_{q\overline{q}}$	$+0.0619^{+0.0028}_{-0.0027}$	$+0.0587^{+0.0020}_{-0.0019}$	
$f^5_{q\overline{q}}$	$+0.0476^{+0.0024}_{-0.0023}$	$+0.0395 \pm 0.0015$	
$f^6_{q\overline{q}}$	$+0.1001^{+0.0042}_{-0.0041}$	$+0.1143^{+0.0033}_{-0.0032}$	
p_1^1	$+0.66 \pm 0.16$	$+0.82 \pm 0.11$	The slope of the ΔE PDF for continuum component in each tagging quality region.
p_1^2	$+0.21 \pm 0.23$	$+0.14 \pm 0.16$	
p_1^3	$+0.22 \pm 0.27$	$+0.40 \pm 0.19$	
p_1^4	$+0.06 \pm 0.28$	$+0.39 \pm 0.21$	
p_1^5	-0.45 ± 0.32	-0.39 ± 0.26	
p_1^6	-1.29 ± 0.22	-1.22 ± 0.15	
α	-16.5 ± 1.3	-16.1 ± 0.9	The parameter of ARGUS function.
U_+^-	$+1.23^{+0.13}_{-0.11}$		The time-integrated Dalitz plot coefficients. They are just nuisance parameters in this fit.
U_+^0	$+0.32^{+0.06}_{-0.05}$		
$U_{+-}^{+,\rm Re}$	$+0.92^{+0.74}_{-0.72}$		
$U_{+0}^{+,\rm Re}$	$+0.36^{+0.49}_{-0.46}$		
$U_{-0}^{+,\rm Re}$	$+0.23^{+0.50}_{-0.47}$		
$U_{+-}^{+,\rm Im}$	$+1.29 \pm 0.71$		
$U_{+0}^{+,\rm Im}$	$-0.44^{+0.36}_{-0.32}$		
$U_{-0}^{+,\rm Im}$	$-1.44^{+0.44}_{-0.39}$		
$f_{b\to c}$	$+0.0259$ (fixed)	$+0.0266$ (fixed)	The fractions of the $B\overline{B}$ backgrounds from $b \to c$ and $b \to u$ transitions.
$f_{b\to u}$	$+0.0556$ (fixed)	$+0.0557$ (fixed)	

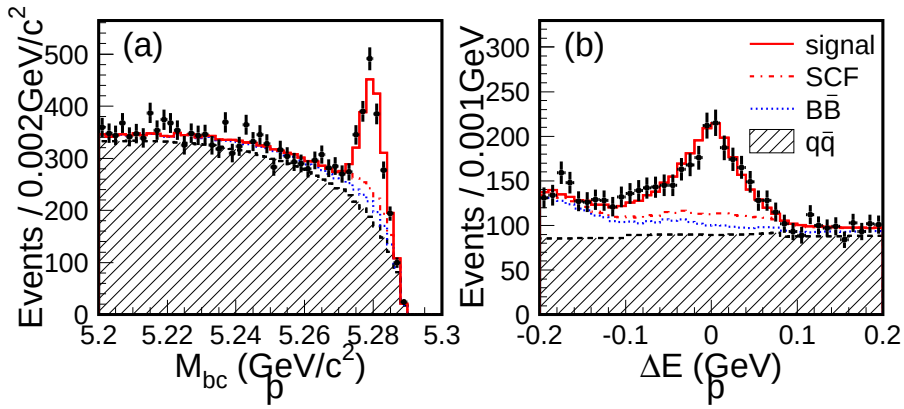


Figure 5.9: The (a) $M_{\rm bc}$ and (b) ΔE distributions within the ΔE and $M_{\rm bc}$ signal regions. The histograms are cumulative. Solid, dot-dashed, dotted and dashed hatched histograms correspond to correctly reconstructed signal, SCF, $B\overline{B}$ background, and continuum background PDF's, respectively.

Chapter 6

Time-Dependent Dalitz Plot Fit

In this chapter, we describe the time-dependent Dalitz plot fit and its result. Before performing the analysis, we need to determine the kinematics of ρ, in particular the lineshape of ρ, which is one of the most important aspect of this analysis; Sec. 1 describes the formalism and the procedure to determine the lineshape. The subsequent sections describe the fitting procedure, result, and systematic uncertainties.

1 Kinematics of ρ Meson Decay

As described in Sec. 4-3 of chapter 2, the functions ${}^{(}\overline{f}{}^{)}_{\kappa}(s_+, s_-)$, which take into account the dynamics in $B^0 \to \rho^{\kappa}\pi^{\overline{\kappa}}$ decay, play an important role in this analysis. In this section, we discuss the detail of them.

1-1 Formalism

The function ${}^{(}\overline{f}{}^{)}_{\kappa}(s_+, s_-)$ can be factorized into two parts as

$$ {}^{(}\overline{f}{}^{)}_{\kappa}(s_+, s_-) = T_J^{\kappa}\, {}^{(}\overline{F}{}^{)}{}^{\kappa}_{\pi}(s_{\kappa}) \quad (\kappa = +, -, 0) \,, \tag{6.1} $$

where T_J^{κ} and ${}^{(}\overline{F}{}^{)}{}^{\kappa}_{\pi}(s_{\kappa})$ correspond to the helicity distribution and the lineshape of ρ^{κ}, respectively.

Helicity Distribution

In the case of pseudoscalar-vector ($J = 1$) decay, T_J^{κ} is given by

$$ T_1^{\kappa} = -4|\vec{p}_j||\vec{p}_k|\cos\theta^{jk} \,, \tag{6.2} $$

$$ \left(\begin{array}{l} T_1^{+} = -4|\vec{p}_+||\vec{p}_-|\cos\theta^{+-} \,, \\ T_1^{-} = -4|\vec{p}_0||\vec{p}_+|\cos\theta^{0+} \,, \\ T_1^{0} = -4|\vec{p}_-||\vec{p}_0|\cos\theta^{-0} \,, \end{array} \right) \tag{6.3} $$

where $\vec{p}_j, \vec{p}_k$ are the three momenta of the π^j and π^k in the rest frame of ρ^{κ} (or the $\pi^i\pi^j$ system), and the $\theta^{jk}(\equiv \theta_{\kappa})$ is the angle between $\vec{p}_j$ and $\vec{p}_k$ (see Fig. 6.1).

An equivalent alternative expression for the T_1^{κ} is

$$ T_1^{\kappa} = s^{ki} - s^{jk} + \frac{({m_{B^0}}^2 - {m_{\pi^k}}^2)({m_{\pi^j}}^2 - {m_{\pi^i}}^2)}{s^{ij}} \,, \tag{6.4} $$

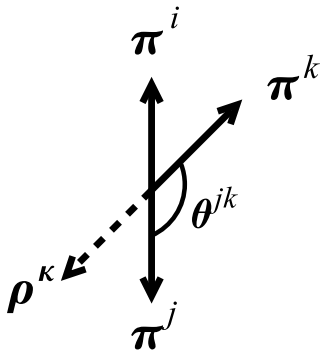


Figure 6.1: The relation between three pions in the rest frame of ρ^κ.

with

$$s^{0+} \equiv (p_0 + p_+)^2 = s_+ \ , \quad s^{+-} \equiv (p_+ + p_-)^2 = s_0 \ , \quad \text{and} \quad s^{-0} \equiv (p_- + p_0)^2 = s_- \ . \quad (6.5)$$

In a Dalitz plot analysis, this expression written in terms of the Dalitz plot variables is sometimes convenient. The derivation of the equivalence can be found in Sec. 1 of appendix D.

Lineshape of ρ

The lineshape is parameterized with Breight-Wigner functions corresponding to the $\rho(770)$, $\rho(1450)$, and $\rho(1700)$ resonances:

$$\overset{(-)}{F}{}^{\kappa}_{\pi}(s) = BW_{\rho(770)} + \overset{(-)}{\beta}_{\kappa} BW_{\rho(1450)} + \overset{(-)}{\gamma}_{\kappa} BW_{\rho(1700)} \ , \quad (6.6)$$

where the amplitudes $\overset{(-)}{\beta}_{\kappa}$ and $\overset{(-)}{\gamma}_{\kappa}$ (denoting the relative size of two resonances) are complex numbers. We use Gounaris-Sakurai (GS) model [51] for the Breit-Wigner shape (see Sec. 2 of appendix D for the detail) and world average [52] for the mass and width of each resonance. Though the $\overset{(-)}{\beta}_{\kappa}$ and $\overset{(-)}{\gamma}_{\kappa}$ can be different for each of six decay modes of $B^0(\overline{B}^0) \to \rho^\kappa \pi^{\overline{\kappa}}$ in general, we assume no such variation, i.e.,

$$\overset{(-)}{F}{}^{\kappa}_{\pi}(s) = F_\pi(s) \equiv BW_{\rho(770)} + \beta\, BW_{\rho(1450)} + \gamma\, BW_{\rho(1700)}, \quad (6.7)$$

in our nominal fit, and address the possible variation in the systematic error. Equation (6.7) leads to the relation of $\overline{f}_\kappa(s_+, s_-) = f_\kappa(s_+, s_-)$, which is assumed in Sec. 4-3 of chapter 2.

1-2 Determination of ρ Lineshape

Using our data sample, we determine the (β, γ) used in our nominal fit and possible deviations of $(\overset{(-)}{\beta}_{\kappa}, \overset{(-)}{\gamma}_{\kappa})$ from the determined (β, γ).

Data Sample and Selection Criteria

The data sample and selection criteria used here are the same as those described in chapter 5, except that we adopt a wider Dalitz plot mass window here to include the Dalitz plot region corresponding to the radial excitations; each event is required to satisfy $0.55\,\mathrm{GeV}/c^2 < \sqrt{s_0} < 1.5\,\mathrm{GeV}/c^2$, $\sqrt{s_+} < 1.5\,\mathrm{GeV}/c^2$, or $\sqrt{s_-} < 1.5\,\mathrm{GeV}/c^2$ (see Fig. 6.2). The signal fraction and the parameters of the continuum background of this data sample are determined in the same way as described in chapter 5.

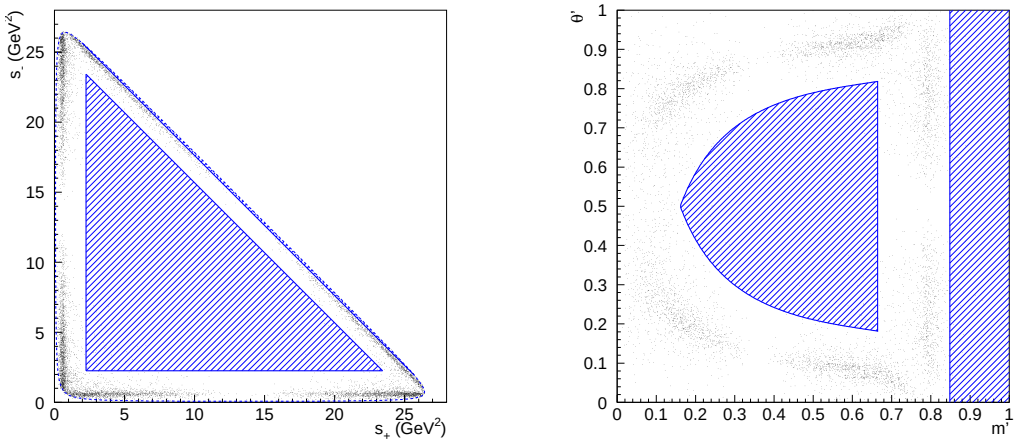


Figure 6.2: The mass window for the lineshape determination in usual Dalitz plot (left) and square Dalitz plot (right), overlayed with MC-generated signal distribution. Hatched region corresponds to the vetoed region.

Lineshape Determination

Using the data sample described above, we determine the ρ lineshape, i.e. the phases and amplitudes of the coefficients β and γ in equation (6.7), which we use for all of the decay amplitudes. In this fit, we use the PDG values [52] for the masses and widths of the $\rho(770)$, $\rho(1450)$, and $\rho(1700)$. The fit yields the result listed in Table 6.1. As can be seen in the table, the free parameters are the amplitudes (β, γ) and the time-integrated Dalitz plot parameters of U_+^X. The latter are nuisance parameters in the fit here[1]. The mass distributions and fit results are shown in Fig. 6.3. Figure 6.4 schematically shows how the radial excitations contribute to our fit result. Note that the values given here for β and γ and their errors are not meaningful measurements of physics parameters but rather are quantities needed for the time-dependent Dalitz fit. This is because these parameters are determined from the interference region, the interference between $\rho^+\pi^-$ and $\rho^-\pi^+$, etc., and depend on the unfounded common lineshape assumption of Eq. (6.7). However, because statistics are still low, the time-dependent Dalitz analysis would not be possible if we were to discard the common lineshape assumption.

Thus, it is important to determine the common or *average* lineshape as well as obtain an upper limit on the deviation from the average lineshape for each of the six decay amplitudes, that is, the deviation of $(\overline{\beta}_\kappa, \overline{\gamma}_\kappa)$ from the nominal (β, γ). For this purpose, we put constraints on additional amplitudes that describe 1) the excess in the high mass region, $\sqrt{s} > 0.9\,\mathrm{GeV}/c^2$, and 2) the interferences between radial excitations and the lowest resonance, the $\rho(770)$: interferences between $\rho(770)^+\pi^-$ and $\rho(1450)^-\pi^+$, etc. The nominal fit is performed with the average lineshape determined above, fixing all of the additional amplitudes to zero. When floating the additional amplitudes for the other resonances, we obtain results consistent with zero for all of the additional amplitudes but with large uncertainties compared to the errors for

[1]Since (β, γ) are highly correlated with the time-integrated Dalitz plot parameters, in particular with the interfering parameters, it is important to set them free and estimate the uncertainty of (β, γ) properly. Note that the fitted values of the Dalitz plot parameters here are not important and not necessarily consistent with the values of our final time-dependent Dalitz plot fit result. This is because the fitted Dalitz plot parameters here are strongly affected by the contribution from radial excitations, $\rho(1450)$ and $\rho(1700)$, which we intentionally avoid in the final fit by adopting narrower mass window in the Dalitz plot than that used here.

the average lineshape parameters above. The detail of the study can be found in appendix E. In the systematic error study, we use the fit result with the additional lineshape parameters floating including their uncertainties.

Table 6.1: Result of the lineshape fit. The large phase differences between $\rho(770)$ and $\rho(1450)$, $\arg\beta$, and between $\rho(1450)$ and $\rho(1700)$, $\arg\gamma - \arg\beta$, indicate that the interferences between them are destructive.

Parameter Name	Values and Errors
$\lvert\beta\rvert$	$0.313^{+0.067}_{-0.055}$
$\lvert\gamma\rvert$	$0.082^{+0.042}_{-0.032}$
$\arg\beta$	$\left(+219^{+16}_{-18}\right)^\circ$
$\arg\gamma$	$\left(+102^{+26}_{-32}\right)^\circ$
U_+^-	$+1.30^{+0.13}_{-0.12}$
U_+^0	$+0.35^{+0.06}_{-0.06}$
$U_+^{+-,\mathrm{Re}}$	$+0.60^{+0.69}_{-0.72}$
$U_+^{+0,\mathrm{Re}}$	$-0.09^{+0.34}_{-0.33}$
$U_+^{-0,\mathrm{Re}}$	$+0.26^{+0.37}_{-0.39}$
$U_+^{+-,\mathrm{Im}}$	$+1.23^{+0.72}_{-0.73}$
$U_+^{+0,\mathrm{Im}}$	$-0.19^{+0.36}_{-0.39}$
$U_+^{-0,\mathrm{Im}}$	$-1.48^{+0.41}_{-0.41}$

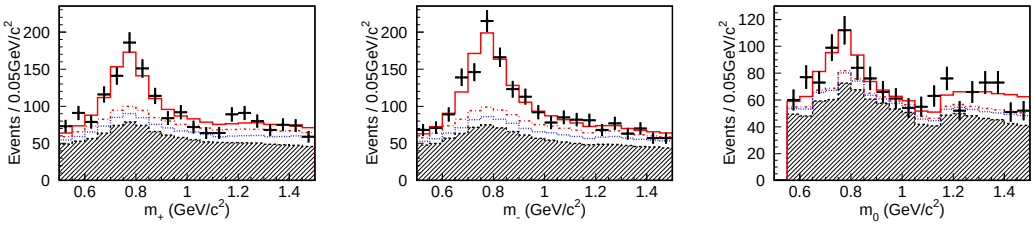


Figure 6.3: Mass distributions and fitted lineshapes in $\rho^+\pi^-$ (left), $\rho^-\pi^+$ (middle), and $\rho^0\pi^0$ (right) enhanced regions. The histograms are cumulative. Solid, dot-dashed, dotted and dashed hatched histograms correspond to correctly reconstructed signal, SCF, $B\overline{B}$, and continuum PDFs, respectively. Note that there are feed-downs from other quasi-two-body components than those of interest, especially in the high-mass regions. For example, the high-mass region ($m_0 \gtrsim 1.0\,\mathrm{GeV}/c^2$) of the $\rho^0\pi^0$ enhanced region (right) includes large contributions from $\rho^\pm\pi^\mp$.

2 Event-by-Event PDF for Time-Dependent Dalitz Fit

To determine the 26 time-dependent Dalitz plot parameters, we define the following event-by-event PDF:

$$P(\vec{x}) = f_{\mathrm{sig}}\mathcal{P}_{\mathrm{sig}}(\vec{x}) + f_{q\overline{q}}\mathcal{P}_{q\overline{q}}(\vec{x}) + f_{B\overline{B}}\mathcal{P}_{B\overline{B}}(\vec{x})\,, \tag{6.8}$$

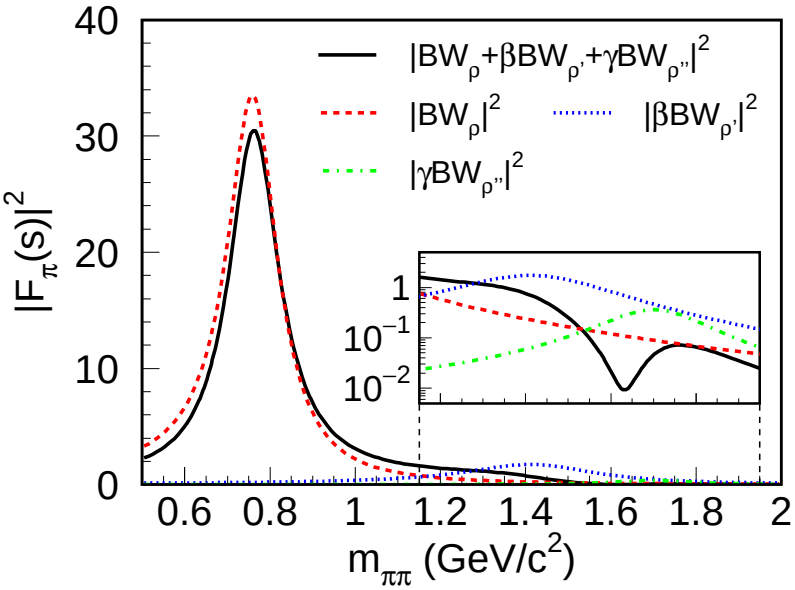


Figure 6.4: A schematic figure of the fit result of the lineshape and the contributions from radial excitations. Note that our definition of $F_\pi(s)$ does not include the factor $1/(1+\beta+\gamma)$ as in Eq. (6.7). One can see that the $\rho(770)$ and $\rho(1450)$ destructively interfere with each other near $\sqrt{s} \equiv m_{\pi\pi} = 1.4(\mathrm{GeV}/c^2)$, which means that the $\rho(1450)$ has a large impact on the phase of $F_\pi(s)$ although the absolute value of $|F_\pi(s)|$ is not much affected.

where $\mathcal{P}_{\rm sig}$, $\mathcal{P}_{q\overline{q}}$, and $\mathcal{P}_{B\overline{B}}$ are PDF's for signal, continuum background, and $B\overline{B}$ background, respectively, and $f_{\rm sig}$, $f_{q\overline{q}}$, and $f_{B\overline{B}}$ are the corresponding fractions that satisfy

$$f_{\rm sig} + f_{q\overline{q}} + f_{B\overline{B}} = 1 \ . \tag{6.9}$$

The vector $\vec{x}$, the argument of the PDF's, represents a set of event-by-event variables:

$$\vec{x} \equiv (\Delta E, M_{\rm bc}; m', \theta'; \Delta t, q_{\rm tag}, l; p_{\pi^0}, E_{\rm beam}) \ , \tag{6.10}$$

where ΔE is the energy difference between the B_{CP} candidate and the beam energy, $M_{\rm bc}$ is the beam-energy constrained mass calculated using the beam energy in place of the reconstructed energy, (m', θ') are the square Dalitz plot variables, Δt is the proper time difference between the decays of B_{CP} and $B_{\rm tag}$, $q_{\rm tag}$ and l are the flavor of $B_{\rm tag}$ and its quality obtained by the flavor-tagging procedure, p_{π^0} is the momentum of the π^0 of the $\pi^+\pi^-\pi^0$ final state measured in the laboratory frame, and $E_{\rm beam}$ is the run-dependent beam energy.

In this section, we briefly describe the PDF for each component. The detailed study of them can be found in appendix A.

2-1 Signal PDF

The PDF of the signal component, $\mathcal{P}_{\rm sig}$, consists of a PDF for correctly reconstructed events, $\mathcal{P}_{\rm true}$, and PDF's of the SCF components, $\mathcal{P}_{\rm CR}$ and $\mathcal{P}_{\rm NR}$:

$$\mathcal{P}_{\rm sig}(\vec{x}) = \frac{\mathcal{P}_{\rm true}(\vec{x}) + \sum\limits_{i={\rm CR,\ NR}} \mathcal{P}_i(\vec{x})}{n_{\rm true} + \sum\limits_{i={\rm CR,\ NR}} n_i} \ , \tag{6.11}$$

where CR and NR represent two types of SCF, the $\pi^\pm$ (charged) replaced and π^0 (neutral) replaced, respectively; and $n_{\rm sig}$, $n_{\rm CR}$, and $n_{\rm NR}$ are the integrals of the PDF's in the signal region. We describe each component in the following.

PDF for correctly reconstructed events

In terms of the event fractions for the $l^{\rm th}$ flavor tagging region ($\mathcal{F}^l_{\rm true}$), the Dalitz plot dependent efficiency (ϵ^l), the π^0 momentum dependent efficiency correction taking account of the difference between data and MC (ϵ'), wrong tag fraction (w_l), and the wrong tag fraction difference between B^0 and $\overline{B}^0$ (Δw_l), the PDF for correctly reconstructed events is given by

$$\begin{aligned}\mathcal{P}_{\rm true}(\vec{x}) &= \mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; m', \theta'; \Delta t, q_{\rm tag}, l; p_{\pi^0}) \\ &= \mathcal{F}^l \cdot \mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0}) \cdot \epsilon^l(m', \theta')\, \epsilon'(p_{\pi^0}) \cdot |{\rm det}\boldsymbol{J}(m', \theta')| \cdot \mathcal{P}^l_{\rm true}(m', \theta'; \Delta t, q_{\rm tag}),\end{aligned} \tag{6.12}$$

where $|{\rm det}\boldsymbol{J}(m', \theta')|$ is the Jacobian for the square Dalitz plot defined in Eq. (4.22) and

$$\begin{aligned}&\mathcal{P}^l_{\rm true}(m', \theta'; \Delta t, q_{\rm tag}) \\ &= \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}} \cdot \Bigg\{ (1 - q_{\rm tag}\Delta w_l)(|A_{3\pi}|^2 + |\overline{A}_{3\pi}|^2) \\ &\quad + q_{\rm tag}(1 - 2w_l) \cdot \left[-(|A_{3\pi}|^2 - |\overline{A}_{3\pi}|^2)\cos(\Delta m_d \Delta t) + 2{\rm Im}\left(\frac{q}{p}\overline{A}_{3\pi}A^*_{3\pi}\right)\sin(\Delta m_d \Delta t) \right] \Bigg\} \ ,\end{aligned} \tag{6.13}$$

which is based on Eq. (2.161) and follows the parameterization described in Sec. 4-2 of chapter 4. For the Δt PDF, the above equation is convolved with the resolution function as described in Sec. 3-4 of chapter 4.

The ΔE-$M_{\rm bc}$ PDF is normalized such that

$$\iint_{\rm SR} d\Delta E\, dM_{\rm bc}\, \mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0}) = 1 \quad (\forall p_{\pi^0}) \ , \tag{6.14}$$

where $\iint_{\rm SR} d\Delta E\ dM_{\rm bc}$ represents the integration over the ΔE-$M_{\rm bc}$ signal region. This normalization condition is convenient since we define the Dalitz plot efficiency for events inside the signal region. With the PDF being also normalized to be unity in the Δt-$q_{\rm tag}$ direction, the integral inside the signal region, $n_{\rm true}$, is,

$$n_{\rm true} = \sum_l n^l_{\rm true} \ , \tag{6.15}$$

$$\begin{aligned} n^l_{\rm true} &\equiv \sum_{q_{\rm tag}} \int d\Delta t \iint_{\rm SR} d\Delta E\, dM_{\rm bc} \iint_{\rm SDP,\ Veto} dm'\, d\theta'\, \mathcal{P}^l_{\rm true}(\Delta E, M_{\rm bc}; m', \theta'; \Delta t, q_{\rm tag}; p_{\pi^0}) \\ &= \mathcal{F}^l_{\rm true} \iint_{\rm SDP,\ Veto} dm'\, d\theta'\, \epsilon^l_{\rm true}(m', \theta')\, \epsilon'(p_{\pi^0})\, |{\rm det}\boldsymbol{J}|\, \left(|A_{3\pi}|^2 + |\overline{A}_{3\pi}|^2 \right) \ , \end{aligned} \tag{6.16}$$

where the correlation between p_{π^0} and m' is properly taken into account in the integration of the last line. The notation $\iint_{\rm SDP,\ Veto} dm'\ d\theta'$ means integration over the square Dalitz plot with the vetoed region in the Dalitz plot being taken into account.

The π^0 momentum dependent ΔE-$M_{\rm bc}$ PDF, $\mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0})$, is modeled using MC-simulated events in a binned histogram interpolated in the p_{π^0} direction, to which a small correction obtained with $B^0 \to D^{(*)-}\rho^+$ is applied to account for the difference between MC and data.

The Dalitz plot distribution is smeared and distorted by detection efficiencies and detector resolutions. We obtain the signal Dalitz plot efficiency from MC to take the former into account. We introduce a dependence of the efficiency on the r region, $\epsilon^l_{\rm true}$, since a significant dependence is observed in MC. Small corrections, $\epsilon'(p_{\pi^0})$, are also applied to the MC-determined efficiency to account for differences between MC and data. We use $\overline{B}^0 \to \rho^- D^{(*)+}$, $\overline{B}^0 \to \pi^- D^{*-}$, $B^- \to \rho^- D^0$ and $B^- \to \pi^- D^0$ decays to obtain the correction factors. The detector resolutions are small compared to the widths of $\rho(770)$ resonances; this is confirmed by MC to be a negligibly small effect.

PDF for SCF events

The SCF (self cross feed) are wrongly reconstructed signal events, with tracks from $B_{\rm tag}$ wrongly included in $f_{CP} = \pi^+\pi^-\pi^0$, or with fake π^0's consisting of wrong γ combinations or fake γ's. Since almost all SCF events have only a single wrong track for each, we can categorize the SCF into two types by the charge of the wrong track, $\pi^\pm$ (charged) replaced (CR) or π^0 (neutral) replaced (NR). Approximately 20% of signal candidates are SCF's, which are subdivided into $\sim 4\%$ of NR SCF and $\sim 16\%$ of CR SCF. This fraction is sizable and thus it is important to model the SCF well. The time-dependent PDF for SCF events is defined as

$$\begin{aligned} \mathcal{P}_i(\vec{x}) &= \mathcal{P}^l_i(\Delta E, M_{\rm bc}; m', \theta'; \Delta t, q_{\rm tag}) \\ &= \mathcal{F}^l_i \cdot \mathcal{P}_i(\Delta E, M_{\rm bc}; s_i) \cdot \mathcal{P}_i(m', \theta'; \Delta t, q_{\rm tag}) \ , \quad (i = {\rm NR}, {\rm CR}) \end{aligned} \tag{6.17}$$

$$\left(\quad s_{\rm CR} \equiv \max(s_+, s_-) \ , \quad s_{\rm NR} \equiv s_0 \quad \right) \tag{6.18}$$

where $\mathcal{F}^l_i$, $\mathcal{P}_i(\Delta E, M_{\rm bc}; s_i)$, and $\mathcal{P}_i(m', \theta'; \Delta t, q_{\rm tag})$ are the event fraction in l-th tagging r-bin region, a ΔE-$M_{\rm bc}$ PDF with the dependence on Dalitz plot, and a Dalitz-Δt PDF, respectively.

The reason why the Dalitz dependence of the ΔE-$M_{\rm bc}$ PDF is modeled well can be found in Sec. 2-2 of appendix A. Here, the Dalitz-Δt PDF is

$$\begin{aligned}\mathcal{P}_i(m',\theta';\Delta t, q_{\rm tag}) = \frac{e^{-|\Delta t|/\tau_i}}{4\tau_i}\cdot\Big\{&(1-q_{\rm tag}\Delta {w_l}^i)\,P_i^{\rm Life}(m',\theta') \\ &+ q_{\rm tag}(1-2{w_l}^i)\cdot\Big[-P_i^{\rm Cos}(m',\theta')\cos(\Delta m_d\Delta t)+P_i^{\rm Sin}(m',\theta')\sin(\Delta m_d\Delta t)\Big]\Big\}\,,\end{aligned} \tag{6.19}$$

where τ_i, ${w_l}^i$, and $\Delta {w_l}^i$ are the effective lifetime, wrong-tag fraction, and wrong-tag fraction difference for SCF's, respectively; and $P_i^{\rm Life}(m',\theta')$, $P_i^{\rm Cos}(m',\theta')$, and $P_i^{\rm Sin}(m',\theta')$ are the Dalitz plot dependent coefficients of the time-dependences of $e^{-|\Delta t|/\tau_i}$, $e^{-|\Delta t|/\tau_i}\cos(\Delta t\Delta m_d)$, and $e^{-|\Delta t|/\tau_i}\sin(\Delta t\Delta m_d)$, respectively. All of them are described later.

The ΔE-$M_{\rm bc}$ PDF is normalized inside the signal region as

$$\iint_{\rm SR} d\Delta E\; dM_{\rm bc}\;\mathcal{P}_i(\Delta E, M_{\rm bc}; s_i) = 1 \quad (\forall s_i)\;. \tag{6.20}$$

With the PDF's in the Δt-$q_{\rm tag}$ direction being normalized to unity and $\sum_l \mathcal{F}_i^l = 1$, the integral inside the signal region, n_i, is

$$\begin{aligned} n_i &\equiv \sum_l\sum_{q_{\rm tag}}\int d\Delta t\iint_{\rm SR} d\Delta E\; dM_{\rm bc}\;\iint_{\rm SDP,\ Veto} dm'\;d\theta'\;\mathcal{P}_i^l(\Delta E, M_{\rm bc}; m',\theta';\Delta t, q_{\rm tag}) \\ &= \iint_{\rm SDP,\ Veto} dm'\;d\theta'\;P_i^{\rm Life}(m',\theta')\;. \end{aligned} \tag{6.21}$$

We find that the ΔE-$M_{\rm bc}$ distribution for SCF has a sizable correlation with Dalitz plot variables, but with only one of its two dimensions. We thus introduce a model with dependences on the Dalitz plot variable s_i. The variable $s_{\rm CR} = s_\pm \equiv \max(s_+, s_-)$ is used, because the CR SCF can be divided into a π^+ replaced SCF and a π^- replaced SCF, where s_- (s_+) is used for π^+ (π^-) replaced SCF. Here, we exploit the fact that almost all of the π^+ (π^-) replaced SCF distributes in the region of $s_+ < s_-$ ($s_+ > s_-$). For the NR SCF, $s_{\rm NR} = s_0$. This parameterization models the correlation quite well, with each of the parameters s_i reasonably related to the kinematics of replaced tracks.

Since the track (π) replacement changes the measured kinematic variables, the SCF events "migrate" in the Dalitz plot from the correct (or generated) position to the observed position. Using MC, we determine resolution functions $R_i(m'_{\rm obs},\theta'_{\rm obs}; m'_{\rm gen},\theta'_{\rm gen})$ to describe this "migration" effect, where $(m'_{\rm obs},\theta'_{\rm obs})$ and $(m'_{\rm gen},\theta'_{\rm gen})$ are the observed and the generated (correct) positions in the Dalitz plot, respectively. Together with the efficiency function $\epsilon_i(m'_{\rm gen},\theta'_{\rm gen})$, which is also obtained with MC, the Dalitz plot PDF for SCF is described as

$$\begin{aligned} P_i^j(m',\theta') &= \left[(R_i\cdot\epsilon_i)\otimes P_{\rm gen}^j\right](m',\theta') \\ &\equiv \iint_{\rm SDP} dm'_{\rm gen}\;d\theta'_{\rm gen}\;R_i(m',\theta';m'_{\rm gen},\theta'_{\rm gen})\cdot\epsilon_i(m'_{\rm gen},\theta'_{\rm gen})\cdot P_{\rm gen}^j(m'_{\rm gen},\theta'_{\rm gen})\;, \end{aligned} \tag{6.22}$$

$$(\ i = {\rm CR}, {\rm NR}\;,\quad j = {\rm Life}, {\rm Cos}, {\rm Sin}\ )$$

where

$$P_{\rm gen}^{\rm Life}(m'_{\rm gen},\theta'_{\rm gen}) = |{\rm det}\boldsymbol{J}|(|A_{3\pi}|^2+|\overline{A}_{3\pi}|^2)\;, \tag{6.23}$$

$$P_{\rm gen}^{\rm Cos}(m'_{\rm gen},\theta'_{\rm gen}) = |{\rm det}\boldsymbol{J}|(|A_{3\pi}|^2-|\overline{A}_{3\pi}|^2)\;, \tag{6.24}$$

$$P_{\rm gen}^{\rm Sin}(m'_{\rm gen},\theta'_{\rm gen}) = |{\rm det}\boldsymbol{J}|2{\rm Im}\left[\frac{q}{p}\overline{A}_{3\pi}A_{3\pi}^*\right]\;. \tag{6.25}$$

For the NR SCF, the shape of the Δt PDF defined in equation (6.19) is exactly the same as correctly reconstructed signal, i.e., $\tau_{\rm NR} = \tau_{B^0}$, ${w_l}^{\rm NR} = w_l$, and $\Delta {w_l}^{\rm NR} = \Delta w_l$, since the replaced track, π^0, is not used for either vertexing or flavor tagging. On the other hand, for the CR SCF, the Δt PDF is different from correctly reconstructed signal, since the replaced $\pi^\pm$ is used for both vertexing and flavor tagging. Thus, we use MC-simulated CR SCF events to obtain $\tau_{\rm CR}$, ${w_l}^{\rm CR}$, and $\Delta {w_l}^{\rm CR}$, which are different from those of correctly reconstructed signal events. In particular, $\Delta {w_l}^{\rm CR}$ is opposite in sign for the π^+ and π^- replaced SCF's, which is due to the fact that the replaced $\pi^\pm$ tends to be directly used for flavor tagging in the slow pion category.

2-2 Continuum PDF

The PDF for the continuum background is

$$\begin{aligned}\mathcal{P}(\vec{x}) &= \mathcal{P}^l_{q\overline{q}}(\Delta E, M_{\rm bc}; m', \theta'; \Delta t, q_{\rm tag}) \\ &= \mathcal{F}^l_{q\overline{q}} \cdot \mathcal{P}^l_{q\overline{q}}(\Delta E, M_{\rm bc}) \cdot \mathcal{P}_{q\overline{q}}(m', \theta'; \Delta E, M_{\rm bc}) \cdot \left[\frac{1 + q_{\rm tag} A^l_{q\overline{q}}(m', \theta'; \Delta E, M_{\rm bc})}{2}\right] \cdot \mathcal{P}_{q\overline{q}}(\Delta t),\end{aligned} \tag{6.26}$$

where $\mathcal{F}^l_{q\overline{q}}$, $\mathcal{P}^l_{q\overline{q}}(\Delta E, M_{\rm bc})$, $\mathcal{P}_{q\overline{q}}(m', \theta'; \Delta E, M_{\rm bc})$, $A^l_{q\overline{q}}(m', \theta'; \Delta E, M_{\rm bc})$, and $\mathcal{P}_{q\overline{q}}(\Delta t)$ are the event fraction in l-th r-region obtained in the signal yield fit[2], a ΔE-$M_{\rm bc}$ PDF, a Dalitz plot PDF with the dependence on ΔE-$M_{\rm bc}$ taken into account, Dalitz plot dependent flavor asymmetry, and a Δt PDF. All the terms on the right hand side of the equation are normalized to be unity as

$$\sum_l \mathcal{F}^l_{q\overline{q}} = 1 \,, \tag{6.27}$$

$$\iint_{\rm SR} d\Delta E \, dM_{\rm bc} \, \mathcal{P}^l_{q\overline{q}}(\Delta E, M_{\rm bc}) = 1 \,, \tag{6.28}$$

$$\iint_{\rm SDP,\ Veto} dm' \, d\theta' \, \mathcal{P}_{q\overline{q}}(m', \theta'; \Delta E, M_{\rm bc}) = 1 \quad (\forall \Delta E, \forall M_{\rm bc}) \,, \tag{6.29}$$

$$\sum_{q_{\rm tag}} \frac{1 + q_{\rm tag} A^l(m', \theta'; \Delta E, M_{\rm bc})}{2} = 1 \quad (\forall \Delta E, \forall M_{\rm bc}) \,, \tag{6.30}$$

$$\int d\Delta t \, \mathcal{P}_{q\overline{q}}(\Delta t) = 1 \,, \tag{6.31}$$

so that

$$\sum_l \sum_{q_{\rm tag}} \int d\Delta t \iint_{\rm SR} d\Delta E \, dM_{\rm bc} \iint_{\rm SDP,\ Veto} dm' \, d\theta' \, \mathcal{P}^l_{q\overline{q}}(\Delta E, M_{\rm bc}; m', \theta'; \Delta t, q_{\rm tag}) = 1 \,. \tag{6.32}$$

Since the allowed kinematic region is dependent on ΔE and $M_{\rm bc}$, the Dalitz plot distribution is dependent on ΔE and $M_{\rm bc}$. We define a ΔE-$M_{\rm bc}$ independent PDF, $\mathcal{P}_{q\overline{q}}(m'_{\rm scale}, \theta')$, where $m'_{\rm scale}$ is a modified version of the square Dalitz plot variable m', originally defined by Eq. (4.20), with the kinematic effect taken into account as

$$m'_{\rm scale} \equiv \frac{1}{\pi} \arccos\left(2 \frac{m_0 - m_0^{\rm min}}{m_0^{\rm max} - m_0^{\rm min} + \Delta E + \Delta M_{\rm bc}} - 1\right) \,, \tag{6.33}$$

where

$$\Delta M_{\rm bc} \equiv M_{\rm bc} - m_{B^0} \,. \tag{6.34}$$

[2]Strictly speaking, what is determined in the signal yield fit is the product of $f_{q\overline{q}}$ and $\mathcal{F}^l$, $f^l_{q\overline{q}} \equiv f_{q\overline{q}} \cdot \mathcal{F}^l$.

Using the ΔE-$M_{\rm bc}$ independent PDF, $\mathcal{P}_{q\overline{q}}(m',\theta';\Delta E, M_{\rm bc})$ is described as

$$\mathcal{P}_{q\overline{q}}(m',\theta';\Delta E, M_{\rm bc}) = \begin{cases} \dfrac{1}{N(\Delta E+\Delta M_{\rm bc})}\cdot\dfrac{\sin(\pi m')}{\sin(\pi m'_{\rm scale})}\cdot\mathcal{P}_{q\overline{q}}(m'_{\rm scale},\theta') & (m_0\in\mathcal{M})\ , \\ 0 & (\text{otherwise})\ , \end{cases} \tag{6.35}$$

$$\left(\quad \mathcal{M}\equiv\left[m_0^{\rm min},\ \min(m_0^{\rm max}, m_0^{\rm max}+\Delta E+\Delta M_{\rm bc})\right] \quad \right)$$

where $N(\Delta E+\Delta M_{\rm bc})$ and $\sin(\pi m')/\sin(\pi m'_{\rm scale})$ are a normalization factor and the Jacobian for the parameter transformation of $m'_{\rm scale}\mapsto m'$, respectively. We obtain the $\mathcal{P}_{q\overline{q}}(m'_{\rm scale},\theta')$ distribution from data in part of the sideband region, $-0.1\,{\rm GeV} < \Delta E < 0.2\,{\rm GeV}$ and $5.2\,{\rm GeV}/c^2 < M_{\rm bc} < 5.26\,{\rm GeV}/c^2$, where the contribution from $B\overline{B}$ background is negligible.

Since we find significant flavor asymmetry dependent on the location in the Dalitz plot, we introduce the following term to take account of it:

$$\frac{1+q_{\rm tag}A^l_{q\overline{q}}(m',\theta';\Delta E, M_{\rm bc})}{2}\ , \tag{6.36}$$

which is r-region (l) dependent. The asymmetry is anti-symmetric in the direction of θ', i.e., $A^l_{q\overline{q}}>0$ ($A^l_{q\overline{q}}<0$) in the region of $\theta'>0.5$ ($\theta<0.5$), and the size of the asymmetry is $\sim 20\%$ at most in the best r-region. Note that the anti-symmetric property means that the introduced effect is not CP-violating. This asymmetry is due to the jet-like topology of continuum events; when an event has a high momentum π^- (π^+) on the CP side, the highest momentum π on the tag side tends to have + (−) charge. The highest momentum π on the tag side with + (−) charge tags the flavor of $B_{\rm tag}$ as B^0 ($\overline{B}^0$)[3]. Since an event with a high momentum π^- (π^+) resides in the region $\theta'>0.5$ ($\theta'<0.5$), a continuum event in the region $\theta'>0.5$ ($\theta'<0.5$) tends to be tagged as B^0 ($\overline{B}^0$). We again parameterize the $A^l_{q\overline{q}}(m',\theta')$ in a ΔE-$M_{\rm bc}$ independent way as

$$A^l_{q\overline{q}}(m',\theta';\Delta E, M_{\rm bc}) = A^l_{q\overline{q}}(m'_{\rm scale},\theta')\ , \tag{6.37}$$

and model with a two-dimensional polynomial, whose coefficients are determined by a fit to data in the sideband region.

2-3 $B\overline{B}$ background PDF

The $B\overline{B}$ background PDF is a linear combination of the all decay modes that are expected to contribute:

$$\mathcal{P}_{B\overline{B}}(\vec{x}) = \frac{\sum_k \epsilon_k Br_k \mathcal{P}_k(\vec{x})}{\sum_k \epsilon_k Br_k}\ , \tag{6.38}$$

where k is an index over the modes; and ϵ_k, Br_k, and $\mathcal{P}_k(\vec{x})$ are the efficiency, branching fraction, and PDF for each mode, respectively.

The treatment of the PDF for each $B\overline{B}$ mode is different for CP-eigenstate modes and flavor-specific or charged modes. The PDF for CP-eigenstate modes is

$$\begin{aligned} \mathcal{P}_k(\vec{x}) &= \mathcal{P}^l_k(\Delta E, M_{\rm bc}; m',\theta';\Delta t, q_{\rm tag}) \\ &= \mathcal{F}^l_k\cdot\mathcal{P}_k(\Delta E, M_{\rm bc})\cdot\mathcal{P}_k(m',\theta')\cdot\mathcal{P}^l_k(\Delta t, q_{\rm tag})\ , \end{aligned} \tag{6.39}$$

where $\mathcal{F}^l_k$, $\mathcal{P}_k(\Delta E, M_{\rm bc})$, $\mathcal{P}_k(m',\theta')$, and $\mathcal{P}_k(\Delta t, q_{\rm tag})$ are the event fraction in l-th r-region, a ΔE-$M_{\rm bc}$ PDF, a Dalitz plot PDF, and a time-dependent CP-violation PDF, respectively.

[3]This is because there are decay modes such as $B^0\to D^{(*)-}\pi^+$, where the π^+ has high momentum.

They are normalized as

$$\sum_l \mathcal{F}_k^l = 1 , \tag{6.40}$$

$$\iint_{\rm SR} d\Delta E\, dM_{\rm bc}\, \mathcal{P}_k(\Delta E, M_{\rm bc}) = 1 , \tag{6.41}$$

$$\iint_{\rm SDP,\ Veto} dm'\, d\theta'\, \mathcal{P}_k(m', \theta') = 1 , \tag{6.42}$$

$$\sum_{q_{\rm tag}} \int d\Delta t\, \mathcal{P}_k^l(\Delta t, q_{\rm tag}) = 1 . \tag{6.43}$$

For flavor-specific or charged modes, the PDF is

$$\begin{aligned} \mathcal{P}_k(\vec{x}) &= \mathcal{P}_k^l(\Delta E, M_{\rm bc}; m', \theta'; \Delta t, q_{\rm tag}) \\ &= \mathcal{F}_k^l \cdot \mathcal{P}_k(\Delta E, M_{\rm bc}) \sum_{q_{\rm flv}} \mathcal{P}_k(m', \theta'; q_{\rm flv}) \cdot \mathcal{P}_k^l(\Delta t, q_{\rm tag}, q_{\rm flv}) , \end{aligned} \tag{6.44}$$

where the Dalitz plot PDF $\mathcal{P}_k(m', \theta'; q_{\rm flv})$ is dependent on the flavor of the CP (fully reconstructed) side B, $q_{\rm flv}$, and $\mathcal{P}_k^l(\Delta t, q_{\rm tag}, q_{\rm flv})$ is a mixing PDF (lifetime PDF with flavor asymmetry) for flavor-specific (charged) modes. They are normalized as

$$\iint_{\rm SDP,\ Veto} dm'\, d\theta'\, \mathcal{P}_k(m', \theta'; q_{\rm flv}) = 1 \quad (\forall q_{\rm flv}) , \tag{6.45}$$

$$\sum_{q_{\rm tag}} \sum_{q_{\rm flv}} \int d\Delta t\, \mathcal{P}_k^l(\Delta t, q_{\rm tag}, q_{\rm flv}) = 1 . \tag{6.46}$$

Note that the normalization conditions above lead to the following relation

$$\sum_l \sum_{q_{\rm tag}} \int d\Delta t \iint_{\rm SR} d\Delta E\, dM_{\rm bc} \iint_{\rm SDP,\ Veto} dm'\, d\theta'\, \mathcal{P}_{B\overline{B}}(\vec{x}) = 1 . \tag{6.47}$$

The ΔE-$M_{\rm bc}$ PDF and Dalitz plot PDF are obtained mode-by-mode from MC. We assume the Dalitz plot PDF's of the CP-eigenstate modes to satisfy following symmetry

$$\mathcal{P}_k(m', \theta') = \mathcal{P}_k(m', 1 - \theta') , \tag{6.48}$$

and those of flavor-specific and charged modes to satisfy

$$\mathcal{P}_k(m', \theta'; q_{\rm flv} = +1) = \mathcal{P}_k(m', 1 - \theta'; q_{\rm flv} = -1) . \tag{6.49}$$

3 Unbinned Maximum Likelihood Fit and the Result

With the PDF defined above, we form the likelihood function

$$\mathcal{L} \equiv \prod_i P(\vec{x}_i) , \tag{6.50}$$

where i is an index over events. We perform an unbinned-maximum-likelihood fit and determine the 26 Dalitz plot parameters using the likelihood function with the signal fraction and the lineshape parameters obtained in Sec. 1 and chapter 5, respectively.

A fit to the 2,824 events in the signal region yields the result listed in Table 6.2. The correlation matrix of the 26 parameters after combining statistical and systematic errors is

shown in tables 6.3-6.5. Figure 6.5 shows the projections of the square Dalitz plot in data with the fit result superimposed. We also show the mass and helicity distributions for each $\rho\pi$ enhanced region along with projections of the fit (Fig. 6.6). Figure 6.7 shows the Δt distributions and background subtracted asymmetries. We define the asymmetry in each Δt bin by $(N_+ - N_-)/(N_+ + N_-)$, where $N_{+(-)}$ corresponds the background subtracted number of events with $q_{\rm tag} = +1(-1)$. The $\rho^-\pi^+$ enhanced region shows a significant asymmetry, corresponding to a non-zero value of U_-^-. Note that this is not a CP-violating effect, since $\rho^-\pi^+$ is not a CP-eigenstate. No sin-like asymmetry is observed in any of the regions.

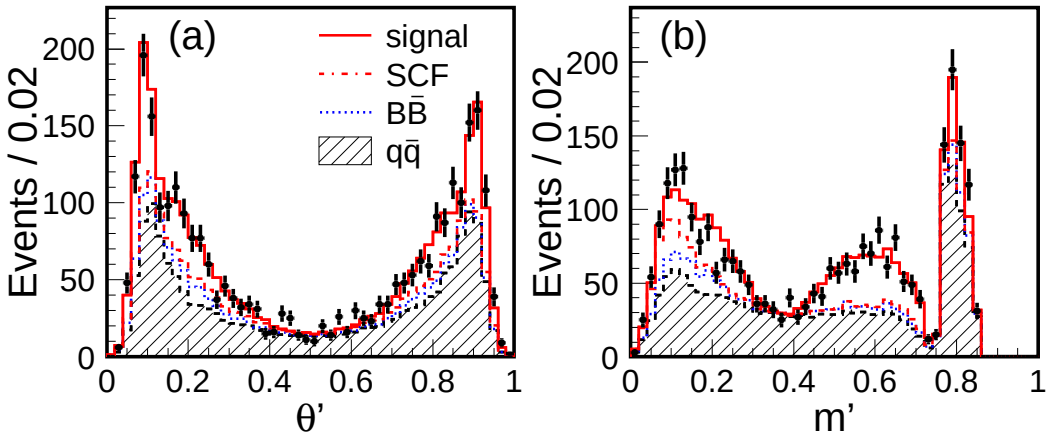


Figure 6.5: Distributions of (a) θ' and (b) m' with fit results. The histograms are cumulative. Solid, dot-dashed, dotted and dashed hatched histograms correspond to correctly reconstructed signal, SCF, $B\overline{B}$, and continuum PDFs, respectively.

Treatment of statistical errors

With a toy MC study, we check the pull distribution, where the pull is defined as the residual divided by the MINOS error. Here, the MINOS error, which corresponds to the deviation from the best fit parameter when $-2\ln(\mathcal{L}/\mathcal{L}_{\rm max})$ is changed by one, is an estimate of the statistical error. Although the pull is expected to follow a Gaussian distribution with unit width, we find that the width of the pull distribution tends to be significantly larger than one for the interference terms due to small statistics. To restore the pull width to unity, we multiply the MINOS errors of the interference terms by a factor of 1.17, which is the average pull width for the interference terms obtained above, and quote the results as the statistical errors. For the non-interfering terms, we quote the MINOS errors without the correction factor.

Qualitative interpretation of the resultant parameters

Here, we qualitatively discuss the meaning of the fit result. As can be seen in Table 6.2, most of the parameters corresponding to flavor asymmetry, U_κ^-, $U_{\kappa\sigma}^{-,\rm Im(Re)}$, I_κ, and $I_{\kappa\sigma}^{\rm Im(Re)}$, are zero consistent. This implies $\phi_2 \sim 90^\circ$ or 180° as described in the following.

As can be found in appendix B, amplitudes $A^\pm$ and $\overline{A}^\pm$ are related to ϕ_2 as

$$\arg\left(A^{\pm *}\overline{A}^\pm\right) = \arg\left(e^{2i\phi_2}A^{\pm *}\widetilde{A}^\pm\right) = 2\phi_2 + \theta_\pm \ . \tag{6.51}$$

Here, $\theta_\pm$ is not helicity but the phase related to strong interaction. (See Fig. B.1 and Eq. (B.18))

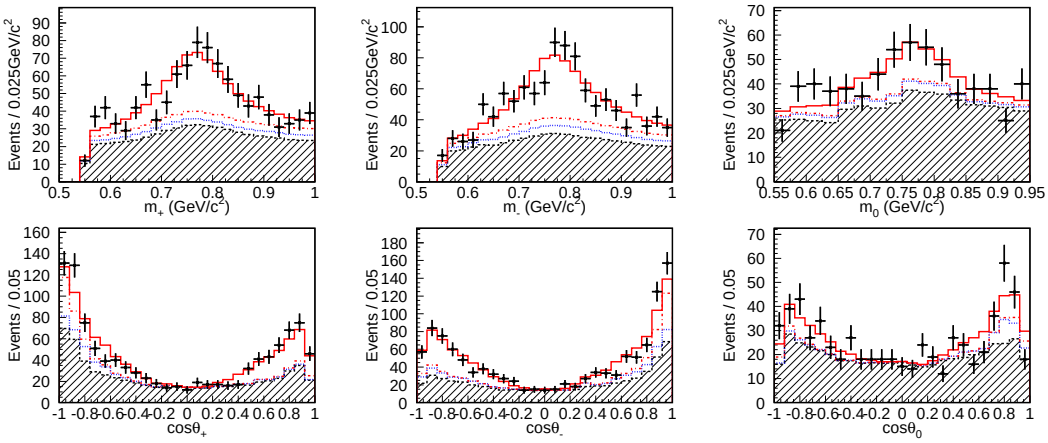


Figure 6.6: Mass (upper) and helicity (lower) distribution of $\rho^+\pi^-$ (left), $\rho^-\pi^+$ (middle), and $\rho^0\pi^0$ (right) enhanced regions. The histograms are cumulative. Solid, dot-dashed, dotted and dashed hatched histograms correspond to correctly reconstructed signal, SCF, $B\overline{B}$, and continuum PDFs, respectively.

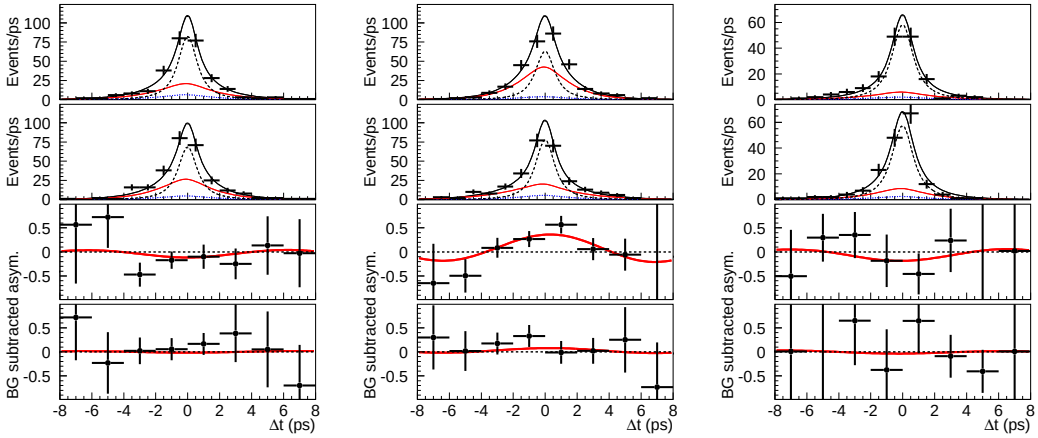


Figure 6.7: Proper time distributions of good tag ($r > 0.5$) regions for $q_{\rm tag} = +1$ (upper) and $q_{\rm tag} = -1$ (middle upper), in $\rho^+\pi^-$ (left), $\rho^-\pi^+$ (middle), and $\rho^0\pi^0$ (right) enhanced regions, where solid (red), dotted, and dashed curves correspond to signal, continuum, and $B\overline{B}$ PDFs. The middle lower and lower plots show the background subtracted asymmetries in the good tag ($r > 0.5$) and poor tag ($r < 0.5$) regions, respectively. The significant asymmetry in the $\rho^-\pi^+$ enhanced region (middle) corresponds to a non-zero value of U_-^-.

Table 6.2: Result of the time-dependent Dalitz plot fit. First and second errors in the middle column correspond to the statistical and systematic errors, respectively. The right column describes which term the coefficient corresponds to, omitting constant factors, a common factor of $e^{-|\Delta t|/\tau_{B^0}}$, and the effects of wrong tag fraction and Δt resolution function.

Name	Value	Coefficient of
U_+^+	$+1$ (fixed)	$\lvert f_+\rvert^2$
U_-^+	$+1.27 \pm 0.13 \pm 0.09$	$\lvert f_-\rvert^2$
U_0^+	$+0.29 \pm 0.05 \pm 0.04$	$\lvert f_0\rvert^2$
$U_{+-}^{+,\mathrm{Re}}$	$+0.49 \pm 0.86 \pm 0.52$	$\mathrm{Re}[f_+ f_-^*]$
$U_{+0}^{+,\mathrm{Re}}$	$+0.29 \pm 0.50 \pm 0.35$	$\mathrm{Re}[f_+ f_0^*]$
$U_{-0}^{+,\mathrm{Re}}$	$+0.25 \pm 0.60 \pm 0.33$	$\mathrm{Re}[f_- f_0^*]$
$U_{+-}^{+,\mathrm{Im}}$	$+1.18 \pm 0.86 \pm 0.34$	$\mathrm{Im}[f_+ f_-^*]$
$U_{+0}^{+,\mathrm{Im}}$	$-0.57 \pm 0.35 \pm 0.51$	$\mathrm{Im}[f_+ f_0^*]$
$U_{-0}^{+,\mathrm{Im}}$	$-1.34 \pm 0.60 \pm 0.47$	$\mathrm{Im}[f_- f_0^*]$
U_+^-	$+0.23 \pm 0.15 \pm 0.07$	$q_{\mathrm{tag}} \cos(\Delta m_d \Delta t) \cdot \lvert f_+\rvert^2$
U_-^-	$-0.62 \pm 0.16 \pm 0.08$	$q_{\mathrm{tag}} \cos(\Delta m_d \Delta t) \cdot \lvert f_-\rvert^2$
U_0^-	$+0.15 \pm 0.11 \pm 0.08$	$q_{\mathrm{tag}} \cos(\Delta m_d \Delta t) \cdot \lvert f_0\rvert^2$
$U_{+-}^{-,\mathrm{Re}}$	$-1.18 \pm 1.61 \pm 0.72$	$q_{\mathrm{tag}} \cos(\Delta m_d \Delta t) \cdot \mathrm{Re}[f_+ f_-^*]$
$U_{+0}^{-,\mathrm{Re}}$	$-2.37 \pm 1.36 \pm 0.60$	$q_{\mathrm{tag}} \cos(\Delta m_d \Delta t) \cdot \mathrm{Re}[f_+ f_0^*]$
$U_{-0}^{-,\mathrm{Re}}$	$-0.53 \pm 1.44 \pm 0.65$	$q_{\mathrm{tag}} \cos(\Delta m_d \Delta t) \cdot \mathrm{Re}[f_- f_0^*]$
$U_{+-}^{-,\mathrm{Im}}$	$-2.32 \pm 1.74 \pm 0.91$	$q_{\mathrm{tag}} \cos(\Delta m_d \Delta t) \cdot \mathrm{Im}[f_+ f_-^*]$
$U_{+0}^{-,\mathrm{Im}}$	$-0.41 \pm 1.00 \pm 0.47$	$q_{\mathrm{tag}} \cos(\Delta m_d \Delta t) \cdot \mathrm{Im}[f_+ f_0^*]$
$U_{-0}^{-,\mathrm{Im}}$	$-0.02 \pm 1.31 \pm 0.83$	$q_{\mathrm{tag}} \cos(\Delta m_d \Delta t) \cdot \mathrm{Im}[f_- f_0^*]$
I_+	$-0.01 \pm 0.11 \pm 0.04$	$q_{\mathrm{tag}} \sin(\Delta m_d \Delta t) \cdot \lvert f_+\rvert^2$
I_-	$+0.09 \pm 0.10 \pm 0.04$	$q_{\mathrm{tag}} \sin(\Delta m_d \Delta t) \cdot \lvert f_-\rvert^2$
I_0	$+0.02 \pm 0.09 \pm 0.05$	$q_{\mathrm{tag}} \sin(\Delta m_d \Delta t) \cdot \lvert f_0\rvert^2$
I_{+-}^{Re}	$+1.21 \pm 2.59 \pm 0.98$	$q_{\mathrm{tag}} \sin(\Delta m_d \Delta t) \cdot \mathrm{Im}[f_+ f_-^*]$
I_{+0}^{Re}	$+1.15 \pm 2.26 \pm 0.92$	$q_{\mathrm{tag}} \sin(\Delta m_d \Delta t) \cdot \mathrm{Im}[f_+ f_0^*]$
I_{-0}^{Re}	$-0.92 \pm 1.34 \pm 0.80$	$q_{\mathrm{tag}} \sin(\Delta m_d \Delta t) \cdot \mathrm{Im}[f_- f_0^*]$
I_{+-}^{Im}	$-1.93 \pm 2.39 \pm 0.89$	$q_{\mathrm{tag}} \sin(\Delta m_d \Delta t) \cdot \mathrm{Re}[f_+ f_-^*]$
I_{+0}^{Im}	$-0.40 \pm 1.86 \pm 0.85$	$q_{\mathrm{tag}} \sin(\Delta m_d \Delta t) \cdot \mathrm{Re}[f_+ f_0^*]$
I_{-0}^{Im}	$-2.03 \pm 1.62 \pm 0.81$	$q_{\mathrm{tag}} \sin(\Delta m_d \Delta t) \cdot \mathrm{Re}[f_- f_0^*]$

Table 6.3: Correlation matrix (1) of the 26 fitted parameters, with statistical and systematic errors combined.

	U_-^+	U_0^+	$U_{+-}^{+,\mathrm{Re}}$	$U_{+0}^{+,\mathrm{Re}}$	$U_{-0}^{+,\mathrm{Re}}$	$U_{+-}^{+,\mathrm{Im}}$	$U_{+0}^{+,\mathrm{Im}}$	$U_{-0}^{+,\mathrm{Im}}$
U_-^+	+1.00							
U_0^+	+0.22	+1.00						
$U_{+-}^{+,\mathrm{Re}}$	+0.06	+0.04	+1.00					
$U_{+0}^{+,\mathrm{Re}}$	+0.10	+0.02	+0.02	+1.00				
$U_{-0}^{+,\mathrm{Re}}$	−0.04	−0.11	+0.01	+0.01	+1.00			
$U_{+-}^{+,\mathrm{Im}}$	+0.08	+0.03	+0.12	+0.02	−0.00	+1.00		
$U_{+0}^{+,\mathrm{Im}}$	−0.03	−0.08	−0.00	+0.13	+0.02	−0.00	+1.00	
$U_{-0}^{+,\mathrm{Im}}$	−0.14	−0.08	−0.02	−0.02	+0.10	−0.01	+0.01	+1.00
U_+^-	+0.05	+0.02	+0.00	−0.02	+0.00	−0.02	−0.01	−0.01
U_-^-	−0.23	−0.08	−0.03	−0.04	−0.02	−0.03	+0.01	+0.03
U_0^-	+0.05	+0.10	+0.01	+0.00	−0.04	+0.01	−0.06	−0.08
$U_{+-}^{-,\mathrm{Re}}$	−0.03	−0.01	−0.03	−0.00	−0.00	−0.04	+0.00	+0.01
$U_{+0}^{-,\mathrm{Re}}$	−0.04	−0.01	−0.01	−0.12	−0.00	+0.00	−0.01	+0.02
$U_{-0}^{-,\mathrm{Re}}$	−0.02	−0.04	−0.00	−0.00	+0.06	−0.00	+0.01	+0.08
$U_{+-}^{-,\mathrm{Im}}$	−0.04	−0.02	−0.05	−0.01	−0.00	+0.00	+0.00	+0.01
$U_{+0}^{-,\mathrm{Im}}$	−0.03	−0.09	−0.01	−0.01	+0.02	−0.00	−0.04	+0.01
$U_{-0}^{-,\mathrm{Im}}$	+0.01	−0.02	+0.00	+0.00	+0.00	−0.00	+0.00	−0.25
I_+	+0.00	+0.00	−0.02	−0.01	−0.00	−0.05	−0.01	−0.00
I_-	+0.06	+0.03	−0.01	+0.01	−0.02	+0.05	−0.00	+0.04
I_0	+0.01	+0.01	+0.00	+0.02	−0.00	+0.00	−0.02	−0.02
I_{+-}^{Re}	−0.04	−0.01	+0.01	−0.00	+0.00	−0.16	+0.00	+0.00
I_{+0}^{Re}	+0.00	+0.02	+0.00	−0.13	−0.01	−0.00	−0.00	−0.00
I_{-0}^{Re}	−0.06	+0.01	−0.01	−0.01	−0.12	+0.00	−0.01	−0.29
I_{+-}^{Im}	−0.02	−0.01	+0.13	−0.00	+0.00	+0.00	+0.00	+0.00
I_{+0}^{Im}	−0.01	−0.03	+0.00	+0.00	+0.01	+0.01	+0.04	+0.01
I_{-0}^{Im}	−0.06	−0.04	−0.01	−0.02	−0.09	−0.02	+0.01	+0.08

Table 6.4: Correlation matrix (2) of the 26 fitted parameters, with statistical and systematic errors combined.

	U_+^-	U_-^-	U_0^-	$U_{+-}^{-,\mathrm{Re}}$	$U_{+0}^{-,\mathrm{Re}}$	$U_{-0}^{-,\mathrm{Re}}$	$U_{+-}^{-,\mathrm{Im}}$	$U_{+0}^{-,\mathrm{Im}}$	$U_{-0}^{-,\mathrm{Im}}$
U_+^-	+1.00								
U_-^-	−0.06	+1.00							
U_0^-	+0.00	−0.01	+1.00						
$U_{+-}^{-,\mathrm{Re}}$	−0.07	+0.01	−0.00	+1.00					
$U_{+0}^{-,\mathrm{Re}}$	−0.21	+0.03	−0.08	+0.02	+1.00				
$U_{-0}^{-,\mathrm{Re}}$	+0.01	−0.12	−0.16	−0.00	+0.02	+1.00			
$U_{+-}^{-,\mathrm{Im}}$	+0.03	+0.03	−0.00	+0.20	−0.01	−0.00	+1.00		
$U_{+0}^{-,\mathrm{Im}}$	−0.02	+0.01	−0.03	+0.00	+0.01	+0.01	+0.00	+1.00	
$U_{-0}^{-,\mathrm{Im}}$	+0.00	−0.03	+0.02	−0.00	−0.00	+0.14	−0.00	+0.01	+1.00
I_+	−0.02	−0.01	+0.00	+0.03	+0.01	+0.00	−0.02	+0.01	+0.00
I_-	−0.00	−0.01	+0.01	+0.03	−0.00	−0.05	+0.00	−0.01	−0.07
I_0	+0.00	−0.01	+0.07	−0.00	−0.02	+0.02	−0.00	−0.05	−0.06
I_{+-}^{Re}	+0.02	+0.01	−0.00	−0.02	−0.00	+0.00	−0.15	+0.00	+0.00
I_{+0}^{Re}	−0.01	+0.00	+0.02	+0.00	+0.09	−0.01	−0.00	+0.16	+0.01
I_{-0}^{Re}	−0.01	+0.08	+0.08	+0.01	−0.00	−0.12	+0.01	−0.00	+0.21
I_{+-}^{Im}	+0.02	+0.04	−0.00	+0.04	−0.00	−0.00	+0.04	+0.00	−0.00
I_{+0}^{Im}	+0.03	+0.00	−0.03	−0.01	−0.28	+0.00	+0.01	−0.03	+0.01
I_{-0}^{Im}	−0.00	+0.01	−0.01	−0.00	+0.01	+0.18	+0.00	+0.02	+0.11

Table 6.5: Correlation matrix (3) of the 26 fitted parameters, with statistical and systematic errors combined.

	I_+	I_-	I_0	I_{+-}^{Re}	I_{+0}^{Re}	I_{-0}^{Re}	I_{+-}^{Im}	I_{+0}^{Im}	I_{-0}^{Im}
I_+	+1.00								
I_-	−0.06	+1.00							
I_0	+0.00	+0.01	+1.00						
I_{+-}^{Re}	−0.04	−0.06	−0.00	+1.00					
I_{+0}^{Re}	+0.04	−0.00	−0.14	−0.00	+1.00				
I_{-0}^{Re}	−0.02	+0.21	+0.01	−0.01	+0.00	+1.00			
I_{+-}^{Im}	−0.07	−0.01	−0.00	−0.35	−0.00	+0.00	+1.00		
I_{+0}^{Im}	−0.15	+0.01	−0.09	+0.01	−0.23	−0.00	+0.01	+1.00	
I_{-0}^{Im}	+0.01	−0.14	−0.23	+0.01	+0.04	−0.06	+0.00	+0.03	+1.00

On the other hand, I_κ is related to $A^{\pm *}\overline{A}^{\pm}$ by the definition in Eq. (4.26):

$$I_\kappa \sim \mathrm{Im}\left[\overline{A}^\kappa A^{\kappa *}\right] \ .$$

Consequently, the zero consistent values of $I_\pm$ indicates

$$\sin(2\phi_2 + \theta_\pm) \sim 0 \tag{6.52}$$

and thus

$$2\phi_2 + \theta_\pm \sim 180^\circ \text{ or } 360^\circ. \tag{6.53}$$

Our fit result favors $\theta_\pm \sim 0$ as follows. For simplicity, we suppose an extreme case where all of U_κ^- and $U_{\kappa\sigma}^{-,\mathrm{Im(Re)}}$ are zero. With Eqs. (4.25) and (4.27), $U_\kappa^- = 0$ leads to

$$|A^\kappa| = |\overline{A}^\kappa| \ . \tag{6.54}$$

and $U_{\kappa\sigma}^{-,\mathrm{Im(Re)}} = 0$ leads to

$$A^\kappa A^{\sigma *} = \overline{A}^\kappa \overline{A}^{\sigma *} \ . \tag{6.55}$$

Combining Eqs. (6.54) and (6.55) with the isospin relation of Eq. (B.16), we obtain

$$A^\kappa = \widetilde{A}^\kappa \ , \tag{6.56}$$

which corresponds to

$$\theta_\pm = 0 \tag{6.57}$$

since

$$\theta_\pm \equiv \arg\left(A^{\pm *}\widetilde{A}^{\pm}\right) \ . \tag{6.58}$$

Though some of the U_κ^- and $U_{\kappa\sigma}^{-,\mathrm{Im(Re)}}$ in Table 6.2 are significantly non-zero, the situation is basically similar to this and $\theta_\pm \sim 0$ is favored; Fig. 6.8 shows the constraint on $\theta_\pm$ as the results of the full Dalitz+Pentagon analysis performed in Sec. 2 of chapter 7, where $\theta_\pm \sim 0$ corresponds to small $\Delta\chi^2$.

With $\theta_\pm \sim 0$ and $2\phi_2 + \theta_\pm \sim 180^\circ$ or 360°, our result favors $\phi_2 \sim 90^\circ$ or 180°.

4 Systematic Uncertainty

Tables 6.6-6.8 list the systematic errors for the 26 time-dependent Dalitz plot parameters. The total systematic error is obtained by adding each source of systematic uncertainty in quadrature.

Radial excitations (ρ' and ρ'')

The largest contribution for the interference terms tends to come from radial excitations ($\rho(1450)$ and $\rho(1700)$, or ρ' and ρ''). The systematic error related to the radial excitations can be categorized into three classes: 1) uncertainties coming from the lineshape variation, i.e., the lineshape difference between each decay amplitude, 2) uncertainties in external parameters, $m_{\rho(1450)}$, $\Gamma_{\rho(1450)}$, $m_{\rho(1700)}$, $\Gamma_{\rho(1700)}$, and 3) uncertainties in the common lineshape parameters β and γ used for the nominal fit.

In our nominal fit, we assume all of 6 decay amplitudes have the same contribution from $\rho(1450)$ and $\rho(1700)$, i.e., we assume Eq. (6.7). This assumption, however, is not well grounded. As described in Sec. 1, in general, the contributions from $\rho(1450)$ and $\rho(1700)$, i.e., $^{(}\overline{\beta)}_\kappa$ and

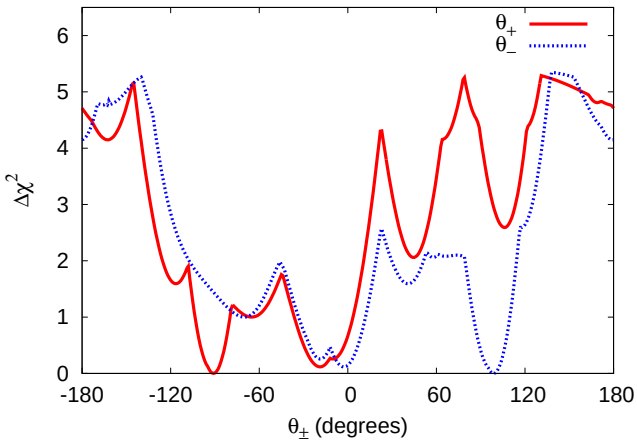


Figure 6.8: $\Delta\chi^2 \equiv \chi^2(\theta_\pm) - \chi^2_{\rm min}$ vs. $\theta_\pm$ as a result of Dalitz+Pentagon analysis, the detail of which can be found in Sec. 2 of chapter 7. The regions with small $\Delta\chi^2$ correspond to the favored region for $\theta_\pm$. As can be seen in the figure, $\theta_\pm = 0$ is included in the allowed region.

$^{(}\overline{\gamma}^{)}_\kappa$, can be different for each of the decay amplitudes and thus the systematic uncertainty from this assumption must be addressed. For this purpose, we rewrite Eq. (6.7) as

$$^{(}\overline{F}^{)\kappa}_\pi(s) = BW^{GS}_{\rho(770)}(s) + (\beta + \Delta^{(}\overline{\beta}^{)}_\kappa) BW^{GS}_{\rho(1450)}(s) + (\gamma + \Delta^{(}\overline{\gamma}^{)}_\kappa) BW^{GS}_{\rho(1700)}(s)\ . \qquad (6.59)$$

The variation of the contributions from radial excitations is described by non-zero $\Delta^{(}\overline{\beta}^{)}_\kappa$ and $\Delta^{(}\overline{\gamma}^{)}_\kappa$, which are 12 complex variables. We generate various toy MC samples, where the input A^κ and $\overline{A}^\kappa$ are fixed but the values of $\Delta^{(}\overline{\beta}^{)}_\kappa$ and $\Delta^{(}\overline{\gamma}^{)}_\kappa$ are randomly varied according to the constraints on $\Delta^{(}\overline{\beta}^{)}_\kappa$ and $\Delta^{(}\overline{\gamma}^{)}_\kappa$; these constraints are obtained from the results in Sec. 1-1 of Sec. 1, and are combined with the isospin relation [17, 18], which improves the constraints. The statistics for each pseudo-experiment are set to be large enough so that the statistical uncertainty is negligible. We assign the variations and the biases of the fit results due to the $\Delta^{(}\overline{\beta}^{)}_\kappa$ and $\Delta^{(}\overline{\gamma}^{)}_\kappa$ variation as systematic errors.

For the masses and widths of the $\rho(1450)$ and $\rho(1700)$, we quote the values from the PDG [52]. To estimate the systematic error coming from uncertainties in their parameters, we generate toy MC varying the input masses and widths and fit them with the masses and widths of the nominal fit. Here, we vary the masses and widths by twice the PDG error, $\pm 50\,{\rm MeV}/c^2$ and $\pm 40\,{\rm MeV}/c^2$ ($\pm 120\,{\rm MeV}/c^2$ and $\pm 200\,{\rm MeV}/c^2$) for $\rho(1450)$ and $\rho(1700)$ masses (widths), respectively. This is because the variations between independent experiments are much larger than the 1σ PDG errors. We quote the mean shift of the Toy MC ensemble as the systematic errors. The contribution from the uncertainties of the mass and width of $\rho(770)$ is also estimated in the same manner.

We perform the toy MC to take account of the systematic errors from the uncertainties in β and γ for the nominal fit, determined in Sec. 1-1, in the same way. Here, the correlation among the four degrees of freedom of (β, γ) is significant and thus we treat the correlation properly, as described in Sec. 2 of appendix H.

SCF

Systematic errors due to SCF are dominated by the uncertainty in the difference between data and MC; these errors are determined from the $B \to D^{(*)}\rho$ control samples that contain a single

π^0 in the final state [53]. We generate toy MC samples varying the amount of SCF's by their 1σ errors, which are $\pm 100\%$ for the CR SCF and $^{+30(60)}_{-30}\%$ for the NR SCF in the DS-I (DS-II); fit the toy MC samples in the same condition as nominal fit; and quote the mean differences of the fit results from the default.

We also take account of the uncertainties from the event fraction for each r region (for CR and NR), the wrong tag fractions (for CR) and lifetime used in the Δt PDF (for CR), which are obtained from MC. We vary these parameters in the data fit and quote differences from the nominal fit as systematic errors.

Signal Dalitz PDF

Systematic errors due to the Dalitz PDF for signal is mainly from the Dalitz plot dependent efficiency. We take account of MC statistics in the efficiency[4] and uncertainty in the π^0 momentum dependent efficiency correction, $\epsilon'(p_{\pi^0})$, obtained from the control samples of the decay modes $\overline{B}^0 \to \rho^- D^{(*)+}$, $\overline{B}^0 \to \pi^- D^{(*)+}$, $B^- \to \rho^- D^{(*)0}$ and $B^- \to \pi^- D^{(*)0}$, where we choose the D subdecay modes such that a single π^0 is included in the final states. The Dalitz plot efficiency obtained from MC is found to have a small charge asymmetry ($\sim 3\%$ at most). We generate toy MC samples with and without the asymmetry, fit them using the nominal PDF with the asymmetry, and quote twice the difference as the systematic error. The Dalitz plot efficiency is r region dependent and obtained as a product with the event fraction in the corresponding region, $\mathcal{F}^l_{\rm sig} \cdot \epsilon^l(m', \theta')$, using MC. The difference in the fraction, $\mathcal{F}^l_{\rm sig}$, for data and MC is estimated to be $\sim 10\%$ at most, using the $B^0 \to D^{*-}\pi^+$ decay mode as a control sample. We vary the fraction in each r region by $\pm 10\%$ and fit the data to estimate the systematic error.

Background Dalitz PDF

The Dalitz plot for continuum background has an uncertainty due to the limited statistics of the sideband events, which we use to model the PDF. We estimate the uncertainty by performing a Toy MC study of sideband events[5]. Systematic uncertainty from the statistics of the $B\overline{B}$ MC, which is used to model the $B\overline{B}$ Dalitz plot PDF, is also taken into account[6].

$B^0 \to \pi^+\pi^-\pi^0$ processes other than $B^0 \to (\rho\pi)^0$

Large contributions to the systematic errors for the non-interfering parameters tends to come from the $B^0 \to \pi^+\pi^-\pi^0$ decay processes that are not $B^0 \to (\rho\pi)^0$. We take account of the contributions from $B^0 \to f_0(980)\pi^0$, $B^0 \to f_0(600)\pi^0$, $B^0 \to \omega\pi^0$, and non-resonant $B^0 \to \pi^+\pi^-\pi^0$. Using the 1σ upper limits[7] as input, we generate toy MC for each mode with the interference between the $B^0 \to (\rho\pi)^0$ and the other $B^0 \to \pi^+\pi^-\pi^0$ mode taken into account. We obtain the systematic error by fitting the Toy MC assuming $B^0 \to (\rho\pi)^0$ only in the PDF. Within physically allowed regions, we vary the CP-violation parameters of the other $B^0 \to \pi^+\pi^-\pi^0$ modes and the relative phase difference between $B^0 \to (\rho\pi)^0$ and the other $B^0 \to \pi^+\pi^-\pi^0$ as the input parameters, and use the largest deviation as the systematic error for each decay mode.

[4]The detailed description on the estimation of the statistical fluctuation can be found in Sec. 1-2 of appendix H.

[5]The detail of the procedure can be found in Sec. 1-1 of appendix H.

[6]The detailed description on the estimation of the statistical fluctuation can be found in Sec. 1-2 of appendix H.

[7]The procedure to obtain the upper limits can be found in appendix G.

Background fraction

Systematic errors due to the event-by-event ΔE-$M_{\rm bc}$ background fractions are studied by varying the PDF shape parameters and the fraction of continuum background, and the correction factor to the signal PDF shape by $\pm 1\sigma$. We also vary the fractions of the $B\overline{B}$ background, which are estimated with MC, by $\pm 50\%$ ($\pm 20\%$) for the $b \to c$ ($b \to u$) transition. We assume smaller uncertainty for the $b \to u$ category since the branching fractions of the decay modes in this category are better measured than those of the $b \to c$ category.

Physics parameters

We use the world average [52, 54, 55] for the following physics parameters: τ_{B^0} and Δm_d (used for the Δt PDF's of signal and $B\overline{B}$ background), the CKM angles of ϕ_1 and ϕ_2 (used in $B\overline{B}$ background), and the branching fractions of $b \to u$ decay modes (used in $B\overline{B}$ background). The systematic error is assigned by varying them by $\pm 1\sigma$. The charge asymmetry of $B^0 \to a_1^{\pm}\pi^{\mp}$, for which there is no measurement and we use zero in the nominal fit, is varied in the physically allowed region, i.e., ± 1.

Background Δt PDF

Systematic errors from uncertainties in the background Δt shapes for both continuum and $B\overline{B}$ backgrounds are estimated by varying each parameter by $\pm 1\sigma$.

Vertex Reconstruction

To determine the systematic error that arises from uncertainties in the vertex reconstruction, the track and vertex selection criteria are varied to search for possible systematic biases. Systematic error due to the IP constraint in the vertex reconstruction is estimated by varying the smearing used to account for the B flight length by $\pm 10\,\mu m$.

Resolution Function for the Δt PDF and Flavor Tagging

Systematic errors due to uncertainties in the resolution function are estimated by varying each resolution parameter obtained from data (MC) by $\pm 1\sigma$ ($\pm 2\sigma$). Systematic errors due to uncertainties in the wrong tag fractions are also studied by varying the wrong tag fraction individually for each r region.

Fit Bias

We observed fit bias due to small statistics for some of the fitted parameters. Since this bias is much smaller than the statistical error, we take it into account in the systematic errors. We estimate the size of the fit bias by toy MC study and quote the bias as the systematic errors. We also confirm that the bias is consistent between toy MC and full detector MC simulation.

Tag-side interference

Finally, we investigate the effects of tag-side interference (TSI), which is the interference between CKM-favored and CKM-suppressed $B \to D$ transitions in the $f_{\rm tag}$ final state [56]. A small correction to the PDF for the signal distribution arises from the interference. We estimate the size of the correction using the $B^0 \to D^{*-}\ell^+\nu$ sample. We then generate MC pseudo-experiments and make an ensemble test to obtain the systematic biases.

Table 6.6: Table of systematic errors (1). The notation "< 0.01" means that the value is small and less than 0.01, and thus invisible in the number of significant digits shown here. We calculate the total systematic error including these small contributions.

	U_-^+	U_0^+	$U_{+-}^{+,\mathrm{Re}}$	$U_{+0}^{+,\mathrm{Re}}$	$U_{-0}^{+,\mathrm{Re}}$	$U_{+-}^{+,\mathrm{Im}}$	$U_{+0}^{+,\mathrm{Im}}$	$U_{-0}^{+,\mathrm{Im}}$
ρ' and ρ''	0.01	0.01	0.31	0.19	0.19	0.21	0.39	0.30
SCF	0.01	0.02	0.31	0.09	0.11	0.11	0.12	0.11
Signal Dalitz	0.06	0.01	0.15	0.20	0.18	0.13	0.10	0.10
BG Dalitz	0.02	0.01	0.17	0.11	0.11	0.14	0.10	0.19
Other $\pi\pi\pi$	0.04	0.02	0.06	0.08	0.07	0.10	0.09	0.07
BG fraction	0.02	0.01	0.08	0.04	0.07	0.06	0.03	0.11
Physics	0.02	< 0.01	0.01	0.01	0.02	0.01	0.01	0.01
BG Δt	< 0.01	< 0.01	0.03	0.01	0.01	0.02	0.01	0.01
Vertexing	0.03	0.01	0.03	0.05	0.02	0.09	0.05	0.07
Resolution	< 0.01	< 0.01	0.03	0.05	0.02	0.02	0.02	0.03
Flavor tagging	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.01
Fit bias	0.02	0.02	0.10	0.11	0.07	0.06	0.24	0.22
TSI	< 0.01	< 0.01	0.01	0.02	0.02	0.02	0.01	< 0.01
Total	0.09	0.04	0.52	0.35	0.33	0.34	0.51	0.47

Table 6.7: Table of systematic errors (2). The notation "< 0.01" means that the value is small and less than 0.01, and thus invisible in the number of significant digits shown here. We calculate the total systematic error including these small contributions.

	U_+^-	U_-^-	U_0^-	$U_{+-}^{-,\mathrm{Re}}$	$U_{+0}^{-,\mathrm{Re}}$	$U_{-0}^{-,\mathrm{Re}}$	$U_{+-}^{-,\mathrm{Im}}$	$U_{+0}^{-,\mathrm{Im}}$	$U_{-0}^{-,\mathrm{Im}}$
ρ' and ρ''	0.01	0.02	0.04	0.53	0.29	0.42	0.70	0.31	0.59
SCF	0.02	0.02	0.02	0.09	0.17	0.17	0.13	0.09	0.18
Signal Dalitz	0.01	0.02	0.01	0.27	0.20	0.14	0.30	0.15	0.19
BG Dalitz	0.04	0.03	0.02	0.28	0.32	0.22	0.30	0.20	0.30
Other $\pi\pi\pi$	0.03	0.03	0.02	0.07	0.08	0.12	0.13	0.08	0.08
BG fraction	0.02	0.04	0.01	0.18	0.17	0.14	0.22	0.13	0.11
Physics	0.01	0.01	<0.01	0.03	0.03	0.03	0.04	0.01	0.04
BG Δt	<0.01	<0.01	<0.01	0.02	0.03	0.02	0.03	0.02	0.03
Vertexing	0.02	0.01	0.05	0.18	0.20	0.17	0.08	0.07	0.11
Resolution	0.01	0.01	<0.01	0.10	0.14	0.28	0.07	0.11	0.26
Flavor tagging	0.01	0.01	<0.01	0.03	0.03	0.03	0.05	0.03	0.02
Fit bias	<0.01	0.02	<0.01	0.03	0.09	0.02	0.27	0.08	0.26
TSI	0.03	0.03	0.01	0.06	0.03	0.01	0.05	0.04	0.02
total	0.07	0.08	0.08	0.72	0.60	0.65	0.91	0.47	0.83

Table 6.8: Table of systematic errors (3). The notation "< 0.01" means that the value is small and less than 0.01, and thus invisible in the number of significant digits shown here. We calculate the total systematic error including these small contributions.

	I_+	I_-	I_0	I_{+-}^{Re}	I_{+0}^{Re}	I_{-0}^{Re}	I_{+-}^{Im}	I_{+0}^{Im}	I_{-0}^{Im}
ρ' and ρ''	0.02	0.02	0.03	0.82	0.64	0.55	0.46	0.56	0.48
SCF	0.01	0.01	0.01	0.18	0.27	0.10	0.38	0.17	0.14
Signal Dalitz	0.01	0.01	0.01	0.28	0.22	0.14	0.27	0.21	0.30
BG Dalitz	0.01	0.01	0.01	0.29	0.35	0.26	0.28	0.26	0.34
Other $\pi\pi\pi$	0.02	0.03	0.01	0.13	0.10	0.10	0.10	0.13	0.14
BG fraction	0.01	0.01	0.01	0.13	0.24	0.19	0.16	0.15	0.25
Physics	0.01	0.01	<0.01	0.04	0.05	0.03	0.04	0.03	0.05
BG Δt	<0.01	<0.01	<0.01	0.05	0.04	0.03	0.05	0.04	0.09
Vertexing	0.02	0.01	0.03	0.11	0.24	0.09	0.31	0.36	0.16
Resolution	0.01	0.01	0.01	0.19	0.22	0.15	0.28	0.20	0.23
Flavor tagging	<0.01	<0.01	<0.01	0.04	0.07	0.04	0.04	0.07	0.03
Fit bias	<0.01	0.01	<0.01	0.11	0.10	0.41	0.25	0.13	0.18
TSI	<0.01	<0.01	<0.01	0.09	0.04	0.06	0.05	0.18	0.05
Total	0.04	0.04	0.05	0.98	0.92	0.80	0.89	0.85	0.81

Chapter 7

Interpretation of the Fit Result

In this chapter, we describe several interpretations of our result of the time-dependent Dalitz plot analysis. Though the main purpose of this analysis is to constrain the CKM angle ϕ_2, which is described in Sec. 2, our analysis also gives useful information on the parameters related to the quasi-two-body processes of $B^0 \to \rho^+\pi^-$, $\rho^-\pi^+$, and $\rho^0\pi^0$, as written in Sec. 1.

1 Quasi-Two-Body Parameters

Among the 27 coefficients determined by the time-dependent Dalitz plot fit, those of the non-interfering terms, U_+^+, U_-^+, U_0^+, U_+^-, U_-^-, U_0^-, I_+, I_-, and I_0 can be interpreted as the CP-violation and non-CP-violation parameters of the quasi-two-body decay processes of $B^0 \to \rho^+\pi^-$, $B^0 \to \rho^-\pi^+$, and $B^0 \to \rho^0\pi^0$. In the following, we extract the quasi-two-body parameters from the result of the time-dependent Dalitz plot analysis.

1-1 Parameters Related to $B^0 \to \rho^\pm\pi^\mp$

Since $\rho^\pm\pi^\mp$ are not CP eigenstates, the time-dependent decay rates of the processes $B^0 \to \rho^\pm\pi^\mp$ are described by Eq. (2.96). We can calculate the CP-violation parameters in the equation from the several of the 27 coefficients determined by the time-dependent Dalitz plot fit as

$$\mathcal{C}^+ = \frac{U_+^-}{U_+^+}\,, \quad \mathcal{C}^- = \frac{U_-^-}{U_-^+}\,, \quad \mathcal{S}^+ = \frac{2I_+}{U_+^+}\,, \quad \mathcal{S}^- = \frac{2I_-}{U_-^+}\,, \quad \mathcal{A}_{\rho\pi}^{CP} = \frac{U_+^+ - U_-^+}{U_+^+ + U_-^+}\,, \tag{7.1}$$

and

$$\mathcal{C} \equiv \frac{\mathcal{C}^+ + \mathcal{C}^-}{2}\,, \quad \Delta\mathcal{C} \equiv \frac{\mathcal{C}^+ - \mathcal{C}^-}{2}\,, \quad \mathcal{S} \equiv \frac{\mathcal{S}^+ + \mathcal{S}^-}{2}\,, \quad \Delta\mathcal{S} \equiv \frac{\mathcal{S}^+ - \mathcal{S}^-}{2}\,. \tag{7.2}$$

As described in Sec. 3-4 of chapter 2, $\mathcal{A}_{\rho\pi}^{CP}$, $\mathcal{C}$, and $\mathcal{S}$ are CP-violation parameters, while $\Delta\mathcal{C}$ and $\Delta\mathcal{S}$ can be non-zero even without CP violation, The charge asymmetry, $\mathcal{A}_{\rho\pi}^{CP}$, is a time- and flavor-integrated CP-violation parameter that can be interpreted as

$$\mathcal{A}_{\rho\pi}^{CP} = \frac{\Gamma(B^0, \overline{B}^0 \to \rho^+\pi^-) - \Gamma(B^0, \overline{B}^0 \to \rho^-\pi^+)}{\Gamma(B^0, \overline{B}^0 \to \rho^+\pi^-) + \Gamma(B^0, \overline{B}^0 \to \rho^-\pi^+)}\,. \tag{7.3}$$

Table 7.1: Correlation matrix of the quasi-two-body parameters, with statistical and systematic errors combined.

	$\mathcal{A}_{\rho\pi}^{CP}$	$\mathcal{C}$	$\Delta\mathcal{C}$	$\mathcal{S}$	$\Delta\mathcal{S}$
$\mathcal{A}_{\rho\pi}^{CP}$	$+1.00$				
$\mathcal{C}$	-0.17	$+1.00$			
$\Delta\mathcal{C}$	$+0.10$	$+0.16$	$+1.00$		
$\mathcal{S}$	$+0.01$	-0.02	-0.00	$+1.00$	
$\Delta\mathcal{S}$	-0.00	-0.01	-0.02	$+0.29$	$+1.00$

From our measurement listed in Table 6.2, we obtain

$$\mathcal{A}_{\rho\pi}^{CP} = -0.12 \pm 0.05 \pm 0.04\,, \tag{7.4}$$
$$\mathcal{C} = -0.13 \pm 0.09 \pm 0.05\,, \tag{7.5}$$
$$\Delta\mathcal{C} = +0.36 \pm 0.10 \pm 0.05\,, \tag{7.6}$$
$$\mathcal{S} = +0.06 \pm 0.13 \pm 0.05\,, \quad \text{and} \tag{7.7}$$
$$\Delta\mathcal{S} = -0.08 \pm 0.13 \pm 0.05\,, \tag{7.8}$$

where first and second errors are statistical and systematic, respectively. The correlation matrix is shown in Table 7.1.

One can transform the parameters into the direct CP violation parameters $\mathcal{A}_{\rho\pi}^{+-}$ and $\mathcal{A}_{\rho\pi}^{-+}$ defined as

$$\mathcal{A}_{\rho\pi}^{+-} \equiv -\frac{\mathcal{A}_{\rho\pi}^{CP} + \mathcal{C} + \mathcal{A}_{\rho\pi}^{CP}\Delta\mathcal{C}}{1 + \Delta\mathcal{C} + \mathcal{A}_{\rho\pi}^{CP}\,\mathcal{C}}\,, \quad \text{and} \tag{7.9}$$
$$\mathcal{A}_{\rho\pi}^{-+} \equiv +\frac{\mathcal{A}_{\rho\pi}^{CP} - \mathcal{C} - \mathcal{A}_{\rho\pi}^{CP}\Delta\mathcal{C}}{1 - \Delta\mathcal{C} - \mathcal{A}_{\rho\pi}^{CP}\,\mathcal{C}}\,, \tag{7.10}$$

which are more convenient for interpretation since

$$\mathcal{A}_{\rho\pi}^{+-} = \frac{\Gamma(\overline{B}^0 \to \rho^-\pi^+) - \Gamma(B^0 \to \rho^+\pi^-)}{\Gamma(\overline{B}^0 \to \rho^-\pi^+) + \Gamma(B^0 \to \rho^+\pi^-)}\,, \quad \text{and} \tag{7.11}$$
$$\mathcal{A}_{\rho\pi}^{-+} = \frac{\Gamma(\overline{B}^0 \to \rho^+\pi^-) - \Gamma(B^0 \to \rho^-\pi^+)}{\Gamma(\overline{B}^0 \to \rho^+\pi^-) + \Gamma(B^0 \to \rho^-\pi^+)}\,. \tag{7.12}$$

We obtain

$$\mathcal{A}_{\rho\pi}^{+-} = +0.21 \pm 0.08 \pm 0.04\,, \quad \text{and} \tag{7.13}$$
$$\mathcal{A}_{\rho\pi}^{-+} = +0.08 \pm 0.16 \pm 0.11\,, \tag{7.14}$$

with a correlation coefficient of $+0.47$. Our result differs from the case with no direct CP asymmetry ($\mathcal{A}_{\rho\pi}^{+-} = 0$ and $\mathcal{A}_{\rho\pi}^{-+} = 0$) by 2.3 standard deviations (Fig. 7.1). More data would be useful to clarify whether direct CP violation is present.

1-2 Parameters Related to $B^0 \to \rho^0\pi^0$

Evidence for $B^0 \to \rho^0\pi^0$ Decay Channel

The existence of the decay channel $B^0 \to \rho^0\pi^0$ has been a matter of argument due to the discrepancy between the branching fractions reported by Belle and BaBar. Though the discrepancy is becoming smaller as the data increase, there is still a difference of $1.7\,\sigma$ between

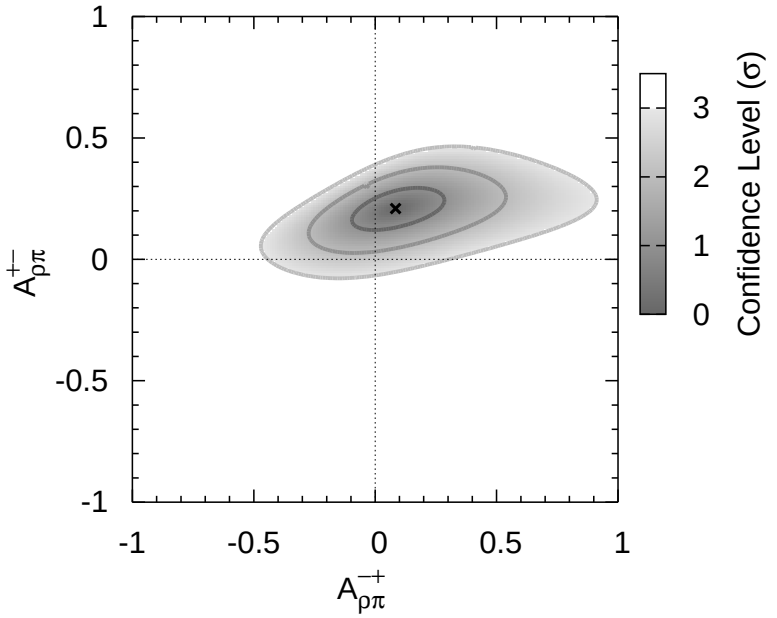


Figure 7.1: Contour plot of the confidence level for the direct CP violation parameters $\mathcal{A}^{+-}_{\rho\pi}$ vs. $\mathcal{A}^{-+}_{\rho\pi}$.

them [57, 58]:

$$\mathcal{B}(B^0 \to \rho^0\pi^0)^{\rm Belle} = 3.12^{+0.88+0.60}_{-0.82-0.76} \quad (\times 10^{-6}) \,, \tag{7.15}$$

$$\mathcal{B}(B^0 \to \rho^0\pi^0)^{\rm BaBar} = 1.4 \pm 0.6 \pm 0.3 \quad (\times 10^{-6}) \,. \tag{7.16}$$

In our analysis, non-zero U_0^+ corresponds to the existence of the $B^0 \to \rho^0\pi^0$ decay channel. As listed in Table 6.2, we obtain

$$U_0^+ = +0.29 \pm 0.05 \pm 0.04 \,, \tag{7.17}$$

which significantly differs from zero. Since the likelihood curve deviates from a Gaussian as shown in Fig. 7.2, we calculate the significance by the convolution of the likelihood curve and a Gaussian that describes the systematic error. We obtain a result that excludes $U_0^+ = 0$ by the significance of $4.8\,\sigma$. This is the most significant evidence for the $B^0 \to \rho^0\pi^0$ decay channel at present.

To compare our result with the dedicated branching fraction measurements, we use the ratio of the branching fractions, $\mathcal{B}(B^0 \to \rho^0\pi^0)/\mathcal{B}(B^0 \to \rho^\pm\pi^\mp)$. The result of the Dalitz plot analysis is related to the ratio as

$$\frac{\mathcal{B}(B^0 \to \rho^0\pi^0)}{\mathcal{B}(B^0 \to \rho^\pm\pi^\mp)} = \frac{U_0^+}{U_+^+ + U_-^+} \,. \tag{7.18}$$

Our measurement yields

$$\frac{\mathcal{B}(B^0 \to \rho^0\pi^0)}{\mathcal{B}(B^0 \to \rho^\pm\pi^\mp)} = 0.130 \pm 0.022 \pm 0.020 \,. \tag{7.19}$$

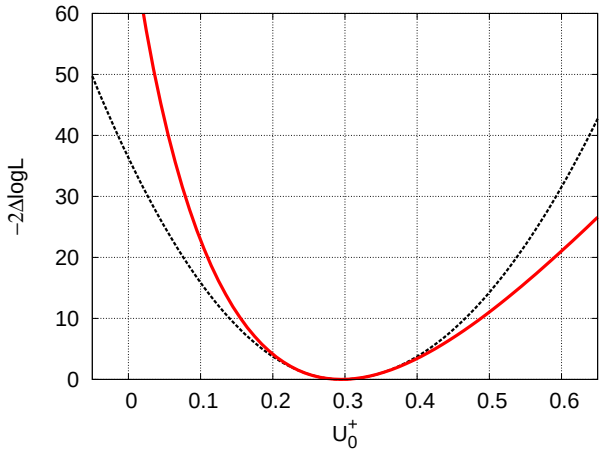


Figure 7.2: The likelihood curve (solid line) as a function of U_0^+. The dotted curve shows a Gaussian curve corresponding the statistical error we quote.

The branching fraction measurements of Belle and BaBar yields

$$\frac{\mathcal{B}(B^0 \to \rho^0\pi^0)^{\rm Belle}}{\mathcal{B}(B^0 \to \rho^\pm\pi^\mp)^{\rm WA}} = \frac{3.12^{+0.88+0.60}_{-0.82-0.76}}{24.0 \pm 2.5} = 0.130^{+0.049}_{-0.046}\ , \quad \text{and} \tag{7.20}$$

$$\frac{\mathcal{B}(B^0 \to \rho^0\pi^0)^{\rm BaBar}}{\mathcal{B}(B^0 \to \rho^\pm\pi^\mp)^{\rm WA}} = \frac{1.4 \pm 0.6 \pm 0.3}{24.0 \pm 2.5} = 0.058 \pm 0.029\ , \tag{7.21}$$

where we use the world average [54, 55] for the denominator. Our result supports the branching fraction measurement of Belle. The Dalitz plot analysis by BaBar also favors rather large contribution from $B^0 \to \rho^0\pi^0$ [59], being consistent with our result:

$$\left.\frac{\mathcal{B}(B^0 \to \rho^0\pi^0)}{\mathcal{B}(B^0 \to \rho^\pm\pi^\mp)}\right|_{\rm BaBar,Dalitz} = 0.103 \pm 0.018 \pm 0.019\ . \tag{7.22}$$

CP Violation Parameters

We also measure the CP violation parameters of $B^0 \to \rho^0\pi^0$. Since $\rho^0\pi^0$ is a CP eigenstate, its time-dependent decay rate is described by Eq. (2.92). The CP violation parameters $\mathcal{A}_{\rho^0\pi^0}$ and $\mathcal{S}_{\rho^0\pi^0}$ are calculated as

$$\mathcal{A}_{\rho^0\pi^0} = -\frac{U_0^-}{U_0^+}\ , \quad \text{and} \quad \mathcal{S}_{\rho^0\pi^0} = \frac{2I_0}{U_0^+}\ . \tag{7.23}$$

We obtain

$$\mathcal{A}_{\rho^0\pi^0} = -0.49 \pm 0.36 \pm 0.28\ , \quad \text{and} \tag{7.24}$$

$$\mathcal{S}_{\rho^0\pi^0} = +0.17 \pm 0.57 \pm 0.35\ , \tag{7.25}$$

with the correlation coefficient of -0.08. This is the first measurement of $\mathcal{S}_{\rho^0\pi^0}$. Our measurement of $\mathcal{A}_{\rho^0\pi^0}$ is consistent with and better than the previous measurement from Belle based on a data set corresponding to $386 \times 10^6\ B\bar{B}$ pairs [57]:

$$\mathcal{A}_{\rho^0\pi^0} = -0.53^{+0.67+0.10}_{-0.84-0.15}\ . \tag{7.26}$$

2 Constraint on ϕ_2

We constrain the CKM angle ϕ_2 from our analysis following the procedure described in Ref. [19]. With three $B^0 \to (\rho\pi)^0$ decay modes, we have 9 free parameters including ϕ_2:

$$\begin{aligned} 9 = &(6 \text{ complex amplitudes} = 12 \text{ d.o.f.}) + \phi_2 \\ &-(1 \text{ global phase}) - (1 \text{ global normalization}) - (1 \text{ isospin relation} = 2 \text{ d.o.f.}) \ , \end{aligned} \tag{7.27}$$

where we make use of an isospin relation that relates neutral B decay processes only [17, 18]. Parameterizing the 6 complex amplitudes with 9 free parameters, we form a χ^2 function using the 26 measurements of our time-dependent Dalitz plot analysis as constraints. We first optimize all the 9 parameters to obtain a minimum χ^2, $\chi^2_{\rm min}$; we then scan ϕ_2 from 0° to 180° optimizing the other 8 parameters, whose resultant minima are defined as $\chi^2(\phi_2)$; and $\Delta\chi^2(\phi_2)$ is defined as $\Delta\chi^2(\phi_2) \equiv \chi^2(\phi_2) - \chi^2_{\rm min}$. Performing a toy MC study following the procedure described in Ref. [60], we obtain the $1-$C.L. plot from the $\Delta\chi^2(\phi_2)$[1]; the result is shown in Fig. 7.3 as a dotted curve[2].

To incorporate all available knowledge, we combine our analysis with the information of the related charged B decay processes. We use the following world average branching fractions and asymmetries: $\mathcal{B}(B^0 \to \rho^\pm\pi^\mp)$, $\mathcal{B}(B^+ \to \rho^+\pi^0)$, $\mathcal{A}(B^+ \to \rho^+\pi^0)$, $\mathcal{B}(B^+ \to \rho^0\pi^+)$, and $\mathcal{A}(B^+ \to \rho^0\pi^+)$ [54, 55], which are not correlated with our 26 observables. With the 31 measurements above, we perform a full combined Dalitz and isospin(pentagon) analysis. Having 5 related decay modes, we have 12 free parameters including ϕ_2:

$$\begin{aligned} 12 = &(10 \text{ complex amplitudes} = 20 \text{ d.o.f.}) + \phi_2 \\ &-(1 \text{ global phase}) - (4 \text{ isospin relations} = 8 \text{ d.o.f.}) \ . \end{aligned} \tag{7.28}$$

The detail of χ^2 formation can be found in appendix B. The obtained $\chi^2_{\rm min}$ is 10.2, which is reasonable for 31(measurements) $-$ 12(free parameters) $=$ 19 degrees of freedom. Following the same procedure as above, we obtain the $1-$C.L. plot in the Fig. 7.3 as a solid curve. We obtain

$$68^\circ < \phi_2 < 95^\circ \tag{7.29}$$

as the 68.3% confidence interval consistent with the SM expectation. A large SM-disfavored region ($0^\circ < \phi_2 < 5^\circ$, $23^\circ < \phi_2 < 34^\circ$, and $109^\circ < \phi_2 < 180^\circ$) also remains.

[1] $\Delta\chi^2(\phi_2)$ is usually expected to follow a χ^2 distribution with one degree of freedom and thus the cumulative χ^2 distribution for one degree of freedom is usually used to convert $\Delta\chi^2(\phi_2)$ into a $1-$C.L. plot. A toy MC study shows, however, that this is not the case for $B \to \rho\pi$, and an analysis with this assumption yields confidence intervals with undercoverage. Thus, we perform a dedicated toy MC study to obtain the confidence interval with correct coverage.

[2] A $1-$C.L. plot reads as follows. Figure 7.4 shows an example of the plot of $1-$C.L. vs. ϕ_2. To calculate a confidence interval for a confidence level of C.L. $= x$, one reads the intersection of a horizontal line of $1-$C.L. $= 1-x$ and the filled region, or the region where the plotted curve is above the line of $1-$C.L. $= 1-x$. For example, a confidence interval for 68.3% C.L. is obtained by the intersection of a line $1-$C.L. $= 0.317$ and the filled region, leading to $70^\circ < \phi_2 < 110^\circ$ in this example. Note that a confidence level can only be defined for an interval (confidence interval) of ϕ_2 and not for a specific value of ϕ_2. Thus, a statement like the following is misleading: *the confidence level for $\phi_2 = 90^\circ$ is 0.* A correct statement is *the confidence level for the confidence interval of $70^\circ < \phi_2 < 110^\circ$ is 68.3%.*

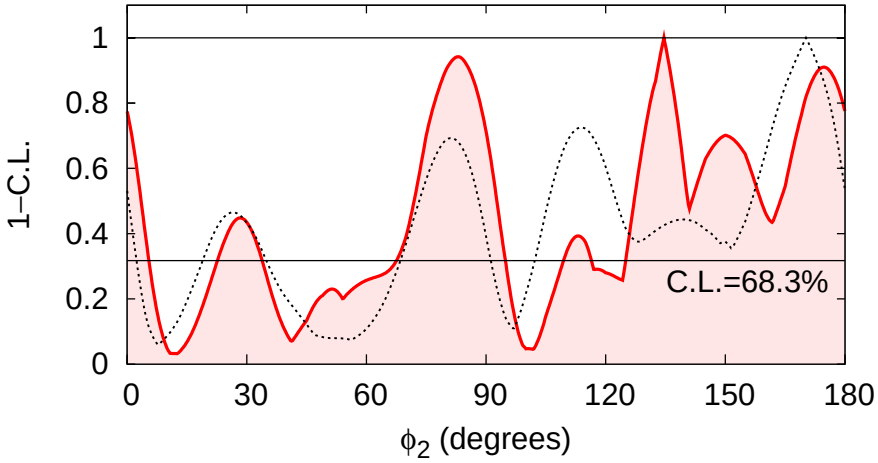


Figure 7.3: $1 - \text{C.L.}$ vs. ϕ_2. Dotted and solid curves correspond to the result from the time-dependent Dalitz plot analysis only and that from Dalitz and isospin (pentagon) combined analysis, respectively.

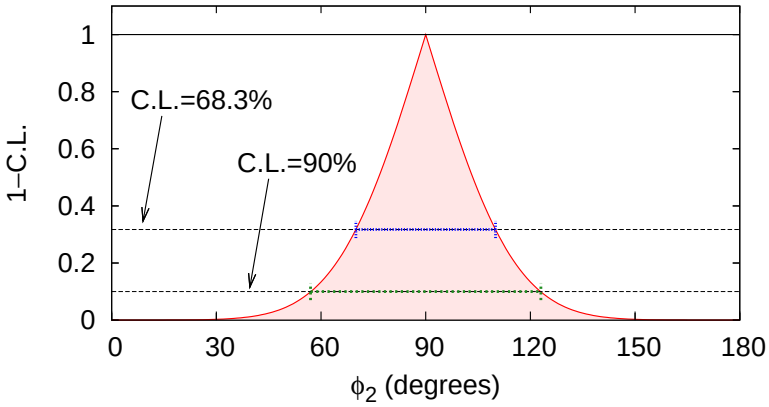


Figure 7.4: An example of $1 - \text{C.L.}$ vs. ϕ_2 plot with a central value of $\phi_2 = 90°$ and a $1\,\sigma$ error of 20°. Dotted (blue) and dot-dashed (green) horizontal error bars correspond to the confidence intervals for the confidence levels of 68.3% and 90%, respectively.

Chapter 8

Discussions and Conclusions

1 Test of Standard Model by ϕ_2 Measurement

As described in chapter 1, the comparison of the direct measurement and indirect measurement, or the standard model (SM) expectation, of a parameter in the CKM model is a good test of SM. In this section, we perform the SM test using ϕ_2 as the test parameter, where the direct measurement incorporates our analysis of $B \to \rho\pi$.

1-1 Direct Measurements of ϕ_2

The direct measurement of the CKM angle ϕ_2 is possible by using the isospin analysis [16] of the decay modes $B \to \pi\pi$ and $B \to \rho\rho$, as well as the analysis of $B \to \rho\pi$. Figures 8.1 and 8.2 show the constraint on ϕ_2 obtained by the isospin analysis using the world averages [54, 55] of the CP-violating asymmetries and branching fractions of $B \to \pi\pi$ and $B \to \rho\rho$, respectively. As 68.3% confidence intervals, they give

$$0^\circ < \phi_2^{\pi\pi} < 7^\circ,\ 83^\circ < \phi_2^{\pi\pi} < 105^\circ,\ 115^\circ < \phi_2^{\pi\pi} < 155^\circ,\ 165^\circ < \phi_2^{\pi\pi} < 180^\circ, \tag{8.1}$$

and

$$0^\circ < \phi_2^{\rho\rho} < 18^\circ, \quad 72^\circ < \phi_2^{\rho\rho} < 112^\circ, \quad 158^\circ < \phi_2^{\rho\rho} < 180^\circ. \tag{8.2}$$

By combining the constraints from $B \to \pi\pi$ and $B \to \rho\rho$, the number of discrete solutions in the constraint from $B \to \pi\pi$ decreases from four to three:

$$0^\circ < \phi_2^{\pi\pi,\rho\rho} < 7^\circ, \quad 83^\circ < \phi_2^{\pi\pi,\rho\rho} < 105^\circ, \quad 165^\circ < \phi_2^{\pi\pi,\rho\rho} < 180^\circ. \tag{8.3}$$

We combine this result with our result of $B \to \rho\pi$ shown in Fig. 8.3, which gives a 68.3% confidence interval of

$$0^\circ < \phi_2^{\rho\pi} < 5^\circ\ , \quad 23^\circ < \phi_2^{\rho\pi} < 34^\circ\ , \quad 68^\circ < \phi_2^{\rho\pi} < 95^\circ\ , \quad 109^\circ < \phi_2^{\rho\pi} < 180^\circ\ . \tag{8.4}$$

Figure 8.4 shows the combined result. By the addition of our analysis, the constraint improves, in particular for the solution of $\phi_2 \sim 90^\circ$:

$$0^\circ < \phi_2 < 4^\circ, \quad 83^\circ < \phi_2 < 95^\circ, \quad 167^\circ < \phi_2 < 180^\circ. \tag{8.5}$$

From the Fig. 8.4, one may consider the solution of $\phi_2 \sim 180^\circ$ is better favored than that of $\phi_2 \sim 90^\circ$. However, the difference is insignificant, corresponding to the significance of $0.16\,\sigma$. Thus, at current statistics we cannot conclude that one of the solutions is more favored than the other.

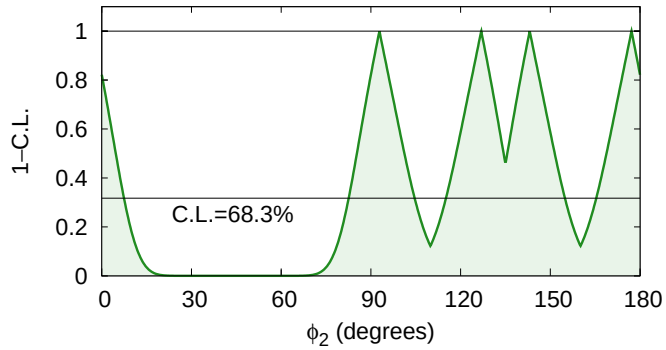


Figure 8.1: $1 - \text{C.L.}$ vs. ϕ_2 obtained by using $B \to \pi\pi$ decay modes.

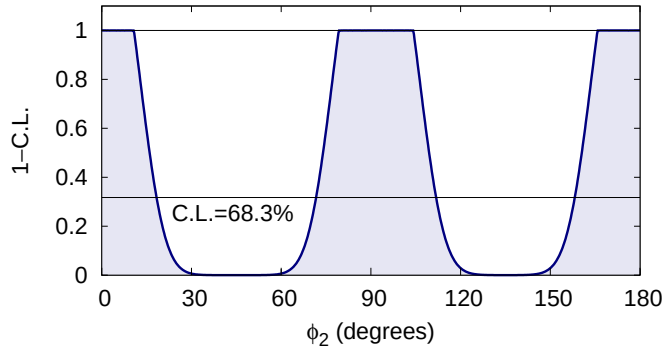


Figure 8.2: $1 - \text{C.L.}$ vs. ϕ_2 obtained by using $B \to \rho\rho$ decay modes.

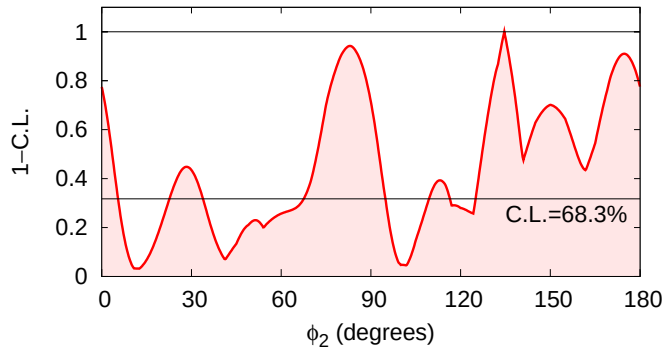


Figure 8.3: $1 - \text{C.L.}$ vs. ϕ_2 obtained from our $B \to \rho\pi$ analysis.

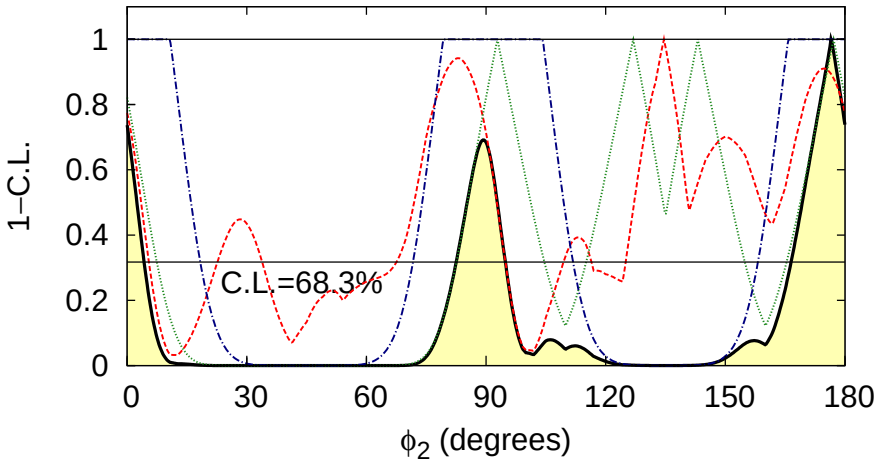


Figure 8.4: $1-$C.L. vs. ϕ_2 obtained by combining the constraints from $B \to \pi\pi$, $B \to \rho\pi$, and $B \to \rho\rho$ (solid curve). Dashed, dotted, and dot-dashed curves correspond to the constraints from $B \to \rho\pi$, $B \to \pi\pi$, and $B \to \rho\rho$, respectively.

1-2 Comparison of the Direct and Indirect Measurements of ϕ_2

As described in chapter 1, the CKM-picture allows us to calculate ϕ_2 as SM expectation from the measurements other than the direct measurements. The indirect measurement of ϕ_2 is mainly determined by the constraints from $\sin 2\phi_1$ and $|V_{td}V_{tb}^*|/|V_{cd}V_{cb}^*|$, as illustrated in Fig. 1.1. The former is from the measurement of the CP violation in the "golden mode," $B^0 \to J/\psi K^0$, at B-factories. The latter is decomposed as

$$\frac{|V_{td}V_{tb}^*|}{|V_{cd}V_{cb}^*|} = \frac{|V_{td}V_{tb}^*|}{|V_{ts}V_{tb}^*|}\frac{1}{|V_{us}|}\cdot\frac{|V_{ts}|}{|V_{cb}|}\frac{1}{|V_{tb}^*|}\frac{|V_{us}|}{|V_{cd}|}\,. \tag{8.6}$$

Here the following relations hold by the unitary and hierarchal properties of the CKM matrix (See Eq. (2.11)):

$$\frac{|V_{ts}|}{|V_{cb}|} \sim \frac{1}{|V_{tb}^*|} \sim \frac{|V_{us}|}{|V_{cd}|} \sim 1\,. \tag{8.7}$$

Consequently,

$$\frac{|V_{td}V_{tb}^*|}{|V_{cd}V_{cb}^*|} \sim \frac{|V_{td}V_{tb}^*|}{|V_{ts}V_{tb}^*|}\frac{1}{|V_{us}|}\,, \tag{8.8}$$

where each factor is obtained by the measurement of B_d-$\overline{B}_d$ mixing parameter

$$\Delta m_d \sim |V_{td}V_{tb}^*|^2\,, \tag{8.9}$$

the measurement of B_s-$\overline{B}_s$ mixing parameter

$$\Delta m_s \sim |V_{ts}V_{tb}^*|^2\,, \tag{8.10}$$

and the measurements of the kaon decays ($|V_{us}|$). The detailed description can be found elsewhere [60]. Note that the important aspect in the measurement of $|V_{td}V_{tb}^*|/|V_{cd}V_{cb}^*|$ is to take the ratio of Δm_d and Δm_s, which leads to a sizable improvement of the precision by the cancellation of the theoretical uncertainty of QCD calculation. The dominant uncertainty of $|V_{td}V_{tb}^*|/|V_{cd}V_{cb}^*|$ comes from the theoretical calculation of the $SU(3)$ breaking effect between B_d and B_s.

The global fit using all the available measurements as of the summer 2006, except for the direct measurements of ϕ_2, yields [11]

$$\phi_2^{\rm ind.} = \left(100.0^{+4.5}_{-7.3}\right)^\circ .$$

On the other hand, the CKM-favored solution from direct measurements is

$$83^\circ < \phi_2^{\rm dir.} < 95^\circ .$$

They have an overlapped region and thus they are consistent with each other; the CKM-picture of CP violation is confirmed to be correct by the comparison of the direct and indirect measurements of ϕ_2 at the precision of $\sigma_{\phi_2} \sim 7^\circ$.

2 Conclusions

In summery, we performed a time-dependent Dalitz plot analysis of $B^0 \to \pi^+\pi^-\pi^0$ decays based on a 414fb^{-1} data sample that contains $449 \times 10^6 B\bar{B}$ pairs collected on the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB energy-asymmetric e^+e^- collider.

Combining the time-dependent Dalitz plot observables obtained from our analysis with the information on charged B decay modes, we perform a full Dalitz and isospin analysis for the first time and obtain a constraint on the CKM angle ϕ_2,

$$68^\circ < \phi_2 < 95^\circ ,$$

as the 68.3% confidence interval consistent with the standard model (SM). A large SM-disfavored region ($0^\circ < \phi_2 < 5^\circ$, $23^\circ < \phi_2 < 34^\circ$ and, $109^\circ < \phi_2 < 180^\circ$) also remains. This result is combined with the measurements from $B \to \pi\pi$ and $B \to \rho\rho$ and gives the 68.3% confidence interval of

$$0^\circ < \phi_2 < 4^\circ, \quad 83^\circ < \phi_2 < 95^\circ, \quad 167^\circ < \phi_2 < 180^\circ ,$$

where our analysis of $B \to \rho\pi$ plays an essential role for the improvement of the precision. This result is compared with the SM expectation of $\phi_2 = \left(100.0^{+4.5}_{-7.3}\right)^\circ$, and found to be consistent with it; the CKM-picture of the CP violation in the quark sector is confirmed to be correct at the precision of $\sigma_{\phi_2} \sim 7^\circ$.

The time-dependent Dalitz plot observables are also interpreted as quasi-two-body CP-violation parameters in $B^0 \to \rho^\pm\pi^\mp$. We obtain

$$\begin{aligned}
\mathcal{A}_{\rho\pi}^{CP} &= -0.12 \pm 0.05 \pm 0.04 , \\
\mathcal{C} &= -0.13 \pm 0.09 \pm 0.05 , \\
\Delta\mathcal{C} &= +0.36 \pm 0.10 \pm 0.05 , \\
\mathcal{S} &= +0.06 \pm 0.13 \pm 0.05 , \quad \text{and} \\
\Delta\mathcal{S} &= -0.08 \pm 0.13 \pm 0.05 ,
\end{aligned}$$

where the first and second errors correspond to statistical and systematic errors, respectively. From $\mathcal{A}_{\rho\pi}^{CP}$, $\mathcal{C}$, and $\Delta\mathcal{C}$ above, direct CP-violation parameters in the $B^0 \to \rho^\pm\pi^\mp$ process are calculated to be

$$\begin{aligned}
\mathcal{A}_{\rho\pi}^{+-} &= +0.21 \pm 0.08 \pm 0.04 , \quad \text{and} \\
\mathcal{A}_{\rho\pi}^{-+} &= +0.08 \pm 0.16 \pm 0.11 .
\end{aligned}$$

The CP-violation parameters of the $B^0 \to \rho^0\pi^0$ process are also obtained to be

$$\begin{aligned}
\mathcal{A}_{\rho^0\pi^0} &= -0.49 \pm 0.36 \pm 0.28 , \quad \text{and} \\
\mathcal{S}_{\rho^0\pi^0} &= +0.17 \pm 0.57 \pm 0.35 ,
\end{aligned}$$

where $\mathcal{S}_{\rho^0\pi^0}$ is measured for the first time.

Appendix A

Detailed Study of the PDF for Each Component

1 Correctly Reconstructed Signal

The PDF for correctly reconstructed signal is described by Eq. (6.12). In this section, we discuss the ΔE-$M_{\rm bc}$ PDF, $\mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0})$, efficiency over the Dalitz plot, $\epsilon^l(m', \theta')$, and correction factor for it, $\epsilon'(p_{\pi^0})$, which is a product of $\epsilon'_{\pi^0}(p_{\pi^0})$ and $\epsilon'_{\rm FF}(p_{\pi^0})$ described in the followings. We also comment on possible detector resolution effect in Dalitz plot.

1-1 ΔE-$M_{\rm bc}$ PDF

We model the ΔE-$M_{\rm bc}$ PDF by binned histogram. Since the resolution of π^0 energy measurement is dependent on the momentum of π^0 in the laboratory frame, the width of the ΔE-$M_{\rm bc}$ distribution, in particular ΔE, is dependent on the π^0 momentum, which we take into account as described in the following. We also take account of the MC-data difference.

On the Correlation between ΔE and $M_{\rm bc}$

Since the final state $\pi^+\pi^-\pi^0$ includes π^0, the tail part of the ΔE-$M_{\rm bc}$ tends to be correlated and the correlation is dependent on the π^0 momentum, p_{π^0}, as shown in Fig. A.1. We adopt to use binned histogram for the ΔE-$M_{\rm bc}$ to take account of this correlation.

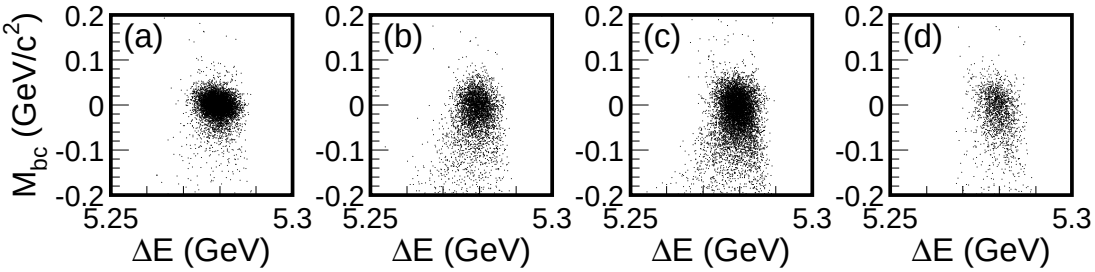


Figure A.1: MC ΔE-$M_{\rm bc}$ distribution of correctly reconstructed signal, for (a) $0.0\,{\rm GeV}/c < p_{\pi^0} < 1.0\,{\rm GeV}/c$, (b) $1.0\,{\rm GeV}/c < p_{\pi^0} < 2.0\,{\rm GeV}/c$, (c) $2.0\,{\rm GeV}/c < p_{\pi^0} < 3.0\,{\rm GeV}/c$, and (d) $3.0\,{\rm GeV}/c < p_{\pi^0} < 4.0\,{\rm GeV}/c$.

Dependence on π^0 Momentum

As shown in figures A.2 and A.3, ΔE-$M_{\rm bc}$ distribution of correctly reconstructed signal is dependent on the momentum of π^0, p_{π^0}. This have to be taken into account since p_{π^0} is strongly correlated with the square Dalitz plot variable m', as shown in Fig. A.4; otherwise, this effect could cause bias in the fit.

First, we subdivide the MC sample into π^0 momentum regions with $0.2\,\mathrm{GeV}/c$ widths. For each region, we prepare ΔE-$M_{\rm bc}$ PDF's, $P_{i_p}(\Delta E, M_{\rm bc})$, where i_p the index over the p_{π^0} with the correspondence of

$$i_p = \left[\frac{p_{\pi^0}\,(\mathrm{GeV}/c) - 0.1}{0.2}\right] , \tag{A.1}$$

where $[x]$ is the floor function, or the greatest integer function. Note that $p_{\pi^0} > 0.1\,\mathrm{GeV}/c$ is required in the event selection. We normalize $P_{i_p}(\Delta E, M_{\rm bc})$ in the grand signal region:

$$\iint_{\rm GS} d\Delta E\, dM_{\rm bc}\, P_{i_p}(\Delta E, M_{\rm bc}) = 1 . \tag{A.2}$$

With the $P_{i_p}(\Delta E, M_{\rm bc})$ above, we define the $\mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0})$ as an interpolation in the p_{π^0} direction:

$$\mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0}) = \frac{f_p\, P_{j_p}(\Delta E, M_{\rm bc}) + (1-f_p)\, P_{j_p+1}(\Delta E, M_{\rm bc})}{f_p\, n_{j_p} + (1-f_p)\, n_{j_p+1}} , \tag{A.3}$$

where

$$j_p = \left[\frac{p_{\pi^0}\,(\mathrm{GeV}/c) - 0.2}{0.2}\right] , \tag{A.4}$$

$$f_p = \frac{0.2\cdot(j_p+2) - p_{\pi^0}\,(\mathrm{GeV}/c)}{0.2} , \tag{A.5}$$

$$n_{j_p} = \iint_{\rm SR} d\Delta E\, dM_{\rm bc}\, P_{j_p}(\Delta E, M_{\rm bc}) . \tag{A.6}$$

Here, the last equation defines the integral over the signal region. The above expression is only valid for $0.2\,\mathrm{GeV}/c < p_{\pi^0} < 3.6\,\mathrm{GeV}/c$; outside the region, we define $\mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0})$ as

$$\mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0}) = \begin{cases} \dfrac{P_{j_p=0}(\Delta E, M_{\rm bc})}{n_{j_p=0}} & (p_{\pi^0} < 0.2\,\mathrm{GeV}/c) , \\ \dfrac{P_{j_p=17}(\Delta E, M_{\rm bc})}{n_{j_p=17}} & (p_{\pi^0} > 3.6\,\mathrm{GeV}/c) . \end{cases} \tag{A.7}$$

The denominator of Eq. (A.3) guarantees the required normalization condition:

$$\iint_{\rm SR} d\Delta E\, dM_{\rm bc} \mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0}) = 1 . \tag{A.8}$$

Correction for the Data-MC Difference (Fudge Factor)

There are differences between data and MC, mainly due to 1) the difference of tracking error and 2) the difference of the energy spread of e^+e^- beams. To take account of the difference, we apply correction factors, which we call Fudge Factor.

We use $B^0 \to D^-\rho^+(D^- \to K^+\pi^-\pi^-)$ and $B^0 \to D^{*-}\rho^+$ $(D^{*-} \to \overline{D}{}^0\pi^-, \overline{D}{}^0 \to K^+\pi^-)$ decay modes, which have large branching fractions and high purity, as control samples to obtain

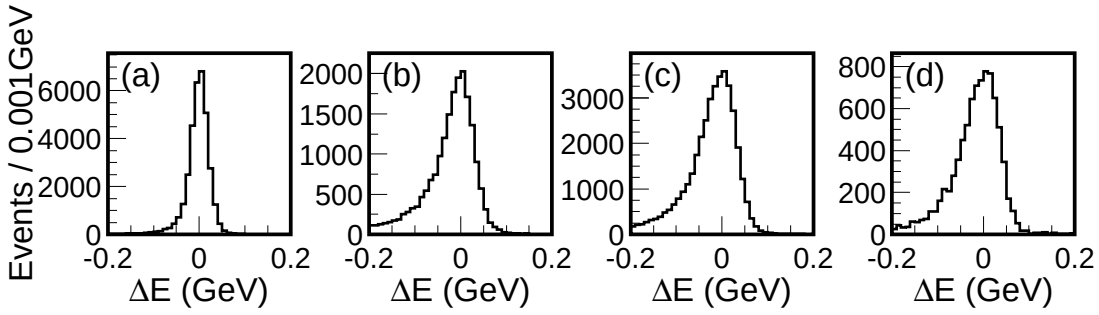


Figure A.2: MC ΔE distribution of correctly reconstructed signal, for (a) $0.0\,\mathrm{GeV}/c < p_{\pi^0} < 1.0\,\mathrm{GeV}/c$, (b) $1.0\,\mathrm{GeV}/c < p_{\pi^0} < 2.0\,\mathrm{GeV}/c$, (c) $2.0\,\mathrm{GeV}/c < p_{\pi^0} < 3.0\,\mathrm{GeV}/c$, and (d) $3.0\,\mathrm{GeV}/c < p_{\pi^0} < 4.0\,\mathrm{GeV}/c$.

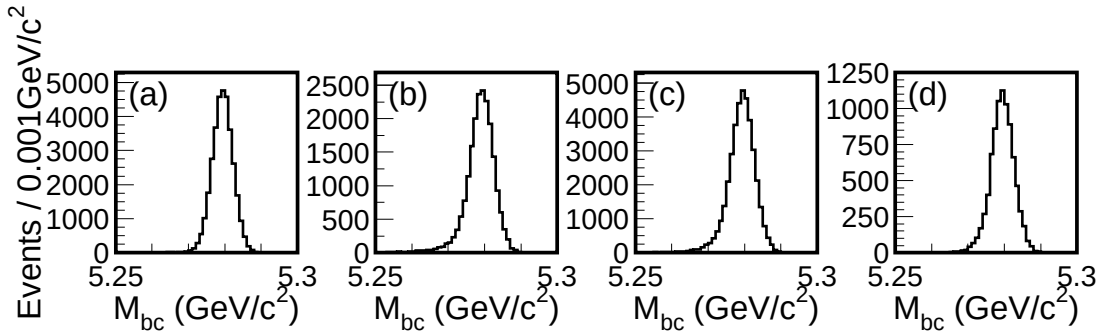


Figure A.3: MC M_{bc} distribution of correctly reconstructed signal, for (a) $0.0\,\mathrm{GeV}/c < p_{\pi^0} < 1.0\,\mathrm{GeV}/c$, (b) $1.0\,\mathrm{GeV}/c < p_{\pi^0} < 2.0\,\mathrm{GeV}/c$, (c) $2.0\,\mathrm{GeV}/c < p_{\pi^0} < 3.0\,\mathrm{GeV}/c$, and (d) $3.0\,\mathrm{GeV}/c < p_{\pi^0} < 4.0\,\mathrm{GeV}/c$.

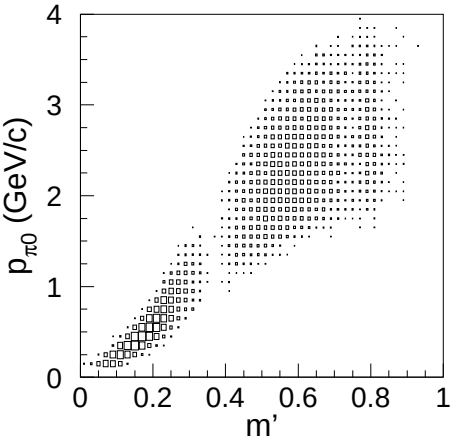


Figure A.4: Correlation between the π_0 momentum, p_{π^0}, and the square Dalitz plot variable m'.

the Fudge Factor. Comparing the distributions of data and MC, we obtain the difference of widths and mean values defined as

$$\Delta\mu \equiv \mu_{\rm Data} - \mu_{\rm MC} \ , \tag{A.9}$$

$$\Delta\sigma \equiv \sigma_{\rm Data}/\sigma_{\rm MC} \ , \tag{A.10}$$

for both of ΔE and $M_{\rm bc}$. Table A.1 lists the values with their statistical uncertainties.

With the obtained Fudge Factor, we define the corrected ΔE-$M_{\rm bc}$ PDF. First, we define the corrected PDF for each p_{π^0} region as

$$P'_{i_p}(\Delta E', {M_{\rm bc}}') = \frac{1}{\mathcal{N}'_{i_p}} P_{i_p}(\Delta E, M_{\rm bc}) \ , \tag{A.11}$$

where

$${M_{\rm bc}}' \equiv \frac{M_{\rm bc} - m_{B^0} - \Delta\mu^{M_{\rm bc}}}{\Delta\sigma^{M_{\rm bc}}} + m_{B^0} \ , \tag{A.12}$$

$$\Delta E' \equiv \frac{\Delta E - \Delta\mu^{\Delta E}}{\Delta\sigma^{\Delta E}} \ . \tag{A.13}$$

We determine the normalization factor $\mathcal{N}'_{i_p}$ such that the corrected PDF satisfy

$$\iint_{\rm GS} d\Delta E \, dM_{\rm bc} \, P'_{i_p}(\Delta E', {M_{\rm bc}}') = 1 \ . \tag{A.14}$$

Then, the corrected ΔE-$M_{\rm bc}$ PDF is defined in the same way as Eqs. (A.3) and (A.3), but $P_{j_p}(\Delta E, M_{\rm bc})$ and n_{j_p} are replaced with $P'_{i_p}(\Delta E', {M_{\rm bc}}')$ and n'_{j_p}, where

$$n'_{j_p} = \iint_{\rm SR} d\Delta E \, dM_{\rm bc} \, P'_{j_p}(\Delta E', {M_{\rm bc}}') \ . \tag{A.15}$$

Table A.1: Fudge Factor obtained by the control sample study.

	DS-I	DS-II
$\Delta\mu^{\Delta E}$ (MeV)	0.3 ± 1.1	-0.9 ± 1.4
$\Delta\sigma^{\Delta E}$	1.18 ± 0.05	1.18 ± 0.07
$\Delta\mu^{M_{\rm bc}}$ (MeV/c^2)	0.38 ± 0.11	0.36 ± 0.13
$\Delta\sigma^{M_{\rm bc}}$	0.93 ± 0.03	0.97 ± 0.04

1-2 Dalitz Plot PDF

The smearing and distortion of the Dalitz plot distribution are understood by the detector resolution and the efficiency dependent on the Dalitz plot position. As discussed in the followings, the effect of the former is negligibly small. As for the latter, we obtain the efficiency from MC-generated events and apply corrections to take account of data-MC difference.

On the Detector Resolution of Dalitz Plot

The measurement precision of the Dalitz plot variables is so good that we can ignore the finite resolution of it. Figure A.5 shows the resolution. It is good enough compared to the width of ρ, whose typical spread is ~ 0.05 in the square Dalitz plot. Using MC, we confirm the fit result is not biased even if we ignore the smearing in Dalitz plot due to the detector resolution.

The main reason of the good resolution is that we do not use the energy of π^0 to calculate the Dalitz plot variables. As implied in Eqs. (4.20) and (4.21), we only use 1) the four-momenta of π^+ and π^-, and 2) the flight direction of π^0 in the rest mass frame of $\pi^+\pi^-$. The usual Dalitz plot variables, s_+, s_-, and s_0, are calculated from the square Dalitz plot variables following appendix C, which corresponds to exploiting the relation of

$$s_+ + s_- + s_0 = {m_{B^0}}^2 + 2{m_{\pi^+}}^2 + {m_{\pi^0}}^2 \tag{A.16}$$

instead of using the π^0 energy.

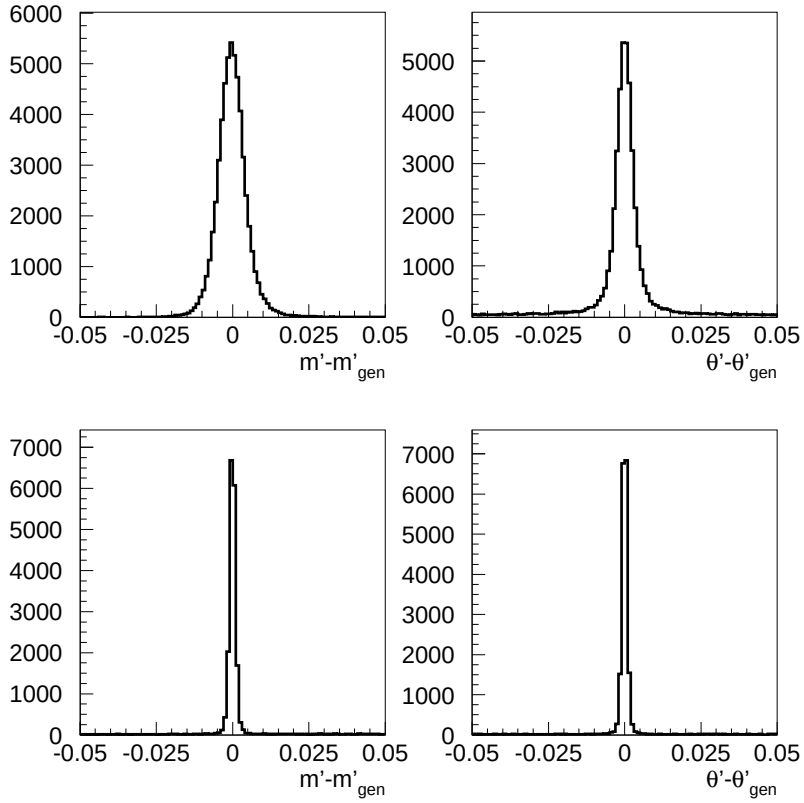


Figure A.5: Residual distribution of the square Dalitz plot variables of m' (left) and θ' (right) obtained from MC. The resolution is dependent on m'; the upper and lower plots correspond to the region of $0.1 < m' < 0.2$ and $0.7 < m' < 0.8$, respectively.

Efficiency

Since the acceptance is not constant over the Dalitz plot, we need to take account of it as a Dalitz plot dependent efficiency, $\epsilon(m', \theta')$, which is shown in Fig. A.6. As shown in Fig. A.7, the

efficiency curve is also dependent on the quality of the flavor tagging, r, and thus we introduce the dependence on the r-bin as $\epsilon(m', \theta'; l)$.

In practice, the efficiency is entangled with the event fraction of each r-region, $\mathcal{F}^l$, and we treat them together. As illustrated in Fig. A.8, it can be understood by two steps: first, the events are distributed into the six r-bins by flavor tagging with the fractions of $\mathcal{F}^l$, and secondly, the detection efficiency $\epsilon(m', \theta'; l)$ is multiplied. Our concern is only with the relation

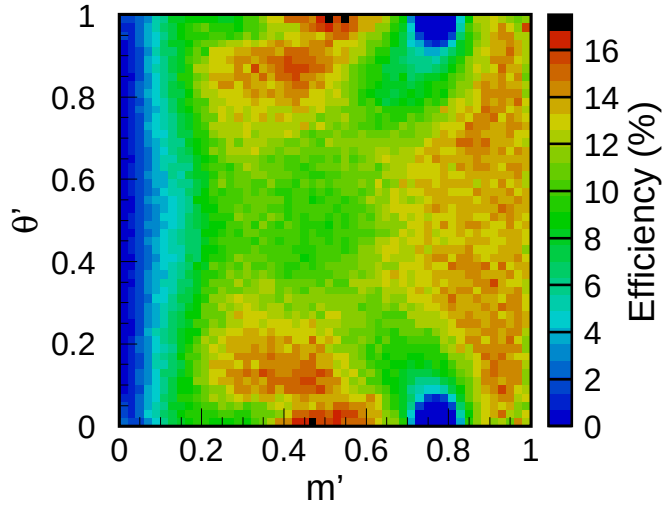


Figure A.6: The Dalitz-dependent efficiency summed over all the tagging quality region. The blue colored low efficiency regions correspond to the kinematic regions where one of the three pions have low momentum.

between the original number of events, N, and the detected number of events, $\mathcal{F}^l\epsilon(m', \theta'; l)N$; we do not need the intermediate information of $\mathcal{F}^l N$. That is, we only have to know about the product $\mathcal{F}^l\epsilon(m', \theta'; l)$ and do not have to disentangle it.

We obtain the $\mathcal{F}^l\epsilon(m', \theta'; l)$ from MC-generated events. With the number of generated events at the Dalitz plot position of (m', θ'), $N_{\rm gen}(m', \theta')$, and that of reconstructed events at (m', θ') in l-th r-bin, $N_{\rm rec}(m', \theta', l)$, we calculate it as

$$\mathcal{F}^l\epsilon(m', \theta'; l) = \frac{N_{\rm rec}(m', \theta', l)}{N_{\rm gen}(m', \theta')} \, . \tag{A.17}$$

As a technical aspect, we use a MC generator that distributes the events flatly over the square Dalitz plot, i.e., $N_{\rm gen}(m', \theta') \sim N(\text{constant})$, to minimize the statistical fluctuation of the MC with a limited computational resource.

On the Charge Asymmetry of Efficiency

The Dalitz-dependent efficiency obtained from MC has charge asymmetry; it has a significant deviation from the following relation

$$\epsilon(m', \theta'; l) = \epsilon(m', 1 - \theta'; l) \, . \tag{A.18}$$

Note that the replacement $\theta' \to 1 - \theta'$ corresponds to $\pi^\pm \to \pi^\mp$. The asymmetry defined as

$$\frac{\epsilon(m', \theta'; l) - \epsilon(m', 1 - \theta'; l)}{\epsilon(m', \theta'; l) + \epsilon(m', 1 - \theta'; l)} \tag{A.19}$$

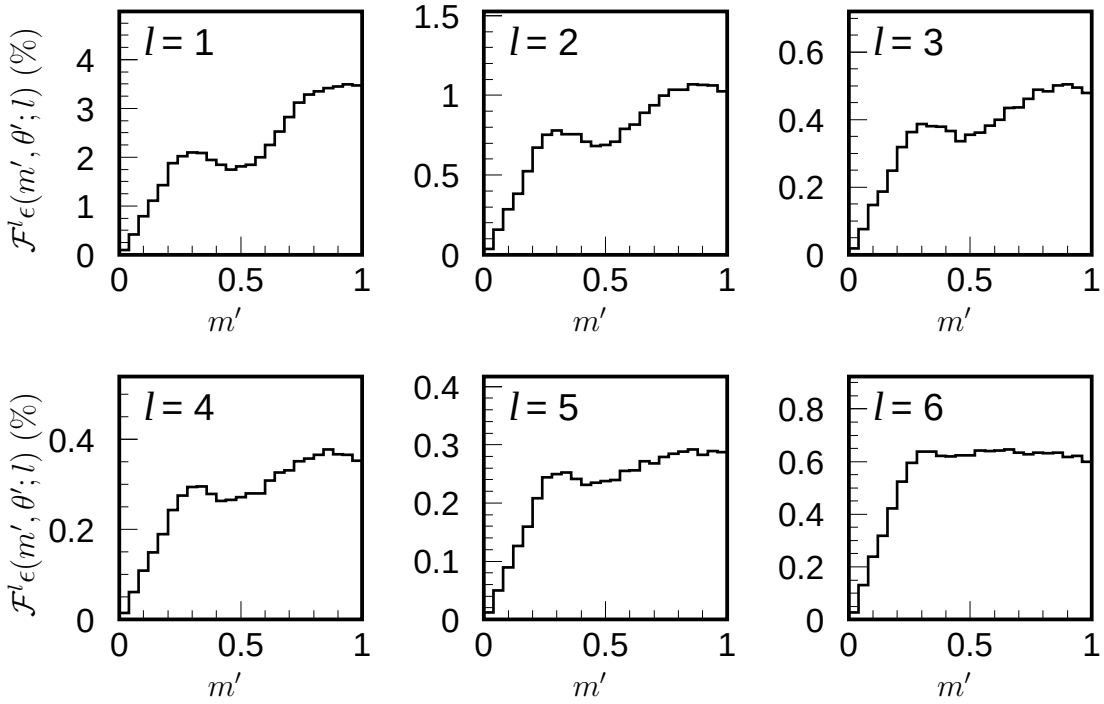


Figure A.7: The slice plot ($0.3 < \theta' < 0.7$) of the Dalitz-dependent efficiency for each tagging quality region. The vertical axis is the product of $\mathcal{F}^l$ and $\epsilon(m', \theta'; l)$ in %.

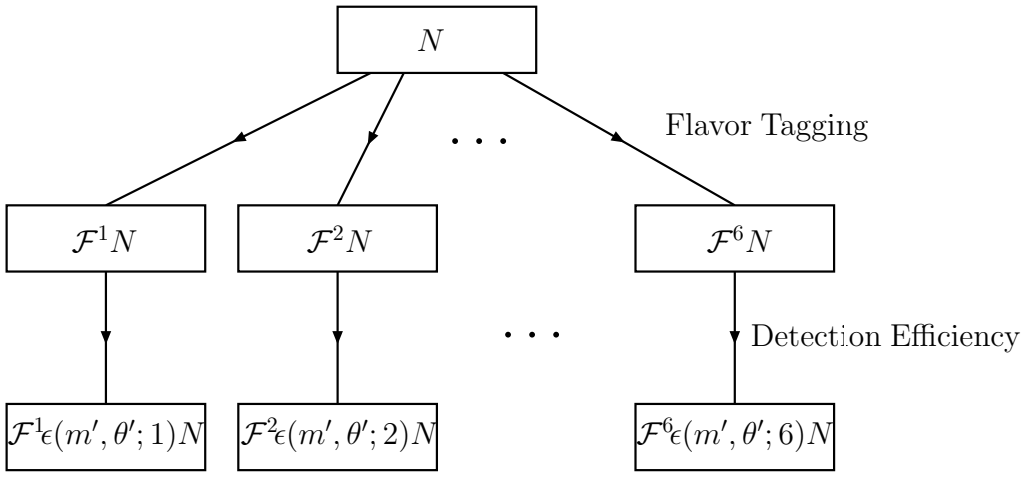


Figure A.8: Schematic illustration of the effect of r-bin event fraction, $\mathcal{F}^l$, and Dalitz-dependent efficiency, $\epsilon(m', \theta'; l)$.

tends to have negative value for $\theta' > 0.5$ and $\sim 3\%$ at most. We use the asymmetric efficiency for the nominal fit and quote twice of the difference between the cases with and without the charge asymmetry as the systematic error.

Efficiency Correction for Data-MC Difference

As described above, we rely on MC to determine the Dalitz-dependent efficiency, and the difference between data and MC should be taken into account. In particular, the electromagnetic cascade shower process is hard to reproduce perfectly, while the detection of charged tracks is well reproduced by MC. Thus, we introduce an efficiency correction factor dependent on the momentum of π^0, $\epsilon'_{\pi^0}(p_{\pi^0})$. We introduce the dependence on π^0 momentum, which is correlated with the square Dalitz plot variable m' as shown in Fig. A.4 [], since an overall efficiency difference is not important in our Dalitz plot analysis.

The correction factor is defined as

$$\epsilon'_{\pi^0}(p_{\pi^0}) = \frac{\epsilon^{\rm Data}(p_{\pi^0})}{\epsilon^{\rm MC}(p_{\pi^0})} \ , \tag{A.20}$$

and obtained using the control samples of the following decay modes:

- $B^+ \to \pi^+ \overline{D}^0$ ($\overline{D}^0 \to K_S \pi^0$),
- $B^0 \to \pi^+ \overline{D}^{*-}$ ($\overline{D}^0 \to K_S \pi^0$),
- $B^+ \to \rho^+ \overline{D}^0$ ($\overline{D}^0 \to K^+\pi^-$, $\overline{D}^0 \to K^+\pi^-\pi^-\pi^+$, $\overline{D}^0 \to K_S\pi^+\pi^-$),
- $B^0 \to \rho^+ D^-$ ($D^- \to K^+\pi^-\pi^-$, $D^- \to K_S\pi^-$), and
- $B^0 \to \rho^+ D^{*-}$ ($D^0 \to K^+\pi^-$, $D^0 \to K^+\pi^-\pi^-\pi^+$, $D^0 \to K_S\pi^+\pi^-$),

where $\overline{D}^{*-} \to \overline{D}^0\pi^-$ is only used for the $\overline{D}^*$ decay. We choose the modes where the decay chain including π^0 consists of two-body decays only. This guarantees that the MC-generated p_{π^0} distribution is exactly the same as the distribution of data, except for the uncertainty of the longitudinal fraction of the $B^0 \to VV$ decays. Figure A.9 shows the correction factor, where the uncertainty in each p_{π^0} region is $\sim 10\%$ at largest. It turns out that the systematic error due to this uncertainty is very small.

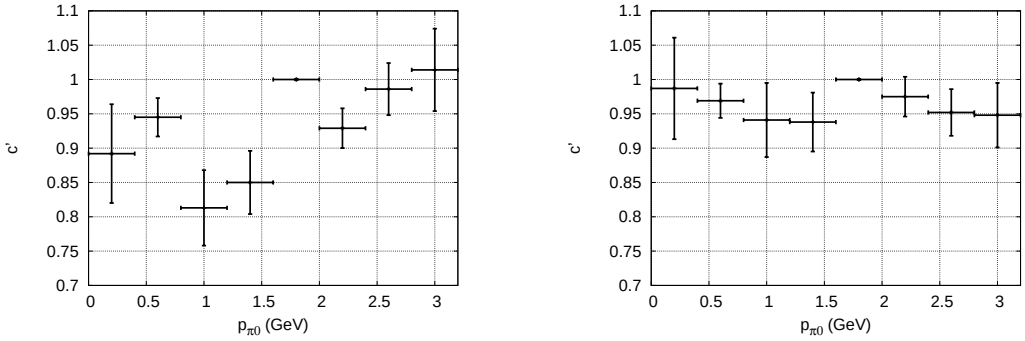


Figure A.9: The efficiency correction factor $\epsilon'(p_{\pi^0})$ for DS-I (left) and DS-II (right) obtained by the control sample study. The vertical axis corresponds to $\epsilon'_{\pi^0}(p_{\pi^0})$.

Efficiency Correction for Fudge Factor Effect

The Fudge Factor for the ΔE-$M_{\rm bc}$ PDF, combined with the p_{π^0} dependence of the ΔE-$M_{\rm bc}$ PDF, requires an efficiency correction. Figure A.10 illustrates the reason. In the region of high p_{π^0}, the ΔE distribution is wide and part of the distribution is outside of the signal region, which corresponds to an inefficiency due to the signal region cut. Note that this inefficiency is taken into account since we calculate the Dalitz-dependent efficiency after the signal region cut applied, but with MC condition, i.e., no Fudge Factor applied in the calculation. By the application of the Fudge Factor, the ΔE distribution becomes wider, since $\Delta\sigma^{\Delta E} > 1$, and the inefficiency is larger than the case without the Fudge Factor correction. This amounts to $\sim 4\%$ difference of the efficiency between the data and MC conditions, or with and without the Fudge Factor correction. In the region of low p_{π^0}, on the other hand, the ΔE distribution is so narrow that both of the distributions with and without Fudge Factor are totally inside the signal region; there is no efficiency difference due to the Fudge Factor. Consequently, the Fudge Factor effect causes the efficiency difference between data and MC that is dependent on p_{π^0}, which is to be corrected.

We calculate the correction factor, $\epsilon'_{\rm FF}(p_{\pi^0})$, as follows. The efficiency of the signal region cut $\epsilon_{\rm SR}(p_{\pi^0})$ is

$$\begin{aligned}\epsilon_{\rm SR}(p_{\pi^0}) &= \frac{\iint_{\rm SR} d\Delta E\, dM_{\rm bc}\, \mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0})}{\iint_{\rm GS} d\Delta E\, dM_{\rm bc}\, \mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0})} \\ &= \frac{1}{\iint_{\rm GS} d\Delta E\, dM_{\rm bc}\, \mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0})}\ ,\end{aligned} \tag{A.21}$$

where the second equality is due to the normalization condition of Eq. (A.8). From Eqs. (A.2) and (A.3), the efficiency for MC condition, i.e., without the Fudge Factor correction, is

$$\epsilon^{\rm MC}_{\rm SR}(p_{\pi^0}) = f_p\, n_{j_p} + (1 - f_p)\, n_{j_p+1}\ . \tag{A.22}$$

The corresponding efficiency for the data condition, i.e., with the Fudge Factor correction, is calculated in the same manner to be

$$\epsilon^{\rm Data}_{\rm SR}(p_{\pi^0}) = f_p\, n'_{j_p} + (1 - f_p)\, n'_{j_p+1}\ . \tag{A.23}$$

Thus, the correction factor $\epsilon'_{\rm FF}(p_{\pi^0})$ is

$$\begin{aligned}\epsilon'_{\rm FF}(p_{\pi^0}) &= \frac{\epsilon^{\rm Data}_{\rm SR}(p_{\pi^0})}{\epsilon^{\rm MC}_{\rm SR}(p_{\pi^0})} \\ &= \frac{f_p\, n'_{j_p} + (1 - f_p)\, n'_{j_p+1}}{f_p\, n_{j_p} + (1 - f_p)\, n_{j_p+1}}\ .\end{aligned} \tag{A.24}$$

On the Integration over the Dalitz Plot

Since p_{π^0} is correlated with m', the correlation has to be taken into account in the integration of Eq. (6.16). For this purpose, we define the efficiency correction averaged over the p_{π^0} defined as

$$\epsilon'(m') \equiv \int dp_{\pi^0} \epsilon'(p_{\pi^0}) P(p_{\pi^0}; m')\ , \tag{A.25}$$

where $P(p_{\pi^0}; m')$ is p_{π^0} distribution for the given m', corresponding to Fig. A.4. It is normalized as

$$\int dp_{\pi^0} P(p_{\pi^0}; m') = 1 \quad (\forall m')\ . \tag{A.26}$$

In the integration of Eq. (6.16), we replace the $\epsilon'(p_{\pi^0})$ with the $\epsilon'(m')$ defined above, to take account of the correlation.

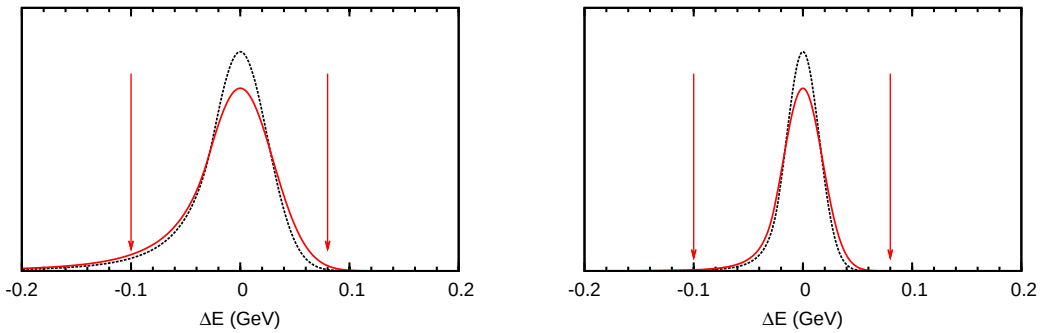


Figure A.10: Schematic illustration of the effect of Fudge Factor for high (left) and low (right) p_{π^0} regions. Solid and dotted curves correspond to the distributions with and without the Fudge Factor correction, respectively. The arrows indicate the borders of the signal region.

Efficiency for Each Resonance

As typical values of the efficiency, we give the efficiencies for $B^0 \to \rho^+\pi^-$, $B^0 \to \rho^-\pi^+$, and $B^0 \to \rho^0\pi^0$ in Table A.2. The efficiency difference between $B^0 \to \rho^+\pi^-$ and $B^0 \to \rho^-\pi^+$ is due to the charge asymmetry.

Table A.2: Efficiency for each $\rho\pi$ resonance.

DS-I		DS-II	
$\rho^+\pi^-$	7.9%	$\rho^+\pi^-$	8.5%
$\rho^-\pi^+$	8.2%	$\rho^-\pi^+$	8.7%
$\rho^0\pi^0$	9.1%	$\rho^0\pi^0$	9.4%

2 SCF

The PDF for SCF is described by Eq. (6.17). In this section, we discuss 1) the Dalitz plot PDF, which is described by Eq. (6.22); 2) the ΔE-$M_{\rm bc}$ PDF with the dependence on Dalitz plot, $\mathcal{P}_i(\Delta E, M_{\rm bc}; s_i)$ with $i =$ CR, NR; and 3) the Δt PDF of Eq. (6.19).

2-1 Dalitz Plot PDF

Figure A.11 shows the distribution of SCF events in the Dalitz plot, where the distribution is clearly separated into three parts. As discussed in Sec. 2-1 of chapter 6, the SCF is categorized into CR and NR SCF's, depending on the charge of the wrongly reconstructed track; $\pi^\pm$ (π^0) is wrongly reconstructed in CR (NR) SCF. The CR SCF can be subdivided into two by the sign of the wrong $\pi^\pm$; we call them CR(+) and CR(−) using the charge of the wrong track. In total, we have three categories, NR, CR(+), and CR(−), corresponding to the three clusters in Fig. A.11. The reason of the clustering is that the SCF events concentrate on the Dalitz plot positions where the wrongly reconstructed tracks have low momentum.

This localization of the SCF events leads to the fact that the fraction of SCF is dependent on the original distribution in Dalitz plot. Table A.3 lists the typical fractions of SCF's for

each $(\rho\pi)^0$ resonance components. One can see that $B^0 \to \rho^0\pi^0$ yields very small fraction of NR SCF, for example, which is because the π^0 from $B^0 \to \rho^0\pi^0$ has high momentum.

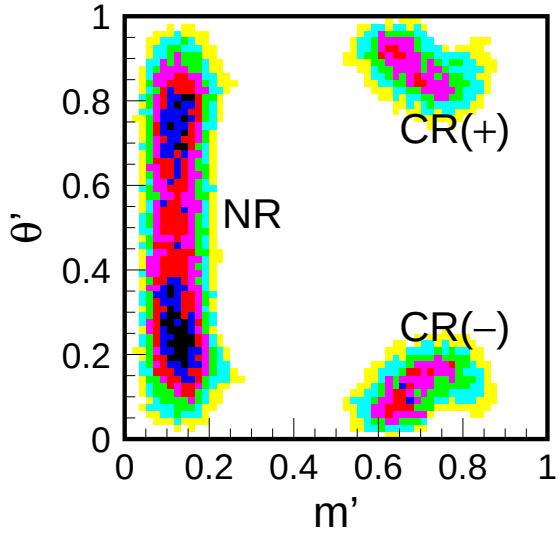


Figure A.11: Dalitz plot distribution of the SCF's.

Table A.3: Fractions of the SCF events with respect to the correctly reconstructed events.

DS-I

	CR	NR
$\rho^+\pi^-$	3.8%	22%
$\rho^-\pi^+$	3.7%	21%
$\rho^0\pi^0$	6.5%	0.3%

DS-II

	CR	NR
$\rho^+\pi^-$	3.5%	21%
$\rho^-\pi^+$	3.4%	20%
$\rho^0\pi^0$	6.1%	0.3%

Migration in Dalitz Plot and Resolution Function

As written in Sec. 2-1 of chapter 6, we describe the migration effect of the SCF events by the resolution functions $R_i(m'_{\rm obs}, \theta'_{\rm obs}; m'_{\rm gen}, \theta'_{\rm gen})$ with $i =$ CR, NR. The resolution function gives the distribution of reconstructed Dalitz plot positions, $(m'_{\rm obs}, \theta'_{\rm obs})$, for a given generated (or correct) Dalitz plot position before the migration, $(m'_{\rm gen}, \theta'_{\rm gen})$, of which Figs. A.12 and A.13 show some examples for several $(m'_{\rm gen}, \theta'_{\rm gen})$. By definition, the resolution function satisfies the following normalization condition:

$$\iint dm'_{\rm obs}\, d\theta'_{\rm obs}\, R_i(m'_{\rm obs}, \theta'_{\rm obs}; m'_{\rm gen}, \theta'_{\rm gen}) = 1 \quad (\forall m'_{\rm gen}, \forall \theta'_{\rm gen})\ . \tag{A.27}$$

In addition to the migration, we also have to take account of the detection efficiency for SCF. Not all of the events are reconstructed as SCF. Much of the events are not detected due to inefficiency, and the primary fraction of the detected events are correctly reconstructed; and only small fractions of the events become SCF. We treat this effect using efficiency functions dependent on the original Dalitz plot position, $\epsilon_i(m'_{\rm gen}, \theta'_{\rm gen})$ with $i =$ CR, NR, which are shown in Fig. A.14.

Products of the resolution functions and the efficiency functions describe the relation between the original Dalitz plot distribution and the observed Dalitz plot distribution of SCF's. We calculate the products using MC-generated events as

$$R_i(m'_{\rm obs},\theta'_{\rm obs};m'_{\rm gen},\theta'_{\rm gen})\cdot\epsilon_i(m'_{\rm gen},\theta'_{\rm gen}) = \frac{N^i_{\rm obs}(m'_{\rm obs},\theta'_{\rm obs})}{N_{\rm gen}(m'_{\rm gen},\theta'_{\rm gen})} \quad (i=\text{CR, NR}) \ , \tag{A.28}$$

where $N_{\rm gen}(m'_{\rm gen},\theta'_{\rm gen})$ and $N^{\rm CR(NR)}_{\rm obs}(m'_{\rm obs},\theta'_{\rm obs})$ are number of events generated at $(m'_{\rm gen},\theta'_{\rm gen})$ and the number of events observed at $(m'_{\rm obs},\theta'_{\rm obs})$ as CR (NR) SCF, respectively. One can easily confirm the validity of Eq. (6.22) by Eq. (A.28). We confirm there is no significant dependence of the efficiency and resolution functions on the flavor tagging quality r. For SCF, we assume no charge asymmetry, i.e., we assume the following relation

$$\begin{aligned} &R_i(m'_{\rm obs},\theta'_{\rm obs};m'_{\rm gen},\theta'_{\rm gen})\cdot\epsilon_i(m'_{\rm gen},\theta'_{\rm gen}) \\ &\qquad = R_i(m'_{\rm obs},1-\theta'_{\rm obs};m'_{\rm gen},1-\theta'_{\rm gen})\cdot\epsilon_i(m'_{\rm gen},1-\theta'_{\rm gen}) \ . \end{aligned} \tag{A.29}$$

2-2 ΔE-$M_{\rm bc}$ PDF

Figures A.15 and A.16 show the ΔE-$M_{\rm bc}$ distributions for CR and NR SCF's, respectively. As expected, the distributions are broader than that of correctly reconstructed signal. As discussed in the followings, we find significant dependence of the ΔE-$M_{\rm bc}$ distribution, ΔE in particular, on the Dalitz plot variables and take account of it.

Dependence on the Dalitz Plot in CR SCF

Examining several parameter transformations, we find that s_- [s_+] is the parameter that is highly correlated with the ΔE-$M_{\rm bc}$ distribution of CR(+) [CR(−)] SCF. This is because s_- [s_+] has one-by-one correspondence with the energy of π^+ [π^-], which is wrongly reconstructed in CR(+) [CR(−)] SCF.

Since CR(+) and CR(−) SCF's are clearly separated in Dalitz plot as shown in Fig. A.11, the following relation is satisfied:

$$s_{\rm CR} \equiv \max(s_+,s_-) = \begin{cases} s_- & \text{for CR(+)} \ , \\ s_+ & \text{for CR(−)} \ , \end{cases} \tag{A.30}$$

which allows us to parameterize the ΔE-$M_{\rm bc}$ distribution of CR SCF by the above defined $s_{\rm CR}$. Figure A.17 shows the dependence of the ΔE distribution on $s_{\rm CR}$. We prepare the ΔE-$M_{\rm bc}$ PDF dependent on $s_{\rm CR}$, $\mathcal{P}_{\rm CR}(\Delta E, M_{\rm bc}; s_{\rm CR})$, in the same manner as the $\mathcal{P}_{\rm true}(\Delta E, M_{\rm bc}; p_{\pi^0})$ described in Sec. 1-1.

Dependence on the Dalitz Plot in NR SCF

We prepare the ΔE-$M_{\rm bc}$ distribution of NR SCF in the same way as CR SCF. Here, ΔE-$M_{\rm bc}$ PDF is parameterized by

$$s_{\rm NR} \equiv s_0 \ , \tag{A.31}$$

and defined as $\mathcal{P}_{\rm NR}(\Delta E, M_{\rm bc}; s_{\rm NR})$. Figure A.18 shows the dependence of ΔE distribution on $s_{\rm NR}$. Note that the tendency of the dependence is different from that of CR SCF. This is because the energy of π^0 is not used to calculate the Dalitz plot variables, while the energy of $\pi^\pm$ is used.

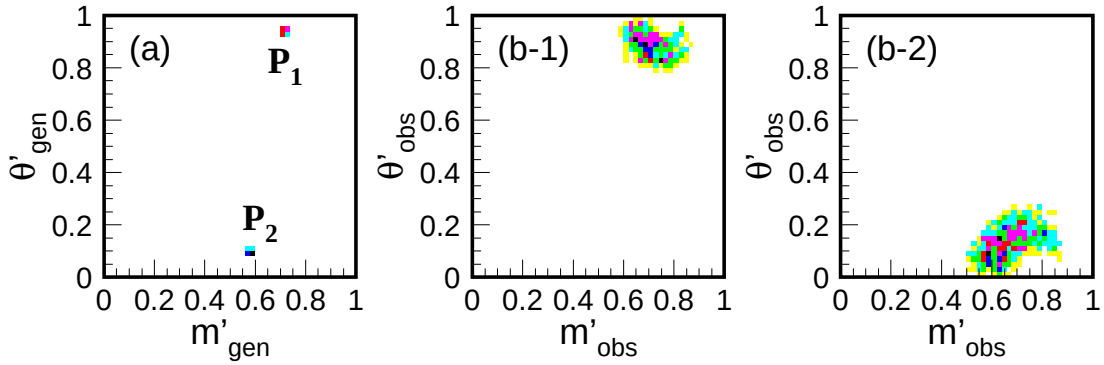


Figure A.12: Migration effect of CR SCF. The plot (a) shows the generated position, and (b-1) and (b-2) show the reconstructed position after the migration. The events generated at the positions of P_1 and P_2 in (a) migrate and distribute as (b-1) and (b-2), respectively.

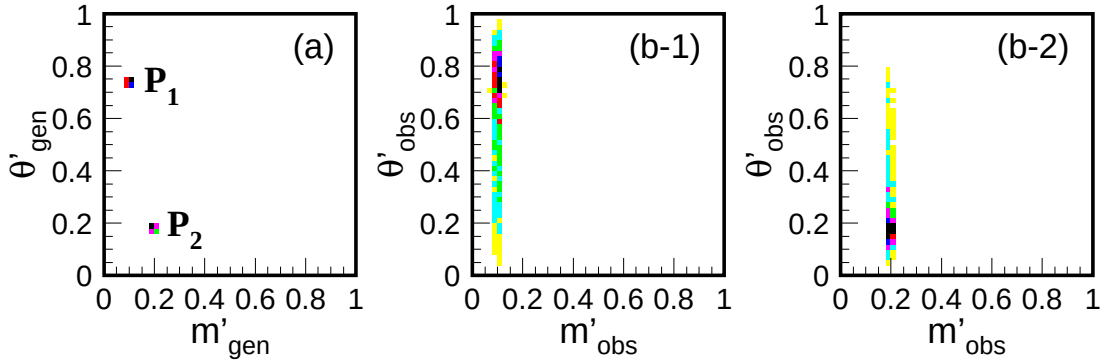


Figure A.13: Migration effect of NR SCF. The plot (a) shows the generated position, and (b-1) and (b-2) show the reconstructed position after the migration. The events generated at the positions of P_1 and P_2 in (a) migrate and distribute as (b-1) and (b-2), respectively. Note that m' is unchanged by the migration of NR SCF, since we only use the four-momenta of π^+ and π^- to calculate m'.

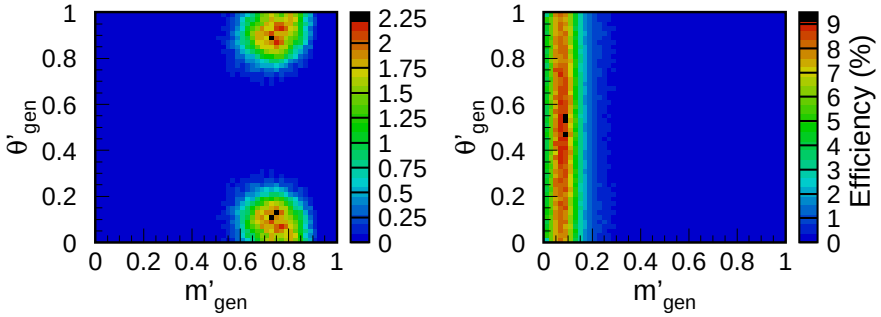


Figure A.14: The Dalitz plot efficiency of SCF, $\epsilon_i(m'_{\rm gen}, \theta'_{\rm gen})$, for CR (left) and NR (right).

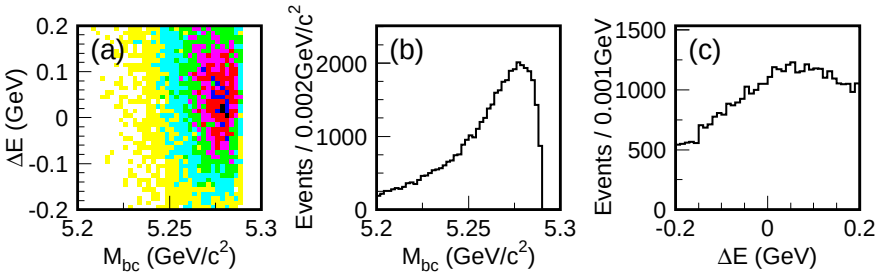


Figure A.15: The distributions of (a) ΔE-$M_{\rm bc}$, (b) $M_{\rm bc}$ projection, and (c) ΔE projection for CR SCF.

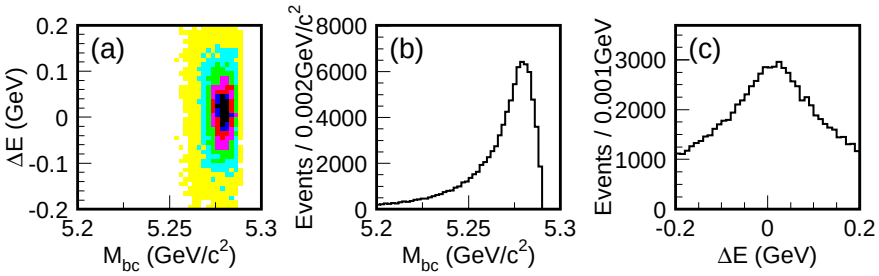


Figure A.16: The distributions of (a) ΔE-$M_{\rm bc}$, (b) $M_{\rm bc}$ projection, and (c) ΔE projection for NR SCF.

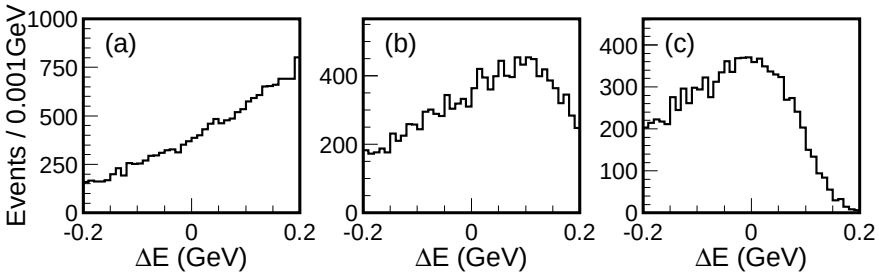


Figure A.17: ΔE distributions of CR SCF for (a) $s_{\rm CR} < 24\,{\rm GeV}^2$, (b) $24\,{\rm GeV}^2 < s_{\rm CR} < 25\,{\rm GeV}^2$, and (c) $25\,{\rm GeV}^2 < s_{\rm CR}$.

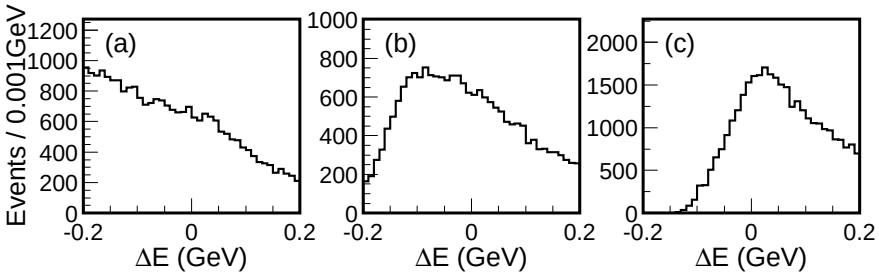


Figure A.18: ΔE distributions of NR SCF for (a) $s_{\rm NR} < 24\,{\rm GeV}^2$, (b) $24\,{\rm GeV}^2 < s_{\rm NR} < 25\,{\rm GeV}^2$, and (c) $25\,{\rm GeV}^2 < s_{\rm NR}$.

2-3 Δt PDF

As described in Sec. 2-1 of chapter 6, the results of vertex reconstruction and the flavor tagging for NR SCF is exactly the same as the correctly reconstructed signal, since the wrongly reconstructed track, π^0, is not used either of the vertex reconstruction or the flavor tagging. In CR SCF, on the other hand, the wrongly reconstructed track, $\pi^{\pm}$, is used for both of them and thus the results are deviated due to the wrong track. In the followings, we describe how we treat this effect.

Vertexing

Low momentum charged tracks are exchanged between B_{CP} and $B_{\rm tag}$ in CR SCF. This always makes the two vertices approach to each other and effectively reduce the B lifetime in the PDF of Eq. (6.19). To take account of it, we introduce an effective lifetime, $\tau_{\rm CR}$, which we determine by fitting MC sample. Figure A.19 shows the fit result, which yields

$$\tau_{\rm CR} = \begin{cases} 1.172 \pm 0.012\,{\rm ps} & \text{for DS-I}\,, \\ 1.052 \pm 0.014\,{\rm ps} & \text{for DS-II}\,. \end{cases} \tag{A.32}$$

Note that the effect in the vertex reconstruction is not large, since the exchanged tracks have low momenta and thus generally have large errors in the tracking parameters; the exchanged tracks have low significance in vertex reconstruction. Thus, the model only with the modified effective B lifetime is good enough for current statistics.

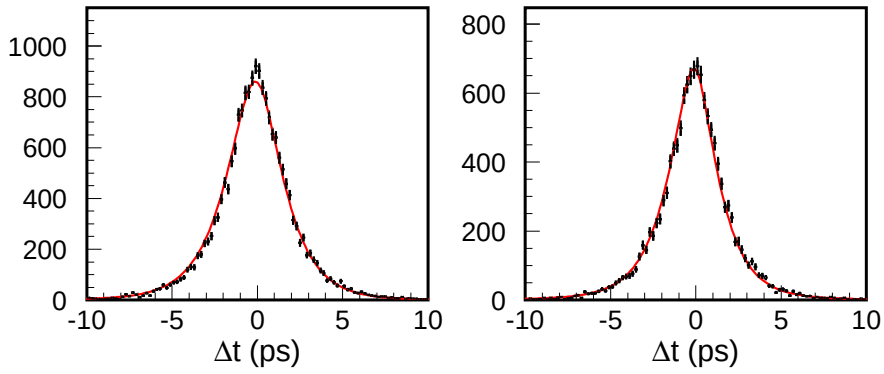


Figure A.19: Δt distributions of CR SCF for DS-I (left) and DS-II (right), overlayed with the fit results.

Wrong Tag Fraction

The wrong reconstruction of the charged track affects flavor tagging as well as vertex reconstruction, since the exchanged track is used for the flavor tagging in the tag-side. Note that the exchanged track has low momentum and thus it is explicitly used for flavor tagging as the information of the slow pion category. (See Sec. 2 of chapter 4.) Table A.4 shows the wrong-tag fractions, ${w_l}^{\rm CR}$, and wrong-tag fraction differences, $\Delta {w_l}^{\rm CR}$, calculated using MC. As discussed in the followings, $\Delta {w_l}^{\rm CR}$ have opposite sign for CR(+) SCF and CR(−) SCF, while ${w_l}^{\rm CR}$ are common; the signs of $\Delta {w_l}^{\rm CR}$ shown in the table correspond to those for CR(+) SCF.

As shown in Fig. A.20, the wrong-tag fraction differences for CR SCF have significantly non-zero values in contrast to those for correctly reconstructed signal, and have opposite sign

for CR(+) and CR(−). This can be understood as follows[1]. In CR(+) SCF, π^+ mesons with low momenta are exchanged between B_{CP} and $B_{\rm tag}$. As described in Sec. 2 of chapter 4, the slow π^+ in tag-side is used to tag the $B_{\rm tag}$ to be $\overline{B}^0$. Consequently, $B_{\rm tag}$ of CR(+) SCF tends to be tagged as $\overline{B}^0$ *irrespective of the true flavor of* $B_{\rm tag}$; the wrong-tag fraction is larger when $B_{\rm tag} = B^0$ than when $B_{\rm tag} = \overline{B}^0$. Thus, $w_+ > w_-$ and $\Delta w > 0$.

Table A.4: Wrong-tag fractions, ${w_l}^{\rm CR}$, and wrong-tag fraction differences, $\Delta {w_l}^{\rm CR}$, for CR(+) SCF. For CR(−) SCF, ${w_l}^{\rm CR}$ are the same as CR(+), while $\Delta {w_l}^{\rm CR}$ have opposite sign.

DS-I			DS-II		
l	${w_l}^{\rm CR}$	$\Delta {w_l}^{\rm CR}$	l	${w_l}^{\rm CR}$	$\Delta {w_l}^{\rm CR}$
1	0.450 ± 0.007	0.040 ± 0.014	1	0.465 ± 0.009	0.037 ± 0.017
2	0.324 ± 0.008	0.071 ± 0.016	2	0.341 ± 0.010	0.059 ± 0.020
3	0.219 ± 0.009	0.090 ± 0.017	3	0.218 ± 0.010	0.101 ± 0.021
4	0.157 ± 0.008	0.055 ± 0.016	4	0.156 ± 0.010	0.089 ± 0.019
5	0.103 ± 0.007	0.065 ± 0.013	5	0.096 ± 0.008	0.077 ± 0.017
6	0.031 ± 0.003	0.028 ± 0.006	6	0.035 ± 0.004	0.044 ± 0.007

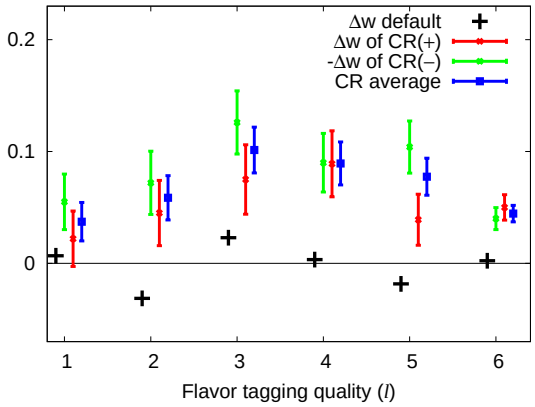


Figure A.20: Wrong-tag fraction differences for CR SCF, in comparison to those of default values, i.e., the values for correctly reconstructed signal.

On the Flavor Asymmetry of Efficiency

In addition to the wrong-tag fraction difference, we also find the efficiency of the CR SCF has significant flavor asymmetry[2], which is shown in Fig. A.21. After investigation, we find that the bias in the fit due to this effect is negligibly small and thus we ignore it.

[1] Here, we only discuss CR(+) for simplicity. The same discussion can be applied to CR(−) flipping the sign of charges.

[2] Note that the effect of flavor-asymmetric efficiency cannot be absorbed by the degree of freedom of the wrong-tag fraction difference. The time-dependent decay width including the flavor-asymmetric efficiency is

$$\frac{d\Gamma}{d\Delta t} = \frac{1}{N}\frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}}\Big\{1 - q_{\rm tag}\left[\Delta w - \Delta\epsilon'(1-2w)\right] + \left[q_{\rm tag}(1-2w-\Delta w\Delta\epsilon') + \Delta\epsilon'\right]\cdot\left[\mathcal{A}\cos(\Delta m_d\Delta t) + \mathcal{S}\cos(\Delta m_d\Delta t)\right]\Big\}, \tag{A.33}$$

The reason of the smallness of the bias is as follows. There also exists small flavor asymmetry of efficiency for correctly reconstructed events in the Dalitz plot region where the CR SCF events reside. The efficiency asymmetry has opposite sign compared to that of CR SCF and they cancels with each other. Note that the number of events of correctly reconstructed signal is much larger than that of CR SCF and the small asymmetry is sufficient to cancel the large asymmetry in the CR SCF. We try all of the following options

1. ignoring both of the asymmetries of correctly reconstructed signal and CR SCF,
2. taking account of one of them, and
3. taking account of both of them,

and find the option 1 has as small bias as the option 3, while the option 2 results in a significant bias. To keep our analysis as simple as possible, we take the option 1.

The efficiency asymmetries described above can be understood by the property of $B_{\rm tag}$ decays. In a sizable fraction of events with $B_{\rm tag} = \overline{B}^0$, the decay chain of $B_{\rm tag}$ includes D^{*+}, which yields slow π^+. Having a low momentum π^+ in tag-side, which is to be exchanged, is a necessary condition to *produce* a CR(+) SCF event, and thus the efficiency for CR(+) SCF is larger in the cases with $B_{\rm tag} = \overline{B}^0$ than in the cases with $B_{\rm tag} = B^0$. Since the efficiency increase of the SCF leads to the efficiency decrease of the correctly reconstructed signal, the tendency is opposite for the correctly reconstructed signal. The same discussion can apply to the CR(−) SCF with flipped sign of charges.

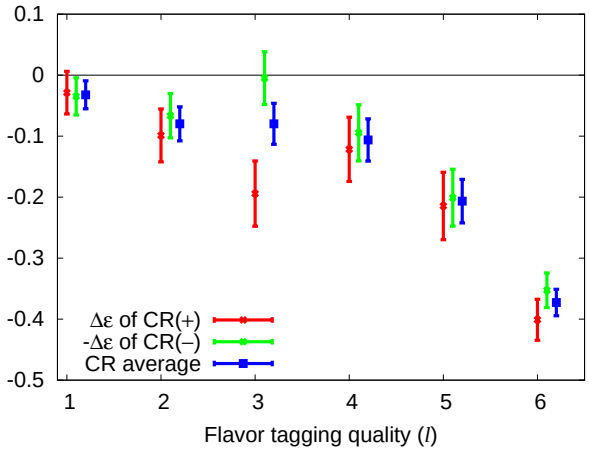


Figure A.21: Flavor asymmetry of efficiency for CR SCF.

3 Continuum Background

The PDF for continuum component is described by Eq. (6.26). In this section, we explain the ΔE-$M_{\rm bc}$ PDF, $\mathcal{P}^l_{q\overline{q}}(\Delta E, M_{\rm bc})$, the Dalitz plot PDF with the dependence on ΔE-$M_{\rm bc}$,

where

$$N = 1 + \frac{\Delta\epsilon' \mathcal{A}}{1 + (\tau_{B^0}\Delta m_d)^2} \ , \tag{A.34}$$

$$\Delta\epsilon' \equiv \frac{\epsilon_+ - \epsilon_-}{\epsilon_+ + \epsilon_-} \ . \tag{A.35}$$

Here, $\epsilon_{+(-)}$ is the efficiency for the events with $B_{\rm tag} = B^0(\overline{B}^0)$.

$\mathcal{P}_{q\overline{q}}(m', \theta'; \Delta E, M_{\rm bc})$, the Dalitz-dependent flavor asymmetry, $A^l_{q\overline{q}}(m', \theta'; \Delta E, M_{\rm bc})$, and the Δt PDF, $\mathcal{P}_{q\overline{q}}(\Delta t)$. Except for the ΔE-$M_{\rm bc}$ PDF, we obtain the PDF's from the data in the ΔE-$M_{\rm bc}$ sideband region and the dependence on the ΔE-$M_{\rm bc}$ is carefully discussed.

3-1 ΔE-$M_{\rm bc}$ PDF

Finding no significant correlation between ΔE and $M_{\rm bc}$, we factorize the ΔE-$M_{\rm bc}$ PDF and describe it as a product of ΔE PDF and $M_{\rm bc}$ PDF:

$$\mathcal{P}^l_{q\overline{q}}(\Delta E, M_{\rm bc}) = \mathcal{P}^l_{q\overline{q}}(\Delta E) \cdot \mathcal{P}_{q\overline{q}}(M_{\rm bc}) \ . \tag{A.36}$$

We adopt 1-st order polynomials and ARGUS parameterization [61] to model the ΔE and $M_{\rm bc}$ PDF's:

$$\mathcal{P}^l_{q\overline{q}}(\Delta E) = \frac{1}{N^l_{\Delta E}} \left(1 + p^l_1 \cdot \Delta E\right) \ , \tag{A.37}$$

$$\mathcal{P}_{q\overline{q}}(M_{\rm bc}) = \frac{1}{N_{M_{\rm bc}}} x\, p(x) \exp\left[\alpha\, p(x)^2\right] \ , \tag{A.38}$$

$$\left(\quad p(x) \equiv \sqrt{1 - \left(\frac{x}{E_{\rm beam}}\right)^2} \quad \right)$$

where $N^l_{\Delta E}$ and $N_{M_{\rm bc}}$ are the normalization factors determined such that

$$\int_{\rm SR} d\Delta E\, \mathcal{P}^l_{q\overline{q}}(\Delta E) \ = \ 1 \quad (\forall l) \ , \tag{A.39}$$

$$\int_{\rm SR} dM_{\rm bc}\, \mathcal{P}_{q\overline{q}}(M_{\rm bc}) \ = \ 1 \ . \tag{A.40}$$

Parameters p^l_1 and α parameterize the shapes of the PDF's and are determined by the fit to data as listed in table 5.2. In ΔE distribution, we find the distribution of continuum component significantly depends on the flavor tagging quality and thus introduce the dependence on it, while no dependence is found in the $M_{\rm bc}$ distribution and thus a common α is used for all regions. Figures A.22 and A.23 show the $M_{\rm bc}$ and ΔE projection plots for each flavor tagging quality bin, l, where one can confirm the (no)dependence on l.

3-2 Dalitz Plot PDF

The Dalitz plot PDF for continuum component is modeled using data in part of ΔE-$M_{\rm bc}$ sideband region, $-0.1\,{\rm GeV} < \Delta E < 0.2\,{\rm GeV}$ and $5.2\,{\rm GeV}/c^2 < M_{\rm bc} < 5.26\,{\rm GeV}/c^2$, where the contribution from the components other than continuum is expected to be very small. The "purity" of continuum events in this region is 96%, with negligible contamination of 4% $B\overline{B}$ background and $< 1\%$ SCF.

Since both the Dalitz plot and ΔE-$M_{\rm bc}$ are kinematic variables, the Dalitz plot distribution of continuum component has sizable dependence on ΔE-$M_{\rm bc}$, which has to be taken into account. Figure A.24 shows the dependence of the m' distribution on ΔE. Another axis of square Dalitz plot, θ', does not show significant dependence. This can be understood by the fact that the size of phasespace for the $\pi^+\pi^-\pi^0$ system is dependent on ΔE and $M_{\rm bc}$. In square Dalitz plot, the size of the phasespace, which corresponds to the rest mass of the $\pi^+\pi^-\pi^0$ system, is related to the maximum value of m_0. The maximum is $m_0^{\rm max} \equiv m_{B^0} - m_{\pi^0}$ when $\Delta E \sim 0$ and $M_{\rm bc} \sim m_{B^0}$, as written in Sec. 4-1 of chapter 4. However, the continuum background component does not satisfy this condition of $\Delta E \sim 0$ and $M_{\rm bc} \sim m_{B^0}$, in particular in the ΔE-$M_{\rm bc}$ sideband region. In such a case, the maximum of m_0 is

$$m_0^{{\rm max},q\overline{q}} = M_{3\pi} - m_{\pi^0} \ , \tag{A.41}$$

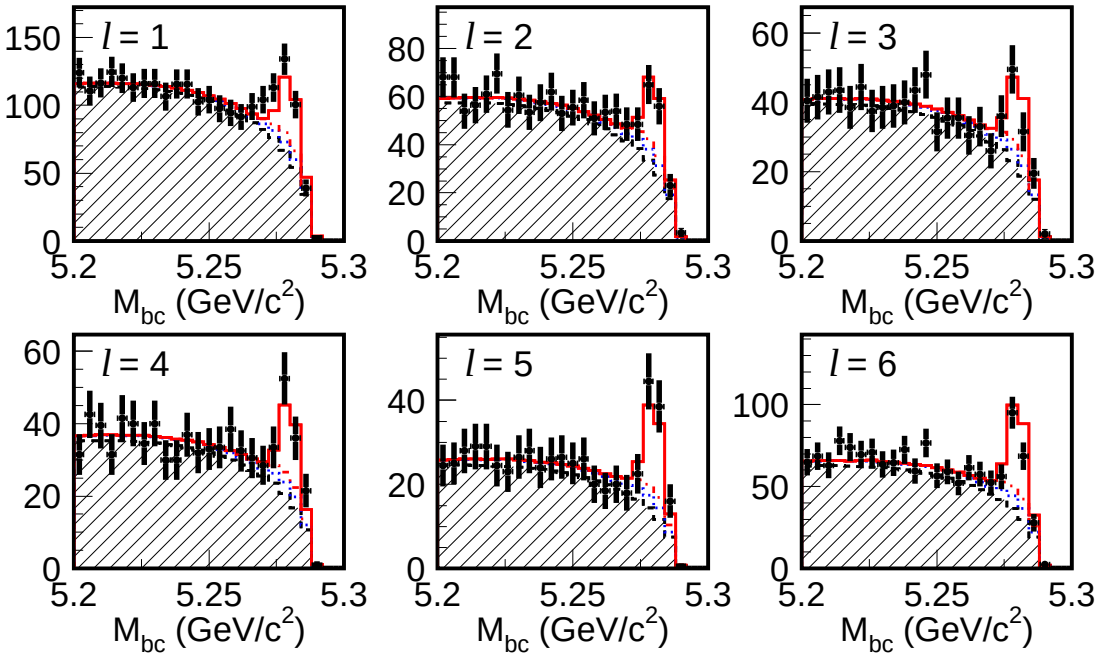


Figure A.22: Projection plot of $M_{\rm bc}$ distribution for each tagging quality region. The notation of the histograms is the same as Fig. 5.9.

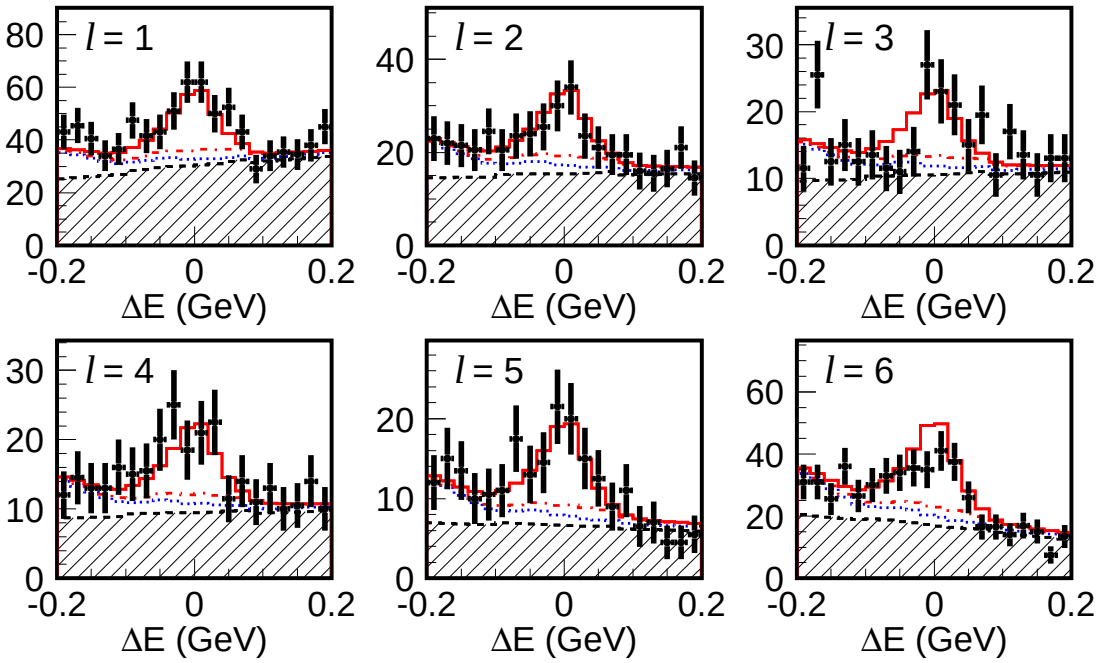


Figure A.23: Projection plot of ΔE distribution for each tagging quality region. The notation of the histograms is the same as Fig. 5.9.

where $M_{3\pi}$ is the rest mass of the $\pi^+\pi^-\pi^0$ system. Using ΔE and $\Delta M_{\rm bc} \equiv M_{\rm bc} - m_{B^0}$, the $M_{3\pi}$ is written as

$$\begin{aligned} M_{3\pi} &= \sqrt{{E_{3\pi}}^2 - |\vec{p}_{3\pi}|^2} \\ &= \sqrt{{m_{B^0}}^2 + 2E_{\rm beam}\Delta E + 2m_{B^0}\Delta M_{\rm bc} + \Delta E^2 + {\Delta M_{\rm bc}}^2} \\ &= m_{B^0} + \Delta E + \Delta M_{\rm bc} + \mathcal{O}\left(\frac{\Delta E^2}{{m_{B^0}}^2}\right) + \mathcal{O}\left(\frac{{\Delta M_{\rm bc}}^2}{{m_{B^0}}^2}\right) + \mathcal{O}\left(\frac{\Delta E \cdot (E_{\rm beam} - m_{B^0})}{{m_{B^0}}^2}\right) , \end{aligned} \tag{A.42}$$

where $(E_{3\pi}, \vec{p}_{3\pi})$ is the four-momentum of $\pi^+\pi^-\pi^0$ system and

$$E_{3\pi} = E_{\rm beam} + \Delta E , \tag{A.43}$$

$$|\vec{p}_{3\pi}|^2 = E_{\rm beam}^2 - {M_{\rm bc}}^2 , \tag{A.44}$$

by definitions of ΔE and $M_{\rm bc}$. Ignoring the second order contributions, the maximum of m_0 is

$$m_0^{{\rm max},q\overline{q}} = m_{B^0} + \Delta E + \Delta M_{\rm bc} - m_{\pi^0} . \tag{A.45}$$

Replacing the $m_0^{\rm max}$ in Eq. (4.20) with above defined $m_0^{{\rm max},q\overline{q}}$, we define the phasespace-scaled square Dalitz plot variable $m'_{\rm scale}$ by Eq. (6.33). As shown in Fig. A.25, the Dalitz plot distribution in terms of the $m'_{\rm scale}$ is independent of ΔE-$M_{\rm bc}$.

Using data in the part of the ΔE-$M_{\rm bc}$ sideband region, we prepare a ΔE-$M_{\rm bc}$ independent PDF of $\mathcal{P}(m'_{\rm scale}, \theta')$. Based on the PDF, we define the PDF in terms of (m', θ') with the dependence on ΔE-$M_{\rm bc}$ by Eq. (6.35).

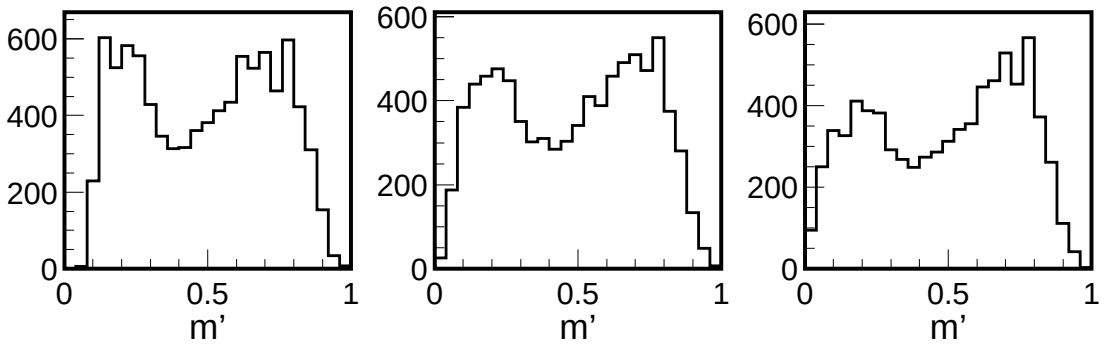


Figure A.24: m' distribution of sideband data. The three plots correspond to the data in $-0.1\,{\rm GeV} < \Delta E < 0.0\,{\rm GeV}$ (left), $0.0\,{\rm GeV} < \Delta E < 0.1\,{\rm GeV}$ (middle), and $0.1\,{\rm GeV} < \Delta E < 0.2\,{\rm GeV}$ (right).

3-3 Dalitz-Dependent Flavor Asymmetry

As written in Sec. 2-2 of chapter 6, the continuum component has significant flavor asymmetry dependent on the Dalitz plot, where the size is $\sim 20\%$ at maximum. We infer that this is due to the jet-like topology of the continuum. Though there is the hadronization process, basically the two mesons that originate from the first-produced $q\overline{q}$ have the highest momenta and opposite charges. Note that it is impossible to include both of the high-momentum tracks to reconstruct the $f_{CP} = \pi^+\pi^-\pi^0$, since ΔE becomes too large to be inside the signal region. Consequently,

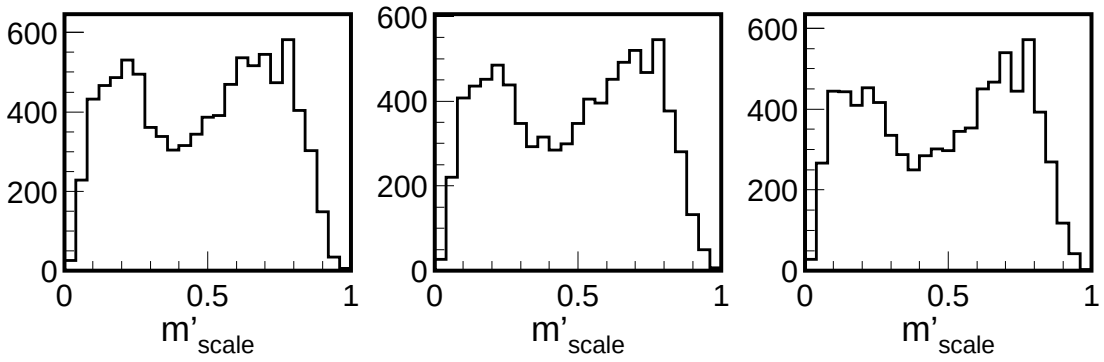


Figure A.25: $m'_{\rm scale}$ distribution of sideband data. The three plots correspond to the data in $-0.1\,{\rm GeV} < \Delta E < 0.0\,{\rm GeV}$ (left), $0.0\,{\rm GeV} < \Delta E < 0.1\,{\rm GeV}$ (middle), and $0.1\,{\rm GeV} < \Delta E < 0.2\,{\rm GeV}$ (right).

the fake B_{CP} and $B_{\rm tag}$ share the two oppositely-charged tracks with high momenta one-by-one; the highest momentum tracks in CP-side and tag-side tends to have opposite charge. To support the inference, we define following asymmetry

$$A_C^{q\overline{q}} \equiv \frac{N_{\rm OC} - N_{\rm SC}}{N_{\rm OC} + N_{\rm SC}}\,, \tag{A.46}$$

where $N_{\rm OC(SC)}$ is the number of events where the highest momentum tracks in CP-side and tag-side have opposite (same) charges. Figure A.26 shows the $A_C^{q\overline{q}}$ in the sideband region, where $A_C^{q\overline{q}}$ exhibits significantly positive values.

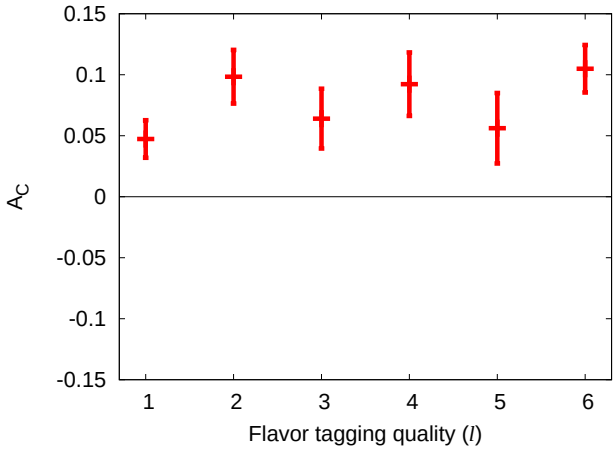


Figure A.26: The track charge asymmetry $A_C^{q\overline{q}}$ for each flavor tagging quality region.

This tendency leads to the Dalitz-dependent flavor asymmetry. CP-side has high momentum π^- (π^+) in the region of $\theta' > 0.5$ ($\theta' < 0.5$). Since $A_C > 0$, the highest momentum pion in tag-side in the corresponding event is π^+ (π^-). Due to the decay processes like $B^0 \to D^{(*)-}\pi^+$, the high momentum π^+ (π^-) in tag-side makes the $B_{\rm tag}$ to be tagged as B^0 ($\overline{B}^0$). Consequently, the continuum events in the region $\theta' > 0.5$ ($\theta' < 0.5$) tends to be tagged as $q_{\rm tag} = +1(-1)$.

We model the asymmetry by the ΔE-$M_{\rm bc}$ independent Dalitz plot, $(m'_{\rm scale}, \theta')$, by polynomials as

$$A^l_{q\overline{q}}(m'_{\rm scale}, \theta') = A(m'_{\rm scale}) \cdot A^l(\theta') \ , \tag{A.47}$$

with

$$\begin{aligned} A(m'_{\rm scale}) &= 1 + B^{m'} m' + C^{m'} m'^2 \ , & \text{(A.48)} \\ A^l(\theta') &= B^{\theta'}_l \left[(\theta' - 0.5) + D^{\theta'} (\theta' - 0.5)^3 \right] \ , & \text{(A.49)} \end{aligned}$$

where the dependence on the flavor tagging quality l is described by $B^{\theta'}_l$. Note that the coefficients of $(\theta' - 0.5)^0, (\theta' - 0.5)^2, \cdots$ are set to be zero, corresponding to the assumption of no CP-violation in the continuum component. We determine the coefficients of the polynomials by the fit to the data in the part of ΔE-$M_{\rm bc}$ sideband region, $-0.1\,{\rm GeV} < \Delta E < 0.2\,{\rm GeV}$ and $5.2\,{\rm GeV}/c^2 < M_{\rm bc} < 5.26\,{\rm GeV}/c^2$. Table A.5 lists the fit result. As shown in Fig. A.27, the fit result describes the distribution well. We also try fitting the coefficients of $(\theta' - 0.5)^0$ and $(\theta' - 0.5)^2$, which are set to be zero in the nominal fit, resulting in zero-consistent values.

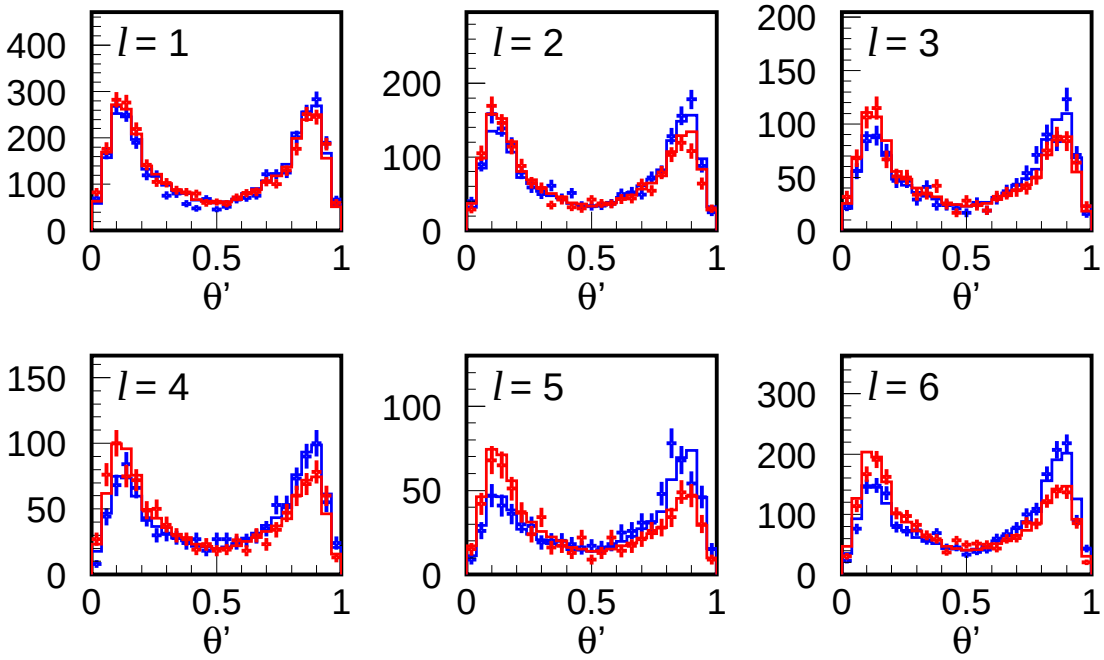


Figure A.27: Projection plots of the square Dalitz plot for sideband data in each flavor tagging quality region, overlayed with the histograms of fitted flavor asymmetry. Histograms and data points with blue (red) color correspond to $q_{\rm tag} = +1$ (-1).

3-4 Δt PDF

The Δt distribution of continuum component, $\mathcal{P}_{q\overline{q}}(\Delta t)$, is modeled by the following empirical PDF:

$$\begin{aligned} \mathcal{P}_{q\overline{q}}(\Delta t) &= [P_{q\overline{q}} \otimes R_{q\overline{q}}](\Delta t) \\ &\equiv \int d\Delta t' P_{qq}(\Delta t - \Delta t') R_{q\overline{q}}(\Delta t') \ , \end{aligned} \tag{A.50}$$

Table A.5: Parameters of the Dalitz-dependent flavor asymmetry of the continuum component.

DS-I	
Name	Value
$B^{m'}$	$+0.2 \pm 0.4$
$C^{m'}$	$-4.4^{+1.5}_{-1.3}$
$B_1^{\theta'}$	-0.04 ± 0.09
$B_2^{\theta'}$	$+0.30^{+0.15}_{-0.13}$
$B_3^{\theta'}$	$+0.29^{+0.18}_{-0.15}$
$B_4^{\theta'}$	$+0.55^{+0.22}_{-0.18}$
$B_5^{\theta'}$	$+0.71^{+0.25}_{-0.21}$
$B_6^{\theta'}$	$+0.87^{+0.20}_{-0.19}$
$D^{\theta'}$	$-2.0^{+1.3}_{-0.9}$

DS-II	
Name	Value
$B^{m'}$	$+1.1 \pm 0.3$
$C^{m'}$	-4.4 ± 0.9
$B_1^{\theta'}$	$+0.17^{+0.07}_{-0.06}$
$B_2^{\theta'}$	$+0.36^{+0.10}_{-0.09}$
$B_3^{\theta'}$	$+0.52^{+0.13}_{-0.11}$
$B_4^{\theta'}$	$+0.67^{+0.14}_{-0.13}$
$B_5^{\theta'}$	$+1.08^{+0.20}_{-0.18}$
$B_6^{\theta'}$	$+0.76^{+0.13}_{-0.12}$
$D^{\theta'}$	$-2.2^{+0.7}_{-0.5}$

where

$$P_{q\overline{q}}(\Delta t) = f_\delta \delta(\Delta t - \mu_\delta) + (1 - f_\delta) \exp\left[-\frac{|\Delta t - \mu_\tau|}{\tau}\right] , \tag{A.51}$$

$$R_{q\overline{q}}(\Delta t) = (1 - f_{\rm tail}) G(\Delta t; S_{\rm main} \cdot \sigma_{\rm vtx}) + f_{\rm tail} G(\Delta t; S_{\rm tail} \cdot \sigma_{\rm vtx}) , \tag{A.52}$$

with

$$G(x; \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{x^2}{2\sigma^2}\right] , \tag{A.53}$$

$$\sigma_{\rm vtx} = \sqrt{{\sigma_{CP}}^2 + {\sigma_{\rm tag}}^2} . \tag{A.54}$$

Here, σ_{CP} and $\sigma_{\rm tag}$ are the event-by-event errors of CP-side and tag-side vertices, respectively. Parameters f_δ, μ_δ, τ, μ_τ, $f_{\rm tail}$, $S_{\rm main}$, and $S_{\rm tail}$ parameterize the shape of the PDF. Since the distribution has significant dependence on the number of tracks used for the vertex reconstruction, we divide the events into the following two categories:

Single-track Either of the vertices of B_{CP} or $B_{\rm tag}$ or both are reconstructed by a single track.

Multi-track Both of the vertices are reconstructed by multiple tracks.

Different set of parameters are used for the two categories. We determine the parameters by the fit to the data in the sideband region. Figure A.28 shows the fit result and Table A.6 lists the determined parameters. We also confirm that the parameters has no significant dependence on ΔE-$M_{\rm bc}$ or the Dalitz plot.

4 $B\overline{B}$ Background

The $B\overline{B}$ background component is described by a linear combination of the PDF's of the decay modes that contribute as background for $\pi^+\pi^-\pi^0$ final state as in Eq. (6.38). Table A.7 lists the modes we take account of and their estimated number of events in the signal region, where the world average branching fractions [54, 55] are used for the estimation. We assume the PDF's are factorisable:

$$\begin{aligned} \mathcal{P}_k(\vec{x}) &= \mathcal{P}_k^l(\Delta E, M_{\rm bc}; m', \theta'; \Delta t, q_{\rm tag}) \\ &= \mathcal{F}_k^l \cdot \mathcal{P}_k(\Delta E, M_{\rm bc}) \cdot \mathcal{P}_k^l(m', \theta'; \Delta t, q_{\rm tag}) , \end{aligned} \tag{A.55}$$

Table A.6: Parameters of Δt PDF for continuum component determined by the fit to sideband data.

DS-I

	Single-track	Multi-track
$S_{\rm main}$	$0.965^{+0.042}_{-0.045}$	$1.076^{+0.028}_{-0.031}$
$S_{\rm tail}$	$2.289^{+0.503}_{-0.382}$	$2.347^{+0.173}_{-0.169}$
$f_{\rm tail}$	$0.130^{+0.074}_{-0.047}$	$0.220^{+0.037}_{-0.033}$
μ_δ	$-0.007^{+0.045}_{-0.042}$	$-0.045^{+0.010}_{-0.010}$
f_δ	$0.684^{+0.087}_{-0.107}$	$0.940^{+0.017}_{-0.020}$
τ	$0.964^{+0.200}_{-0.184}$	$1.838^{+0.268}_{-0.207}$
μ_τ	$+0.014^{+0.139}_{-0.131}$	$-0.035^{+0.216}_{-0.210}$

DS-II

	Single-track	Multi-track
$S_{\rm main}$	$1.034^{+0.032}_{-0.032}$	$1.083^{+0.036}_{-0.039}$
$S_{\rm tail}$	$5.277^{+0.925}_{-0.617}$	$2.136^{+0.115}_{-0.108}$
$f_{\rm tail}$	$0.071^{+0.015}_{-0.014}$	$0.368^{+0.051}_{-0.047}$
μ_δ	$0.096^{+0.038}_{-0.038}$	$-0.003^{+0.007}_{-0.007}$
f_δ	$0.447^{+0.056}_{-0.058}$	$0.816^{+0.019}_{-0.020}$
τ	$0.868^{+0.077}_{-0.072}$	$1.315^{+0.070}_{-0.063}$
μ_τ	$-0.049^{+0.055}_{-0.059}$	$-0.006^{+0.053}_{-0.053}$

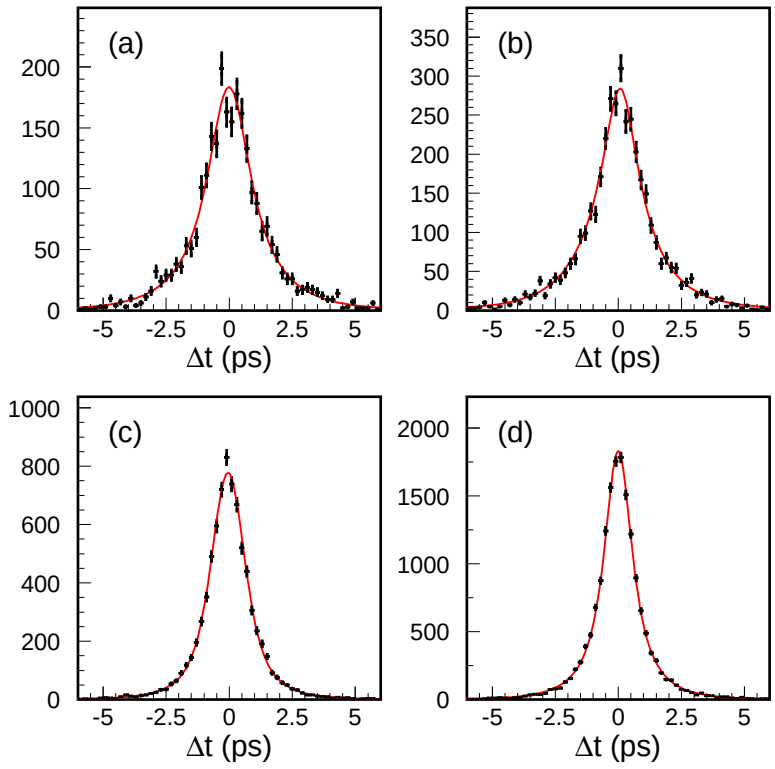


Figure A.28: Δt distributions overlayed with the fit results for (a) DS-I single-track, (b) DS-II single-track, (c) DS-I multi-track, and (d) DS-II multi-track.

where k is the index over the decay modes. In the followings, we describe in particular the treatment of $\mathcal{P}_k^l(m', \theta'; \Delta t, q_{\rm tag})$.

The treatment is different for the CP-eigenstate, flavor-specific, and charged modes, since they exhibit different characteristics in the Dalitz plot and Δt distributions. Figure A.29 shows the Dalitz plot distributions for $B^0 \to \rho^+\rho^-$, $B^0 \to D^-\pi^+$, and $B^+ \to \rho^+\pi^0$, as typical examples of CP-eigenstate, flavor-specific, and charged modes, respectively. As can be seen there, the Dalitz plot distribution of the flavor-specific (charged) mode significantly depends on the flavor (charge) of the B that is mis-reconstructed as B_{CP}, while the CP-eigenstate mode does not have such dependence by definition. Thus, we model the Dalitz plot distribution of flavor-specific (charged) modes, $\mathcal{P}_k^{\rm flv}$, with the dependence on the flavor (charge) of the B wrongly reconstructed as B_{CP}, $q_{\rm flv}$, with the following symmetry

$$\mathcal{P}_k^{\rm flv}(m', \theta'; q_{\rm flv}) = \mathcal{P}_k^{\rm flv}(m', 1-\theta'; -q_{\rm flv}) \ , \tag{A.56}$$

while the Dalitz plot distribution of CP-eigenstate modes, $\mathcal{P}_k^{CP}$, has no such dependence and satisfy the following symmetry condition

$$\mathcal{P}_k^{CP}(m', \theta') = \mathcal{P}_k^{CP}(m', 1-\theta') \ . \tag{A.57}$$

The Δt distributions of flavor-specific and charged modes also have the dependence on the flavor or charge of the B_{CP}, as shown in Fig. A.30. The following mixing PDF, $\mathcal{P}_{\rm mix}^l(\Delta t, q_{\rm tag}, q_{\rm flv})$, (lifetime PDF with flavor asymmetry, $\mathcal{P}_{\rm life}^{{\rm C},l}(\Delta t, q_{\rm tag}, q_{\rm flv})$) is used for the flavor-specific (charged) modes

$$\mathcal{P}_{\rm mix}^l(\Delta t, q_{\rm tag}, q_{\rm flv}) = \frac{1}{8\tau_{B^0}} e^{-|\Delta t|/\tau_{B^0}} \Big\{ 1 - q_{\rm tag}\Delta w_l - q_{\rm flv} q_{\rm tag}(1-2w_l)\cos(\Delta m_d \Delta t) \Big\} \ , \tag{A.58}$$

$$\mathcal{P}_{\rm life}^{{\rm C},l}(\Delta t, q_{\rm tag}, q_{\rm flv}) = \frac{1}{8\tau_{B\overline{B}}} e^{-|\Delta t|/\tau_{B\overline{B}}} \Big\{ 1 - q_{\rm flv} q_{\rm tag} \left(1 - 2w_l^{\rm C}\right) \Big\} \ , \tag{A.59}$$

where $q_{\rm flv}$, $\tau_{B\overline{B}}$, and $w_l^{\rm C}$ are the flavor (charge) of the B wrongly reconstructed as B_{CP}, the effective lifetime for the charged mode, and the effective wrong-tag fraction for the charged mode, respectively. The effective-lifetime, $\tau_{B\overline{B}}$, could be different from the true lifetime of B^+ due to the same reason as the CR SCF (Sec. 2-3), since at least one charged track has to be wrongly assigned to B_{CP} or $B_{\rm tag}$ for charged modes to mimic the $\pi^+\pi^-\pi^0$ final state; we determine the $\tau_{B\overline{B}}$ by the fit to MC-generated events. The wrong-tag fraction for charged modes, $w_l^{\rm C}$, can be different from the nominal wrong tag fraction, w_l, which is for neutral B decays, and we determine the $w_l^{\rm C}$ using MC, too. Table A.8 lists the determined parameters. For CP-eigenstate modes, we use the CP-violating time-dependent PDF

$$\begin{aligned} \mathcal{P}_{CP}^l(\Delta t, q_{\rm tag}) = \frac{1}{4\tau_{B^0}} e^{-|\Delta t|/\tau_{B^0}} \Big\{ & 1 - q_{\rm tag}\Delta w_l \\ & + q_{\rm tag}(1-2w_l)\Big[\mathcal{A}\cos(\Delta m_d \Delta t) + \mathcal{S}\sin(\Delta m_d \Delta t) \Big] \Big\} \ , \end{aligned} \tag{A.60}$$

where $\mathcal{A} = 0$ and $\mathcal{S} = \sin 2\phi$ with the corresponding weak phase of ϕ.

With the Dalitz plot and time-dependent PDF's described above, we define the Dalitz-Δt PDF's for the CP-eigenstate, flavor-specific, and charged modes by

$$\mathcal{P}_k^l(m', \theta'; \Delta t, q_{\rm tag}) = \mathcal{P}_k^{CP}(m', \theta')\, \mathcal{P}_{CP}^l(\Delta t, q_{\rm tag}) \ , \tag{A.61}$$

$$\mathcal{P}_k^l(m', \theta'; \Delta t, q_{\rm tag}) = \sum_{q_{\rm flv}=\pm 1} \mathcal{P}_k^{\rm flv}(m', \theta'; q_{\rm flv})\, \mathcal{P}_{\rm mix}^l(\Delta t, q_{\rm tag}, q_{\rm flv}) \ , \tag{A.62}$$

$$\mathcal{P}_k^l(m', \theta'; \Delta t, q_{\rm tag}) = \sum_{q_{\rm flv}=\pm 1} \mathcal{P}_k^{\rm flv}(m', \theta'; q_{\rm flv})\, \mathcal{P}_{\rm life}^{{\rm C},l}(\Delta t, q_{\rm tag}, q_{\rm flv}) \ , \tag{A.63}$$

respectively.

Table A.7: The decay modes taken into account as the $B\overline{B}$ background, with the expected numbers of events inside signal region, $N_{\rm ev}$, among the 2,824 events.

$b \to u, d, s$ decay modes

Mode	$N_{\rm ev}$
$\rho^+\rho^-$	11.2
$\rho^+\rho^0$	8.9
$a_1(1260)^\pm\pi^\mp$	9.7
$\rho^0\pi^+$	34.4
$\rho^+\pi^0$	13.5
$K_0^*(1430)^0\pi^+$	3.3
$K_0^*(1430)^+\pi^-$	10.5
$K^*(892)^+\pi^-$	8.3
ρ^-K^+	34.7
$\pi^+\pi^0$	3.9
$\eta\pi^+\pi^-$	0.2
$K^0\pi^+$	3.7
$K^+\pi^+\pi^-$	0.1
$\pi^0\eta'(958)$	2.3
$K^*(892)^0\pi^0$	2.8
$K^0\pi^0$	0.6
$K^+\pi^0$	1.7
$\omega(782)\pi^+$	0.8
$\eta'(958)K^+$	0.5
ρ^0K^+	1.4
$K^*(892)^0\gamma$	5.6

$b \to c$ decay modes

Mode	$N_{\rm ev}$
$\overline{D}^0\pi^+$	31.6
$D^-\pi^+$	21.5
$\overline{D}^0\rho^+$	7.0
$\overline{D}^{*0}\pi^+$	2.1
$D^-\rho+$	0.4
$\overline{D}^0\pi^0$	2.5
$J/\psi\pi^0$	2.5
Other charged	1.4
Other flavor-specific	6.0

Table A.8: Parameters used for the Δt PDF of charged mode $B\overline{B}$ background.

DS-I

Name	Value
$\tau_{B\overline{B}}$	1.53 ± 0.06 (ps)
$w_1^{\rm C}$	0.41 ± 0.03
$w_2^{\rm C}$	0.32 ± 0.04
$w_3^{\rm C}$	0.23 ± 0.04
$w_4^{\rm C}$	0.24 ± 0.04
$w_5^{\rm C}$	0.12 ± 0.03
$w_6^{\rm C}$	0.02 ± 0.01

DS-II

Name	Value
$\tau_{B\overline{B}}$	1.38 ± 0.04 (ps)
$w_1^{\rm C}$	0.42 ± 0.02
$w_2^{\rm C}$	0.33 ± 0.03
$w_3^{\rm C}$	0.25 ± 0.03
$w_4^{\rm C}$	$0.14\ ^{+0.03}_{-0.02}$
$w_5^{\rm C}$	0.07 ± 0.02
$w_6^{\rm C}$	0.04 ± 0.01

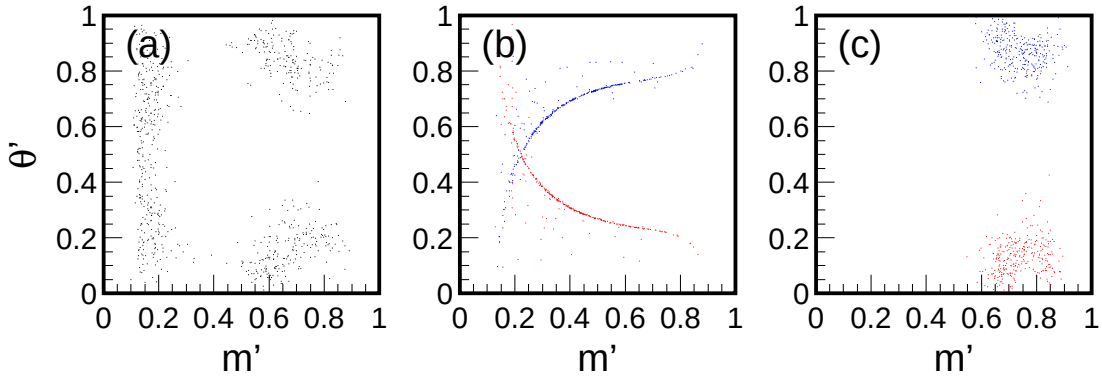


Figure A.29: The Dalitz plot distributions of (a) $B^0 \to \rho^+\rho^-$, (b) $B^0 \to D^-\pi^+$, and (c) $B^+ \to \rho^+\pi^0$, as examples of the $B\overline{B}$ background of CP-eigenstate, flavor-specific, and charged modes, respectively. The points colored red (blue) correspond to the events with $B_{CP} = B^0$ ($\overline{B}^0$) and $B_{CP} = B^+$ (B^-) in the figures (b) and (c), respectively.

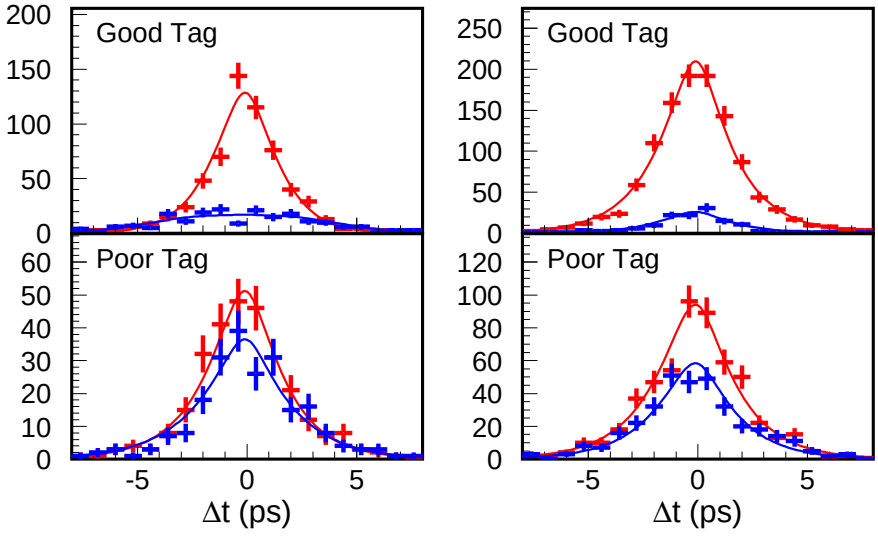


Figure A.30: The Δt distributions of $B\overline{B}$ background component of flavor-specific (left) and charged (right) categories overlayed with the PDF curves, where red (blue) histograms and curves correspond to $q_{\rm tag}q_{\rm flv} = -1$ $(+1)$. The upper and lower plots correspond to good tag ($r > 0.5$) and poor tag ($r < 0.5$) regions, respectively.

5 Outlier

As described in Ref. [46], the resolution function includes an overall very-bad-resolution component $P_{\rm ol}(\Delta t)$, which we call outlier, to take account of a very long tail that is observed both in data and MC. We model the outlier by a Gaussian with $\sigma \sim 30\,{\rm ps}$ and its fraction among all component, $f_{\rm ol}$, is $\sim 10^{-4}$ ($\sim 10^{-2}$) for the events with the vertices reconstructed using multiple (single) tracks. When Δt is the only observable for the fit, the total PDF including the outlier component, $P_{\rm tot}(\Delta t)$, is

$$P_{\rm tot}(\Delta t) = (1 - f_{\rm ol})P(\Delta t) + f_{\rm ol}P_{\rm ol}(\Delta t) \ , \tag{A.64}$$

where $P(\Delta t)$ is the PDF for the main part including both of signal and background contributions, which is analogous to the $P(\vec{x})$ in Eq. (6.8).

In our time-dependent Dalitz plot analysis, the treatment is more complex than that described above. The outlier is understood to be the events with wrongly reconstructed tracks used in the vertex reconstruction. Here, the wrong reconstruction means wrong association with the SVD hits and the other information, such as momentum and dE/dx, are correctly measured. Thus, we model the outlier PDF such that its ΔE-$M_{\rm bc}$, Dalitz plot, and flavor-tagging part is the same as the main part; the ΔE-$M_{\rm bc}$, Dalitz plot, and flavor-tagging part of the outlier PDF, $P_{\rm ol}(\Delta E, M_{\rm bc}; m', \theta'; q_{\rm tag}, l)$, is

$$P_{\rm ol}(\Delta E, M_{\rm bc}; m', \theta'; q_{\rm tag}, l) = \int d\Delta t \, P(\vec{x}) \ , \tag{A.65}$$

where $P(\vec{x})$ is the PDF of the main part defined in Eq. (6.8). Using this PDF, the total PDF including the outlier part, $P_{\rm tot}(\vec{x})$, is written as

$$P_{\rm tot}(\vec{x}) = (1 - f_{\rm ol})P(\vec{x}) + f_{\rm ol}P_{\rm ol}(\Delta E, M_{\rm bc}; m', \theta'; q_{\rm tag}, l) \, P_{\rm ol}(\Delta t) \ . \tag{A.66}$$

We use this $P_{\rm tot}(\vec{x})$ for the fit.

Appendix B

Method of ϕ_2 Constraint

1 Formalism

We define amplitudes as

$$A^+ \equiv A(B^0 \to \rho^+\pi^-) , \tag{B.1}$$
$$A^- \equiv A(B^0 \to \rho^-\pi^+) , \tag{B.2}$$
$$A^0 \equiv A(B^0 \to \rho^0\pi^0) , \tag{B.3}$$
$$A^{+0} \equiv A(B^+ \to \rho^+\pi^0) , \tag{B.4}$$
$$A^{0+} \equiv A(B^+ \to \rho^0\pi^+) , \tag{B.5}$$

and

$$\overline{A}^+ \equiv \frac{p}{q}A(\overline{B}^0 \to \rho^+\pi^-) , \tag{B.6}$$
$$\overline{A}^- \equiv \frac{p}{q}A(\overline{B}^0 \to \rho^-\pi^+) , \tag{B.7}$$
$$\overline{A}^0 \equiv \frac{p}{q}A(\overline{B}^0 \to \rho^0\pi^0) , \tag{B.8}$$
$$A^{-0} \equiv \frac{p}{q}A(B^- \to \rho^-\pi^0) , \tag{B.9}$$
$$A^{0-} \equiv \frac{p}{q}A(B^- \to \rho^0\pi^-) . \tag{B.10}$$

These amplitudes are obtained from 1) 26 measurements determined in the time-dependent Dalitz plot analysis as well as 2) branching fractions and asymmetry measurements, and give a constraint on ϕ_2.

Equations (4.25)-(4.28) define the relations between the amplitudes for the neutral modes and the parameters determined in the time-dependent Dalitz plot analysis. The relations between the branching fractions and asymmetries, and the amplitudes are

$$\mathcal{B}(\rho^\pm\pi^\mp) = c \cdot \left(|A^+|^2 + |A^-|^2 + |\overline{A}^+|^2 + |\overline{A}^-|\right) \cdot \tau_{B^0} , \tag{B.11}$$

$$\mathcal{B}(\rho^+\pi^0) = c \cdot \left(|A^{+0}|^2 + |A^{-0}|^2\right) \cdot \tau_{B^+} , \tag{B.12}$$
$$\mathcal{B}(\rho^0\pi^+) = c \cdot \left(|A^{0+}|^2 + |A^{0-}|^2\right) \cdot \tau_{B^+} , \tag{B.13}$$
$$\mathcal{A}(\rho^+\pi^0) = \frac{|A^{-0}|^2 - |A^{+0}|^2}{|A^{-0}|^2 + |A^{+0}|^2} , \tag{B.14}$$

$$\mathcal{A}(\rho^0\pi^+) = \frac{|A^{0-}|^2 - |A^{0+}|^2}{|A^{0-}|^2 + |A^{0+}|^2} , \tag{B.15}$$

where c is a constant and the lifetimes τ_{B^0} and τ_{B^+} are introduced to take account of the total width difference between B^0 and B^+. Note that we do not use quasi-two-body parameters related to neutral modes except for $\mathcal{B}(\rho^\pm\pi^\mp)$, since they are included in the Dalitz plot parameters.

The amplitudes are expected to follow $SU(2)$ isospin symmetry to a good approximation [17, 18]

$$\begin{aligned} A^+ + A^- + 2A^0 &= \widetilde{A}^+ + \widetilde{A}^- + 2\tilde{A}^0 \\ &= \sqrt{2}(A^{+0} + A^{0+}) = \sqrt{2}(\widetilde{A}^{-0} + \widetilde{A}^{0-}) , \end{aligned} \tag{B.16}$$

$$A^{+0} - A^{0+} - \sqrt{2}(A^+ - A^-) = \widetilde{A}^{-0} - \widetilde{A}^{0-} - \sqrt{2}(\widetilde{A}^- - \widetilde{A}^+) , \tag{B.17}$$

where

$$\widetilde{A}^\kappa \equiv e^{-2i\phi_2}\overline{A}^\kappa , \quad \widetilde{A}^{-0} \equiv e^{-2i\phi_2}A^{-0} , \quad \text{and} \quad \widetilde{A}^{0-} \equiv e^{-2i\phi_2}A^{0-} . \tag{B.18}$$

Note that there is an inconsistency in equation (B.17) between Ref. [17] and Ref. [18]; we follow the treatment of Ref. [17], which we believe is correct.

2 Parameterization

Here we give two examples of the parameterization of the amplitudes. The first example may be more intuitive, while the second example is well behaved in the fit. The results are independent of the parameterizations with respect to the constraint on ϕ_2.

2-1 Amplitude parameterization

We can parameterize the amplitudes as follows [17]

$$A^+ = e^{-i\phi_2}T^+ + P^+ , \tag{B.19}$$
$$A^- = e^{-i\phi_2}T^- + P^- , \tag{B.20}$$
$$A^0 = e^{-i\phi_2}T^0 - \frac{1}{2}(P^+ + P^-) , \tag{B.21}$$
$$\sqrt{2}A^{+0} = e^{-i\phi_2}T^{+0} + P^+ - P^- , \tag{B.22}$$
$$\sqrt{2}A^{0+} = e^{-i\phi_2}(T^+ + T^- + 2T^0 - T^{+0}) - P^+ + P^- , \tag{B.23}$$

and

$$\overline{A}^+ = e^{+i\phi_2}T^- + P^- , \tag{B.24}$$
$$\overline{A}^- = e^{+i\phi_2}T^+ + P^+ , \tag{B.25}$$
$$\overline{A}^0 = e^{+i\phi_2}T^0 - \frac{1}{2}(P^+ + P^-) , \tag{B.26}$$
$$\sqrt{2}A^{-0} = e^{+i\phi_2}T^{+0} + P^+ - P^- , \tag{B.27}$$
$$\sqrt{2}A^{0-} = e^{+i\phi_2}(T^+ + T^- + 2T^0 - T^{+0}) - P^+ + P^- , \tag{B.28}$$

where the overall phase is fixed with the convention $\mathrm{Im}T^+ = 0$. Thus, there are 6 complex amplitudes, T^+, T^-, T^0, P^+, P^-, and T^{+0}, corresponding to 11 degrees of freedom; and ϕ_2, corresponding to 12 degrees of freedom in total. This parameterization automatically satisfies the isospin relations without losing generality, i.e., the isospin relations are the only assumption here.

2-2 Geometric parameterization

We can parameterize the amplitudes using the geometric arrangement of Fig. B.1 that satisfies the isospin relation of equation (B.16). This figure is equivalent to Fig. 3 of Ref. [18], except that the sides corresponding to $B^0 \to \rho^-\pi^+$ and $B^0 \to \rho^0\pi^0$ are swapped. This difference is not physically significant. We apply this modification only to obtain a better behaved parameterization; the parameterization here uses the angles ω_- and θ_- related to the process $B^0(\overline{B}^0) \to \rho^-\pi^+$, which are better behaved than those related to $B^0(\overline{B}^0) \to \rho^0\pi^0$.

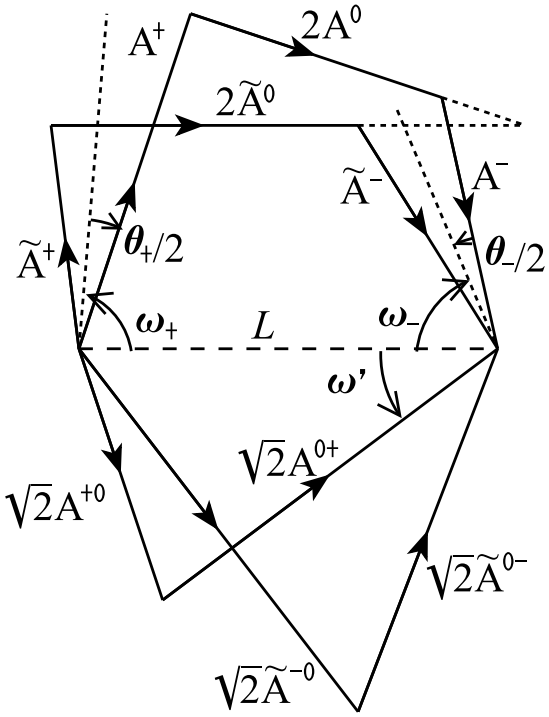


Figure B.1: Complex pentagons formed from the $B \to \rho\pi$ decay amplitudes.

To parameterize the amplitudes, we use ϕ_2 and the following 11 geometric parameters:

$$\omega_+, \omega_-, \omega', \theta_+, \theta_-, b_+, b_-, b', a_+, a_-, L, \tag{B.29}$$

where b and a *imply* branching fraction and asymmetry, respectively. In terms of these parameters, the amplitudes can be described as follows

$$A^+ = e^{i(\omega_+ + \theta_+/2)}\sqrt{b_+(1-a_+)/2}\,, \tag{B.30}$$
$$\widetilde{A}^+ = e^{i(\omega_+ - \theta_+/2)}\sqrt{b_+(1+a_+)/2}\,, \tag{B.31}$$
$$A^- = e^{i(\omega_- + \theta_-/2)}\sqrt{b_-(1-a_-)/2}\,, \tag{B.32}$$
$$\widetilde{A}^- = e^{i(\omega_- - \theta_-/2)}\sqrt{b_-(1+a_-)/2}\,, \tag{B.33}$$
$$A^0 = (L - A^+ - A^-)/2\,, \tag{B.34}$$
$$\widetilde{A}^0 = (L - \widetilde{A}^+ - \widetilde{A}^-)/2\,, \tag{B.35}$$

$$A^{0+} = e^{i\omega'}\sqrt{b'/2}\,, \tag{B.36}$$

$$A^{+0} = \frac{L}{\sqrt{2}} - A^{0+}\,, \tag{B.37}$$

$$\widetilde{A}^{0-} = \frac{L}{2\sqrt{2}} - \left[A^{+0} - A^{0+} - \sqrt{2}(A^{+} - A^{-}) + \sqrt{2}(\widetilde{A}^{-} - \widetilde{A}^{+})\right]/2\,, \tag{B.38}$$

$$\widetilde{A}^{-0} = \frac{L}{\sqrt{2}} - \widetilde{A}^{0-}\,. \tag{B.39}$$

Equation (B.38) exploits the isospin relation of equation (B.17), which Fig. B.1 does not incorporate geometrically. The phase ϕ_2 enters when the $\widetilde{A}$'s are converted into $\overline{A}$'s with equation (B.18). When we perform the analysis only with the time-dependent Dalitz plot observables and without the information from charged decay modes, we remove the parameters ω' and b' from the fit and fix L to be a constant.

This geometric parameterization has a substantial advantage in terms of required computational resources, compared to the parameterization based on the T and P amplitudes described in the previous section. In the procedure to constrain ϕ_2, the minimum χ^2 has to be calculated for each value of ϕ_2. To avoid local minima, initial values of the parameters for the minimization have to be scanned and this inflates the computing time, which increases exponentially with the number of parameters. However, the number of parameters to be scanned decreases in the geometric parameterization. Among 11 parameters except for ϕ_2, five of them, b_+, b_-, b', a_+, and a_-, are related to the branching fractions and asymmetries. Since in most cases they do not have multiple solutions, we do not have to scan the initial values of them. In addition, the optimum initial value for L can also be determined using other parameters and b_0, the nominal branching fraction of $B^0 \to \rho^0\pi^0$, from the following relation

$$b_0 = \left|L - e^{i\omega_+}\sqrt{b_+/2} - e^{i\omega_-}\sqrt{b_-/2}\right|^2\,, \tag{B.40}$$

up to a two-fold ambiguity. Here b_0 is calculated using the input parameters as

$$b_0 = \frac{U_0^+}{U_+^+ + U_-^+} \cdot \frac{\mathcal{B}(\rho^\pm\pi^\mp)}{c \cdot \tau_{B^0}}\,, \tag{B.41}$$

based on equations (4.25) and (B.11). The explicit solution for the optimal initial value of L is

$$L = \mathrm{Re}\gamma \pm \sqrt{b_0 - (\mathrm{Im}\gamma)^2} \quad \left(\quad \gamma \equiv e^{i\omega_+}\sqrt{b_+/2} + e^{i\omega_-}\sqrt{b_-/2} \quad \right)\,. \tag{B.42}$$

When $b_0 - (\mathrm{Im}\gamma)^2 < 0$, there is no real-valued solution and $L = \mathrm{Re}\gamma$ is the optimum initial value. With the optimum values calculated above, the initial value of L does not have to be scanned, except for the two fold ambiguity. Consequently, the number of parameters to be scanned in this parameterization is only five, corresponding to $\omega_+, \omega_-, \omega', \theta_+$, and θ_-, while all of 11 or maybe 10 parameters have to be scanned in the T and P amplitude parameterization. This leads to a substantial reduction of the computational resources required.

Appendix C

Useful Equations Related to Square Dalitz Plot

In this appendix, we derive some relations between the usual Dalitz Plot and the square Dalitz plot. Since the normal Dalitz Plot variables are Lorentz invariant, it is convenient to consider in the ρ^0 rest frame as in the figure C.1, where the Square Dalitz Plot variables are defined. We

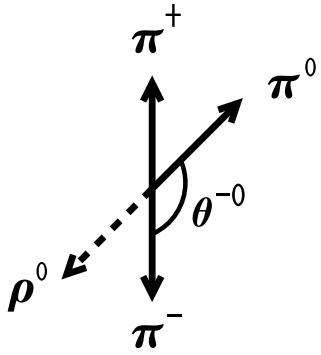


Figure C.1: Kinematics in the ρ^0 rest frame.

follow the definitions in Sec. 4-1 of chapter 4. The parameter transformation from the Dalitz plot to the mass and helicity of $\rho^{\pm}$ is also discussed.

1 Transformation From Usual Dalitz Plot to Square Dalitz Plot

First we derive the square Dalitz plot variables m_0 and $\theta_0(\equiv \theta^{-0})$ as functions of the usual Dalitz Plot variables (s_+, s_-). By Eq. (2.157), the m_0 is expressed as

$$m_0 = \sqrt{s_0} = \sqrt{M^2 - (s_+ + s_-)}\,. \tag{C.1}$$

$$\left(\quad M^2 \equiv m_{B^0}^2 + 2m_{\pi^+}^2 + m_{\pi^0}^2 \quad \right) \tag{C.2}$$

Since we are in the rest frame of ρ^0, the four momenta p_+, p_-, and p_0 in this frame are

$$p_+ \equiv (E_+, \vec{p}_+) = (\frac{m_0}{2}, -\vec{p}_-) \,, \tag{C.3}$$

$$p_- \equiv (E_-, \vec{p}_-) = (\frac{m_0}{2}, \vec{p}_-) \,, \tag{C.4}$$

$$p_0 \equiv (E_0, \vec{p}_0) \,. \tag{C.5}$$

Since

$$s_+ = (E_- + E_0)^2 - (\vec{p}_0 - \vec{p}_-)^2 \,, \tag{C.6}$$

$$s_- = (E_- + E_0)^2 - (\vec{p}_0 + \vec{p}_-)^2 \,, \tag{C.7}$$

the E_0 can be calculated as

$$E_0 = \frac{s_+ + s_- - 2(m_{\pi^+}^2 + m_{\pi^0}^2)}{4E_-} \tag{C.8}$$

$$= \frac{s_+ + s_- - 2(m_{\pi^+}^2 + m_{\pi^0}^2)}{2\sqrt{M^2 - (s_+ + s_-)}} \,. \tag{C.9}$$

The product $\vec{p}_- \cdot \vec{p}_0$ can be expressed using θ_0 as

$$\vec{p}_- \cdot \vec{p}_0 = |\vec{p}_-||\vec{p}_0| \cos\theta_0 \,. \tag{C.10}$$

Subtracting the equation (C.7) from (C.6) and using the equation (C.10), $\cos\theta_0$ is

$$\cos\theta_0 = \frac{s_+ - s_-}{4|\vec{p}_-||\vec{p}_0|} \,, \tag{C.11}$$

$$\left(\quad |\vec{p}_-| = \frac{\sqrt{M^2 - 4m_{\pi^+}^2 - (s_+ + s_-)}}{2} \quad \right) \,, \tag{C.12}$$

$$\left(\quad |\vec{p}_0| = \sqrt{{E_0}^2 - m_{\pi^0}^2} \quad \right) \,. \tag{C.13}$$

These three equations, together with the equations (C.1) and (C.9), describe the parameter transformation of $(s_+, s_-) \mapsto (m_0, \cos\theta_0)$. ■

Then we calculate the Jacobian of this transformation. As one can see in the equations above, the combinations of s_+ and s_- that appear in m_0 and $\cos\theta_0$ are only $s_1 \equiv s_+ + s_-$ and $s_2 \equiv s_+ - s_-$. It is thus convenient to divide the transformation into two steps,

$$(s_+, s_-) \mapsto (s_1, s_2) \tag{C.14}$$

and

$$(s_1, s_2) \mapsto (m_0, \cos\theta_0) \,, \tag{C.15}$$

for the calculation of the Jacobian. The Jacobian of the first transformation is quite trivial:

$$\frac{\partial(s_1, s_2)}{\partial(s_+, s_-)} = \begin{pmatrix} +1 & +1 \\ +1 & -1 \end{pmatrix} \,. \tag{C.16}$$

Since m_0 only depends on s_1 but not on s_2, the Jacobian of the second transformation is

$$\frac{\partial(m_0, \cos\theta_0)}{\partial(s_1, s_2)} = \begin{pmatrix} \partial m_0/\partial s_1 & 0 \\ \partial\cos\theta_0/\partial s_1 & \partial\cos\theta_0/\partial s_2 \end{pmatrix} \,. \tag{C.17}$$

Thus we only have to calculate $\partial m_0/\partial s_1$ and $\partial \cos\theta_0/\partial s_2$ to obtain the determinant of the Jacobian, which is what we need. They are

$$\frac{\partial m_0}{\partial s_1} = \frac{1}{2}\frac{-1}{\sqrt{M^2-(s_+ + s_-)}} = -\frac{1}{2m_0} \tag{C.18}$$

and

$$\frac{\partial \cos\theta_0}{\partial s_2} = \frac{1}{4|\vec{p}_-||\vec{p}_0|} \ . \tag{C.19}$$

The determinant of the Jacobian for the transformation $(s_+, s_-) \mapsto (m_0, \cos\theta_0)$ is

$$\left|\frac{\partial(m_0, \cos\theta_0)}{\partial(s_+, s_-)}\right| = \frac{1}{4|\vec{p}_-||\vec{p}_0|m_0} \ , \tag{C.20}$$

and thus

$$ds_+\, ds_- = 4|\vec{p}_-||\vec{p}_0|\, m_0\, dm_0\, d(\cos\theta_0) \ . \blacksquare \tag{C.21}$$

The Jacobian for the transformation $(s_+, s_-) \mapsto (m', \theta')$ is

$$\left|\frac{\partial(m', \theta')}{\partial(s_+, s_-)}\right| = \frac{1}{4|\vec{p}_-||\vec{p}_0|m_0} \cdot \frac{dm'}{dm_0}\frac{d\theta'}{d(\cos\theta_0)} \ , \tag{C.22}$$

and thus

$$ds_+\, ds_- = 4|\vec{p}_-||\vec{p}_0|\, m_0\, \frac{dm_0}{dm'}\, \frac{d\cos\theta_0}{d\theta'}\, dm'\, d\theta' \ , \tag{C.23}$$

where

$$\frac{d\cos\theta_0}{d\theta'} = -\pi \sin\theta_0 \tag{C.24}$$

and

$$\frac{dm_0}{dm'} = -\frac{m_0^{\max} - m_0^{\min}}{2}\pi \sin\pi m' \tag{C.25}$$

since

$$\cos\pi m' = 2\frac{m_0 - m_0^{\min}}{m_0^{\max} - m_0^{\min}} - 1 \ . \blacksquare \tag{C.26}$$

2 Transformation to the Mass and Helicity of $\rho^{\pm}$

It is also convenient to have equations for transformations into the m and θ in the $\rho^{\pm}$ rest frames: $(s_+, s_-) \mapsto (m_{\pm}, \cos\theta_{\pm})$. Though we only consider the ρ^+ rest frame here as in the figure C.2, it can easily converted into the ρ^- rest frame by swapping the sign + and −. In this frame, the Dalitz Plot variables are m_+ and $\theta_+ (\equiv \theta^{+-})$. The four momenta p_+, p_-, and p_0 in this frame are

$$p_+ \equiv (E_+, \vec{p}_+) \ , \tag{C.27}$$
$$p_0 \equiv (E_0, \vec{p}_0) = (E_0, -\vec{p}_+) \ , \tag{C.28}$$
$$p_- \equiv (E_-, \vec{p}_-) \ . \tag{C.29}$$

The essential difference from the equations (C.3)-(C.5) is that $E_+ \neq E_0$ due to the mass difference between π^+ and π^0. Since

$$s_+ = (E_+ + E_0)^2 = {m_+}^2 \ , \tag{C.30}$$
$$s_- = (E_- + E_0)^2 - (\vec{p}_- - \vec{p}_+)^2 \ , \tag{C.31}$$
$$s_0 = (E_- + E_+)^2 - (\vec{p}_- + \vec{p}_+)^2 \ , \tag{C.32}$$

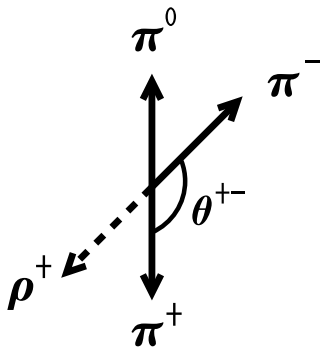


Figure C.2: Kinematics in the ρ^+ rest frame.

following relation holds

$$s_- + s_0 = 3m_{\pi^+}^2 + m_{\pi^0}^2 + 2E_-(E_0 + E_+) \,. \tag{C.33}$$

Using this relation and the equation (2.157), the E_- is calculated as

$$E_- = \frac{s_- + s_0 - (3m_{\pi^+}^2 + m_{\pi^0}^2)}{2(E_0 + E_+)} \tag{C.34}$$

$$= \frac{M^2 - 3m_{\pi^+}^2 - m_{\pi^0}^2 - s_+}{2\sqrt{s_+}} \,. \tag{C.35}$$

Using the following relations

$$E_0 = \sqrt{{m_{\pi^0}}^2 + |\vec{p}_+|^2} \,, \tag{C.36}$$

$$E_+ = \sqrt{{m_{\pi^+}}^2 + |\vec{p}_+|^2} \,, \tag{C.37}$$

and the equation (C.30), the $|\vec{p}_+|$ is described by s_+ as

$$|\vec{p}_+|^2 = \frac{{s_+}^2 + \left({m_{\pi^+}}^2 - {m_{\pi^0}}^2\right)^2}{4s_+} - \frac{m_{\pi^+}^2 + m_{\pi^0}^2}{2} \,. \tag{C.38}$$

From the equation (C.31), $\cos\theta_+$ is

$$\cos\theta_+ = \frac{s_- - \left({m_{\pi^+}}^2 + {m_{\pi^0}}^2\right) - 2E_-E_0}{2|\vec{p}_-||\vec{p}_+|} \,, \tag{C.39}$$

$$\left(\quad |\vec{p}_-| = \sqrt{{E_-}^2 - {m_{\pi^+}}^2} \quad \right) \,. \tag{C.40}$$

These equations, together with the equations (C.30), (C.35), and (C.36), describe the transformation $(s_+, s_-) \mapsto (m_+, \cos\theta_+)$. ■

Then we calculate the Jacobian of this transformation. Since m_+ only depends on s_+ but not on s_-, the Jacobian has the shape of

$$\frac{\partial(m_+, \cos\theta_+)}{\partial(s_+, s_-)} = \begin{pmatrix} \partial m_+/\partial s_+ & 0 \\ \partial\cos\theta_+/\partial s_+ & \partial\cos\theta_+/\partial s_- \end{pmatrix} \,. \tag{C.41}$$

Thus we only need $\partial m_+/\partial s_+$ and $\partial \cos\theta_+/\partial s_-$ to obtain the Jacobian's determinant. They are easily calculated as

$$\frac{\partial m_+}{\partial s_+} = \frac{1}{2m_+} \tag{C.42}$$

and

$$\frac{\partial \cos\theta_+}{\partial s_-} = \frac{1}{2|\vec{p}_-||\vec{p}_+|} \, . \tag{C.43}$$

The determinant of the Jacobian is

$$\left| \frac{\partial(m_+, \cos\theta_+)}{\partial(s_+, s_-)} \right| = \frac{1}{4|\vec{p}_-||\vec{p}_+|m_+} \tag{C.44}$$

and thus

$$ds_+ \, ds_- = 4|\vec{p}_-||\vec{p}_+|m_+ \, dm_+ \, d(\cos\theta_+) \, . \tag{C.45}$$

These results have exactly the same forms as the equations (C.20) and (C.21).[1] ■

[1] Note that the definition of the $\vec{p}_+$ in the equations (C.20) and (C.21) is different from that in the equations (C.44) and (C.45).

Appendix D

Formulas Related to ρ Kinematics

1 Equivalence of Two Expressions for $T^{\kappa}_{J=1}$

As mentioned in Sec. 1-1 of chapter 6, there are two equivalent expressions for $T^{\kappa}_{J=1}$, which take account of the helicity distribution of $B^0 \to \rho^{\kappa}(\to \pi^i\pi^j)\pi^{\overline{\kappa}(=k)}$:

$$T^{\kappa}_1 = -4|\vec{p}_j||\vec{p}_k|\cos\theta^{jk} \ , \tag{D.1}$$

and

$$T^{\kappa}_1 = s^{ki} - s^{jk} + \frac{({m_{B^0}}^2 - {m_{\pi^k}}^2)({m_{\pi^j}}^2 - {m_{\pi^i}}^2)}{s^{ij}} \ . \tag{D.2}$$

In this section, we derive the equivalence.

Since the former is defined in the rest frame of $\pi^i\pi^j$ and the latter is evidently a Lorentz-invariant expression, it is convenient to consider in the rest frame of $\pi^i\pi^j$ (Fig. D.1). In this system, the four-momenta of three pions are

$$\begin{aligned} p_i &= (E_i, \vec{p}_i) = (E_i, -\vec{p}_j) \ , \\ p_j &= (E_j, \vec{p}_j) \ , \\ p_k &= (E_k, \vec{p}_k) \ , \end{aligned} \tag{D.3}$$

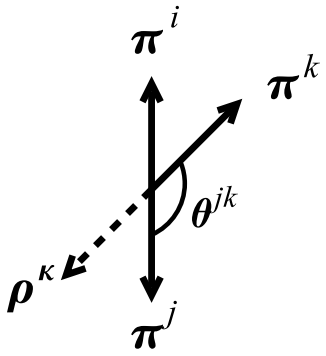


Figure D.1: The relation between three pions in the rest frame of ρ^{κ}.

which leads to

$$\begin{aligned} s^{ki} - s^{jk} &= (p_k + p_i)^2 - (p_j + p_k)^2 \\ &= ({m_{\pi^k}}^2 + {m_{\pi^i}}^2 + 2p_k \cdot p_i) - ({m_{\pi^j}}^2 + {m_{\pi^k}}^2 + 2p_j \cdot p_k) \\ &= {m_{\pi^i}}^2 - {m_{\pi^j}}^2 + 2p_k \cdot \hat{p} \ , \end{aligned} \tag{D.4}$$

$$\left(\quad \hat{p} \equiv p_i - p_j = (E_i - E_j, \ -2\vec{p}_j) \quad \right) \tag{D.5}$$

and

$$\begin{aligned} {m_{B^0}}^2 &= (p_i + p_j + p_k)^2 \\ &= (p_i + p_j)^2 + 2p_k \cdot (p_i + p_j) + {p_k}^2 \\ &= (E_i + E_j)^2 + 2E_k(E_i + E_j) + {m_{\pi^k}}^2 \ . \end{aligned} \tag{D.6}$$

Equations (D.4) and (D.6) leads

$$\begin{aligned} & s^{ij} \cdot (s^{ki} - s^{jk}) + ({m_{B^0}}^2 - {m_{\pi^k}}^2)({m_{\pi^j}}^2 - {m_{\pi^i}}^2) \\ &= (E_i + E_j)^2 \cdot \left({m_{\pi^i}}^2 - {m_{\pi^j}}^2 + 2p_k \cdot \hat{p} \right) \\ &\quad + \Big[(E_i + E_j)^2 + 2E_k(E_i + E_j) \Big] \cdot \left({m_{\pi^j}}^2 - {m_{\pi^i}}^2 \right) \\ &= (E_i + E_j)^2 \cdot \Big[{m_{\pi^i}}^2 - {m_{\pi^j}}^2 + 2E_k(E_i - E_j) - (-4\vec{p}_k \cdot \vec{p}_j) \Big] \\ &\quad + \Big[(E_i + E_j)^2 + 2E_k(E_i + E_j) \Big] \cdot \left({m_{\pi^j}}^2 - {m_{\pi^i}}^2 \right) \\ &= (E_i + E_j)^2 \cdot 4(\vec{p}_k \cdot \vec{p}_j) \\ &\quad + 2E_k(E_i + E_j) \Big[(E_i + E_j)(E_i - E_j) + \left({m_{\pi^j}}^2 - {m_{\pi^i}}^2 \right) \Big] \\ &= s^{ij} \cdot 4(\vec{p}_k \cdot \vec{p}_j) \ , \end{aligned} \tag{D.7}$$

where the last equality is from

$${E_i}^2 - {E_j}^2 = \left({m_{\pi^i}}^2 + |\vec{p}_i|^2 \right) - \left({m_{\pi^j}}^2 + |\vec{p}_j|^2 \right) = {m_{\pi^i}}^2 - {m_{\pi^j}}^2 \ . \tag{D.8}$$

Thus,

$$s^{ki} - s^{jk} + \frac{({m_{B^0}}^2 - {m_{\pi^k}}^2)({m_{\pi^j}}^2 - {m_{\pi^i}}^2)}{s^{ij}} = 4(\vec{p}_k \cdot \vec{p}_j) \ , \tag{D.9}$$

and it is equivalent to the right hand side of equation (D.1), except for the meaningless sign of -1. ■

2 Breit-Wigner Function of Gounaris-Sakurai Parameterization

In this section, we review the relativistic Breit-Wigner function proposed by Gounaris and Sakurai [51]. The Gounaris-Sakurai Breit-Wigner function $BW(s)$ for a resonance with a mass M and width Γ decaying to two pions is

$$BW(s) = \frac{M^2(1 + d \cdot \Gamma/M)}{M^2 - s + f(s) - iM\Gamma(s)} \ , \tag{D.10}$$

where

$$f(s) = \frac{\Gamma M^2}{{p_\pi}^3(M^2)} \left\{ {p_\pi}^2(s) \left[h(s) - h(M^2) \right] + (M^2 - s) \, {p_\pi}^2(M^2) \frac{dh}{ds}\bigg|_{s=M^2} \right\} \ , \tag{D.11}$$

$$h(s) = \frac{2}{\pi}\frac{p_\pi(s)}{\sqrt{s}}\ln\frac{\sqrt{s}+2p_\pi(s)}{2m_\pi}\ , \tag{D.12}$$

and the energy dependence of the width if P-wave type:

$$\Gamma(s) = \Gamma\cdot\left[\frac{p_\pi(s)}{p_\pi(M^2)}\right]^3\left[\frac{M^2}{s}\right]^{1/2}\ . \tag{D.13}$$

Here, $p_\pi(s)$ is the pion momentum in the rest frame of the resonance calculated as

$$p_\pi(s) = \sqrt{\frac{s}{4} - {m_\pi}^2} \tag{D.14}$$

and d is a normalization factor chosen to satisfy $BW(0) = 1$[1]:

$$d = \frac{3}{\pi}\frac{{m_\pi}^2}{{p_\pi}^2(M^2)}\ln\frac{M+2p_\pi(M^2)}{2m_\pi} + \frac{M}{2\pi\, p_\pi(M^2)} - \frac{{m_\pi}^2 M}{\pi\,{p_\pi}^2(M^2)}\ . \tag{D.15}$$

[1]Though this normalization is not essential in our analysis, pion form factor is in general required to satisfy this condition of normalization.

Appendix E

Putting Constraint on the ρ Lineshape Variation

As discussed in Sec. 1 of chapter 6, $B^0(\overline{B}^0) \to (\rho\pi)^0$ consists of six decay processes of

$$\begin{aligned} B^0 &\to \rho^+\pi^- , \\ B^0 &\to \rho^-\pi^+ , \\ B^0 &\to \rho^0\pi^0 , \\ \overline{B}^0 &\to \rho^+\pi^- , \\ \overline{B}^0 &\to \rho^-\pi^+ , \\ \overline{B}^0 &\to \rho^0\pi^0 , \end{aligned}$$

and the lineshape of ρ can be different for all these processes. In the nominal fit, however, we assume Eq. (6.7), i.e., the common lineshape for all the six processes. Since this assumption is not well grounded, we have to estimate the systematic error due to this assumption. In this appendix, we describe the detailed procedure of the systematic error estimation.

1 Formalism

We rewrite the exact equation without the assumption, Eq. (6.6), using the nominal lineshape parameters of β and γ as

$$\overset{(}{F}{}^{\kappa}_{\pi}(s) = BW_{\rho(770)} + (\beta + \Delta\overset{(}{\beta}_{\kappa})\, BW_{\rho(1450)} + (\gamma + \Delta\overset{(}{\gamma}_{\kappa})\, BW_{\rho(1700)} , \tag{E.1}$$

where

$$\Delta\overset{(}{\beta}_{\kappa} \equiv \overset{(}{\beta}_{\kappa} - \beta , \quad \text{and} \tag{E.2}$$

$$\Delta\overset{(}{\gamma}_{\kappa} \equiv \overset{(}{\gamma}_{\kappa} - \gamma , \tag{E.3}$$

describe the deviations of $(\overset{(}{\beta}_{\kappa}, \overset{(}{\gamma}_{\kappa})$ from (β, γ).

For convenience, we define f^{ρ}_{κ}, $f^{\rho'}_{\kappa}$ and $f^{\rho''}_{\kappa}$ as

$$f^{\rho^{(n)}}_{\kappa} \equiv T^{\kappa}_{f} BW_{\rho^{(n)}}(s_{\kappa}) . \tag{E.4}$$

The functions f_{κ} and $\overline{f}_{\kappa}$ are then

$$\overset{(}{f}_{\kappa} = f^{\rm avg}_{\kappa} + \Delta\overset{(}{\beta}_{\kappa} f^{\rho'}_{\kappa} + \Delta\overset{(}{\gamma}_{\kappa} f^{\rho''}_{\kappa} , \tag{E.5}$$

where

$$f_\kappa^{\rm avg} \equiv f_\kappa^\rho + \beta f_\kappa^{\rho'} + \gamma f_\kappa^{\rho''} \ . \tag{E.6}$$

Since all of the diagrams can be classified into two types, the ones with weak phase ϕ_2 and the others with weak phase 0, the amplitudes $A^\kappa(\overline{A}^\kappa)$ can be written as (See Sec. 4-1 of chapter 2)

$$A^\kappa = e^{-i\phi_2}\, T^\kappa + P^\kappa \ , \tag{E.7}$$
$$\overline{A}^\kappa = e^{+i\phi_2}\, T^{\overline{\kappa}} + P^{\overline{\kappa}} \ , \tag{E.8}$$

where $(\overline{+}, \overline{-}, \overline{0}) = (-, +, 0)$, i.e., $T^+ = T^{\overline{-}}$, $T^- = T^{\overline{+}}$, etc. Note that the amplitudes T and P does not simply describe one tree and penguin diagrams but represents several diagrams with common weak phases. Since strong interaction does not violate CP, the lineshapes can be attributed to each T and P amplitudes and thus

$$f_\kappa A^\kappa = e^{-i\phi_2}\, f_\kappa^T T^\kappa + f_\kappa^P P^\kappa \ , \tag{E.9}$$
$$\overline{f}_\kappa \overline{A}^\kappa = e^{+i\phi_2}\, f_{\overline{\kappa}}^T T^{\overline{\kappa}} + f_{\overline{\kappa}}^P P^{\overline{\kappa}} \ , \tag{E.10}$$

where $f_\kappa^{T(P)}$ is

$$f_\kappa^{T(P)} = f_\kappa^{\rm avg} + \Delta\beta_\kappa^{T(P)} f_\kappa^{\rho'} + \Delta\gamma_\kappa^{T(P)} f_\kappa^{\rho''} \ . \tag{E.11}$$

Equations (E.7), (E.8), (E.9), and (E.10) lead to the following relations

$$\Delta\beta_\kappa = \frac{e^{-i\phi_2}\, \Delta\beta_\kappa^T T^\kappa + \Delta\beta_\kappa^P P^\kappa}{e^{-i\phi_2}\, T^\kappa + P^\kappa} \ , \tag{E.12}$$

$$\Delta\overline{\beta}_\kappa = \frac{e^{+i\phi_2}\, \Delta\beta_{\overline{\kappa}}^T T^{\overline{\kappa}} + \Delta\beta_{\overline{\kappa}}^P P^{\overline{\kappa}}}{e^{+i\phi_2}\, T^{\overline{\kappa}} + P^{\overline{\kappa}}} \ , \tag{E.13}$$

$$\Delta\gamma_\kappa = \frac{e^{-i\phi_2}\, \Delta\gamma_\kappa^T T^\kappa + \Delta\gamma_\kappa^P P^\kappa}{e^{-i\phi_2}\, T^\kappa + P^\kappa} \ , \tag{E.14}$$

$$\Delta\overline{\gamma}_\kappa = \frac{e^{+i\phi_2}\, \Delta\gamma_{\overline{\kappa}}^T T^{\overline{\kappa}} + \Delta\gamma_{\overline{\kappa}}^P P^{\overline{\kappa}}}{e^{+i\phi_2}\, T^{\overline{\kappa}} + P^{\overline{\kappa}}} \ . \tag{E.15}$$

Assuming the isospin relation[1]

$$P^0 = -\frac{1}{2}\left(P^+ + P^-\right) \tag{E.16}$$

holds for radial excitations as well as $\rho(770)$, $\Delta\beta_0^P$ and $\Delta\gamma_0^P$ can be described as

$$\Delta\beta_0^P = \frac{\Delta\beta_+^P P^+ + \Delta\beta_-^P P^-}{P^+ + P^-} \ , \tag{E.17}$$

$$\Delta\gamma_0^P = \frac{\Delta\gamma_+^P P^+ + \Delta\gamma_-^P P^-}{P^+ + P^-} \ . \tag{E.18}$$

To summarize, the deviations of the lineshape from the nominal one are described with 10 complex valued parameters: $\Delta\beta_{+,-,0}^T, \Delta\beta_{+,-}^P, \Delta\gamma_{+,-,0}^T, \Delta\gamma_{+,-}^P$, corresponding to 20 real valued degrees of freedom. The purpose here is to put constraint on these parameters and to estimate the impact of the deviation on the final result.

2 Procedure

Here we describe the procedure. It proceeds by three steps, as

[1] With Eqs. (E.7) and (E.8), this relation is equivalent to the first equality of Eq. (2.141).

1. Obtaining constraint on the parameters related to the deviation in higher resonance part by time-integrated (and flavor-integrated) Dalitz Plot fit,
2. Putting constraint on the 10 complex parameters, $\Delta\beta^T_{+,-,0}, \Delta\beta^P_{+,-}, \Delta\gamma^T_{+,-,0}, \Delta\gamma^P_{+,-}$, from the fit result of the first step, assuming T^κ, P^κ, and ϕ_2 obtained by our nominal time-dependent Dalitz plot fit, and
3. Estimating the systematic errors from the variation of the 10 complex parameters.

3 First step: obtaining constraint by real data fit

With the assumption of one common lineshape for all amplitudes, the time-integrated (and flavor-integrated) Dalitz Plot PDF is described as (See Sec. 4-3 of chapter 2)

$$\begin{aligned}\frac{d\Gamma}{ds_+ds_-} \propto & \sum_{\kappa\in\{+,-,0\}} |f_\kappa|^2 \left(|A^\kappa|^2 + |\overline{A}^\kappa|^2\right) \\ & + 2 \sum_{\kappa<\sigma\in\{+,-,0\}} \mathrm{Re}[f_\kappa f_\sigma^*]\mathrm{Re}\left[A^\kappa A^{\sigma *} + \overline{A}^\kappa \overline{A}^{\sigma *}\right] \\ & - 2 \sum_{\kappa<\sigma\in\{+,-,0\}} \mathrm{Im}[f_\kappa f_\sigma^*]\mathrm{Im}\left[A^\kappa A^{\sigma *} + \overline{A}^\kappa \overline{A}^{\sigma *}\right] \ .\end{aligned} \tag{E.19}$$

The effect of the lineshape variation appears as corrections to either the first line (non-interfering terms) or the second and third lines (interfering terms). To obtain the information relevant to the higher resonance deviations, we fit 39 additional parameters that parameterize the corrections together with nominal 8 $(= 9-1)$ time-integrated Dalitz plot coefficients. The 39 parameters can be classified into the two categories: 15 quasi-two-body related parameters and 24 interference related parameters.

3-1 15 quasi-two-body related parameters

With the lineshape variation, the quasi-two-body term becomes

$$|f_\kappa|^2|A^\kappa|^2 + |\overline{f}_\kappa|^2|\overline{A}^\kappa|^2 \ . \tag{E.20}$$

We take following parametrization[2] to describe the deviation of $|f_\kappa|^2$ from $|f^{\mathrm{avg}}_\kappa|^2$

[2]The deviation of the $|f_\kappa|^2$ term from the average lineshape f^{avg}_κ is

$$\begin{aligned}& |f_\kappa|^2 - |f^{\mathrm{avg}}_\kappa|^2 \\ & = |\Delta\beta_\kappa|^2\,|f^{\rho'}_\kappa|^2 + |\Delta\gamma_\kappa|^2\,|f^{\rho''}_\kappa|^2 \\ & \quad +2\mathrm{Re}[\Delta\beta_\kappa]\,\mathrm{Re}[f^{\rho'}_\kappa f^{\mathrm{avg}*}_\kappa] - 2\mathrm{Im}[\Delta\beta_\kappa]\,\mathrm{Im}[f^{\rho'}_\kappa f^{\mathrm{avg}*}_\kappa] \\ & \quad +2\mathrm{Re}[\Delta\gamma_\kappa]\,\mathrm{Re}[f^{\rho'}_\kappa f^{\mathrm{avg}*}_\kappa] - 2\mathrm{Im}[\Delta\gamma_\kappa]\,\mathrm{Im}[f^{\rho''}_\kappa f^{\mathrm{avg}*}_\kappa] \\ & \quad +2\mathrm{Re}[\Delta\beta_\kappa\Delta\gamma^*_\kappa]\,\mathrm{Re}[f^{\rho'}_\kappa f^{\rho''*}_\kappa] - 2\mathrm{Im}[\Delta\beta_\kappa\Delta\gamma^*_\kappa]\,\mathrm{Im}[f^{\rho'}_\kappa f^{\rho''*}_\kappa] \ .\end{aligned} \tag{E.21}$$

The deviation of the $|\overline{f}_\kappa|^2$ can also be written in the same way. Ideally, the parametrization to exploit the information to the full is something like

$$\begin{aligned}& |f_\kappa|^2|A^\kappa|^2 + |\overline{f}_\kappa|^2|\overline{A}^\kappa|^2 \\ & = |f^{\mathrm{avg}}_\kappa|^2\,U^+_\kappa + |f^{\rho'}_\kappa|^2\,U^{2+}_\kappa + |f^{\rho''}_\kappa|^2\,U^{3+}_\kappa \\ & \quad +2\mathrm{Re}[f^{\rho'}_\kappa f^{\mathrm{avg}*}_\kappa]\,U^{21+,\mathrm{Re}}_\kappa - 2\mathrm{Im}[f^{\rho'}_\kappa f^{\mathrm{avg}*}_\kappa]\,U^{21+,\mathrm{Im}}_\kappa \\ & \quad +2\mathrm{Re}[f^{\rho''}_\kappa f^{\mathrm{avg}*}_\kappa]\,U^{31+,\mathrm{Re}}_\kappa - 2\mathrm{Im}[f^{\rho''}_\kappa f^{\mathrm{avg}*}_\kappa]\,U^{31+,\mathrm{Im}}_\kappa \\ & \quad +2\mathrm{Re}[f^{\rho''}_\kappa f^{\rho'*}_\kappa]\,U^{32+,\mathrm{Re}}_\kappa - 2\mathrm{Im}[f^{\rho''}_\kappa f^{\rho'*}_\kappa]\,U^{32+,\mathrm{Im}}_\kappa \ ,\end{aligned} \tag{E.22}$$

where U^+_κ is the nominal non-interfering Dalitz plot Dalitz Plot coefficients of Eq. (4.25) and the other 8 U^{xx+}_κ parameters are related to the deviation. It is technically difficult to fit with this parametrization, however, since some of the function shapes are very similar with one another, as shown in the figure E.1.

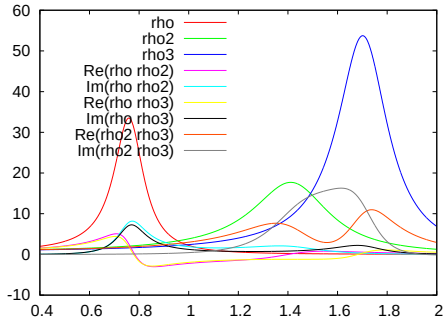


Figure E.1: Function shapes of each quasi-two-body terms.

$$|f_\kappa|^2|A^\kappa|^2 + |\overline{f}_\kappa|^2|\overline{A}^\kappa|^2 = \left(|A^\kappa|^2 + |\overline{A}^\kappa|^2\right)\left[|f_\kappa^{\rm avg}|^2 + p_\kappa^{\rm corr}\right] \ , \tag{E.23}$$

where $p_\kappa^{\rm corr}$ is a correction term written as

$$p_\kappa^{\rm corr} = |T_J^\kappa|^2 \sum_{i=1}^{5} c_\kappa^i p_i^{\rm corr}(s_\kappa) \ , \tag{E.24}$$

$$p_i^{\rm corr}(s) = \begin{cases} 1 & m_i^2 < s < m_{i+1}^2 \\ 0 & \text{otherwise} \end{cases} \ , \tag{E.25}$$

and T_J^κ is the helicity term. This treatment corresponds to model the deviation from the average by a binned histogram PDF. As the binning, or as the definition of m_i, we take the following values

$$m_i \equiv 0.9 + 0.2 \times (i-1) \quad ({\rm GeV}) \ . \tag{E.26}$$

We fit the bins of i = 1-5, corresponding to the region of $0.9 < \sqrt{s} < 1.9({\rm GeV})$, for each of three ρ charges, κ. Thus, the number of the additional coefficients to be fitted, c_κ^i, is 15 in total.

3-2 24 interference related parameters

With the lineshape deviation, the interference terms become

$$2{\rm Re}\left[f_\kappa f_\sigma^* A^\kappa A^{\sigma *} + \overline{f}_\kappa \overline{f}_\sigma^* \overline{A}^\kappa \overline{A}^{\sigma *}\right] \ . \tag{E.27}$$

Neglecting interference between correction terms, this term can be expanded as (here we omit the factor 2 common to all terms)

$$\begin{aligned}
&{\rm Re}\left[f_\kappa f_\sigma^* A^\kappa A^{\sigma *} + \overline{f}_\kappa \overline{f}_\sigma^* \overline{A}^\kappa \overline{A}^{\sigma *}\right] \\
&= {\rm Re}[f_\kappa^{\rm avg} f_\sigma^{\rm avg *}]\, U_{\kappa\sigma}^{+,{\rm Re}} - {\rm Im}[f_\kappa^{\rm avg} f_\sigma^{\rm avg *}]\, U_{\kappa\sigma}^{+,{\rm Im}} \\
&\quad + {\rm Re}[f_\kappa^{\rho'} f_\sigma^{\rm avg *}]\, U_{2\kappa\sigma}^{+,{\rm Re}} - {\rm Im}[f_\kappa^{\rho'} f_\sigma^{\rm avg *}]\, U_{2\kappa\sigma}^{+,{\rm Im}} \\
&\quad + {\rm Re}[f_\kappa^{\rho''} f_\sigma^{\rm avg *}]\, U_{3\kappa\sigma}^{+,{\rm Re}} - {\rm Im}[f_\kappa^{\rho''} f_\sigma^{\rm avg *}]\, U_{3\kappa\sigma}^{+,{\rm Im}} \\
&\quad + {\rm Re}[f_\kappa^{\rm avg} f_\sigma^{\rho' *}]\, U_{\kappa 2\sigma}^{+,{\rm Re}} - {\rm Im}[f_\kappa^{\rm avg} f_\sigma^{\rho' *}]\, U_{\kappa 2\sigma}^{+,{\rm Im}} \\
&\quad + {\rm Re}[f_\kappa^{\rm avg} f_\sigma^{\rho'' *}]\, U_{\kappa 3\sigma}^{+,{\rm Re}} - {\rm Im}[f_\kappa^{\rm avg} f_\sigma^{\rho'' *}]\, U_{\kappa 3\sigma}^{+,{\rm Im}} \ ,
\end{aligned} \tag{E.28}$$

where $U_{\kappa\sigma}^{+,\mathrm{Re(Im)}}$ are the interfering coefficients of Eq. (4.26) and the other 8 parameters are newly introduced coefficients defined as

$$U_{2\kappa\sigma}^{+,\mathrm{Re(Im)}} = \mathrm{Re(Im)}\left[\Delta\beta_\kappa A^\kappa A^{\sigma*} + \Delta\overline{\beta}_\kappa \overline{A}^\kappa \overline{A}^{\sigma*}\right] \tag{E.29}$$

$$U_{3\kappa\sigma}^{+,\mathrm{Re(Im)}} = \mathrm{Re(Im)}\left[\Delta\gamma_\kappa A^\kappa A^{\sigma*} + \Delta\overline{\gamma}_\kappa \overline{A}^\kappa \overline{A}^{\sigma*}\right] \tag{E.30}$$

$$U_{\kappa 2\sigma}^{+,\mathrm{Re(Im)}} = \mathrm{Re(Im)}\left[\Delta\beta_\sigma^* A^\kappa A^{\sigma*} + \Delta\overline{\beta}_\sigma^* \overline{A}^\kappa \overline{A}^{\sigma*}\right] \tag{E.31}$$

$$U_{\kappa 3\sigma}^{+,\mathrm{Re(Im)}} = \mathrm{Re(Im)}\left[\Delta\gamma_\sigma^* A^\kappa A^{\sigma*} + \Delta\overline{\gamma}_\sigma^* \overline{A}^\kappa \overline{A}^{\sigma*}\right] . \tag{E.32}$$

We fit these new 8 $U_{xx}^{+,\mathrm{Re(Im)}}$ parameters in each of 3 interference regions. Consequently, number of new parameters related to interference is 24 in total.

3-3 Fit result

We perform a time-integrated Dalitz Plot fit, where the 39 parameters described above are fitted together with the nominal 8 parameters as nuisance parameters; 47 parameters are fitted simultaneously. Table E.1 shows the fit result of the 39 parameters. Figures E.2 show the mass distributions.

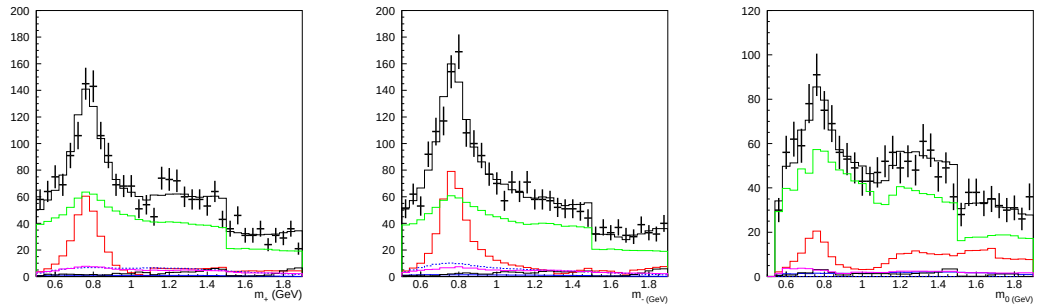


Figure E.2: The mass plots of m_+(left), m_-(middle), and m_0(right), as a result of the 39 parameter fit. The step at $m = 1.5$(GeV) is due to Dalitz veto.

4 Second step: putting constraint on the 10 complex parameters

Using the fitted 39 parameters and assuming T^κ, P^κ, ϕ_2 obtained in our nominal time-dependent Dalitz plot fit, we put a constraint on the 10 complex parameters (model parameters): $\Delta\beta_{+,-,0}^T, \Delta\beta_{+,-}^P, \Delta\gamma_{+,-,0}^T, \Delta\gamma_{+,-}^P$. The method here is a χ^2 fit. The χ^2 is defined as

$$\chi^2 = (\vec{x}_{\mathrm{fit}} - \vec{x}_{\mathrm{model}})^T \mathbf{E}^{-1} (\vec{x}_{\mathrm{fit}} - \vec{x}_{\mathrm{model}}) , \tag{E.33}$$

where $\vec{x}_{\mathrm{fit}}$ and $\mathbf{E}$ are the central values and the error matrix of the fitted 39 parameters. $\vec{x}_{\mathrm{model}}$ is parameters corresponding to the $\vec{x}_{\mathrm{fit}}$, calculated based on the 10 complex valued model parameters, $\Delta\beta_{+,-,0}^T, \Delta\beta_{+,-}^P, \Delta\gamma_{+,-,0}^T, \Delta\gamma_{+,-}^P$.

Among the 39 elements of $\vec{x}_{\mathrm{model}}$, 24 interference related terms are simply calculated using Eqs. (E.7), (E.8), (E.12)-(E.15), (E.17), (E.18), and (E.29)-(E.32). As for the quasi-two-body related 15 parameters, on the other hand, we need to perform numerical calculations to relate

Table E.1: Fit results and errors of the 39 parameters.

$U^{+,Re}_{+2-}$	-0.06 ± 0.35	c^0_+	-1.59 ± 0.82
$U^{+,Re}_{+3-}$	$+0.07 \pm 0.14$	c^1_+	$+0.48 \pm 0.82$
$U^{+,Re}_{2+-}$	-0.04 ± 0.35	c^2_+	$+0.39 \pm 0.61$
$U^{+,Re}_{3+-}$	$+0.26 \pm 0.13$	c^3_+	-0.90 ± 1.09
$U^{+,Im}_{+2-}$	-0.15 ± 0.28	c^4_+	-0.77 ± 0.75
$U^{+,Im}_{+3-}$	$+0.01 \pm 0.12$	c^0_-	$+1.04 \pm 0.98$
$U^{+,Im}_{2+-}$	$+0.86 \pm 0.30$	c^1_-	$+1.01 \pm 0.70$
$U^{+,Im}_{3+-}$	-0.30 ± 0.16	c^2_-	$+0.33 \pm 0.54$
$U^{+,Re}_{+20}$	-0.22 ± 0.25	c^3_-	-0.72 ± 0.78
$U^{+,Re}_{+30}$	$+0.13 \pm 0.11$	c^4_-	$+0.85 \pm 0.81$
$U^{+,Re}_{2+0}$	$+0.38 \pm 0.31$	c^0_0	$+1.44 \pm 2.51$
$U^{+,Re}_{3+0}$	-0.13 ± 0.11	c^1_0	$+0.88 \pm 1.94$
$U^{+,Im}_{+20}$	-0.32 ± 0.21	c^2_0	$+0.61 \pm 1.70$
$U^{+,Im}_{+30}$	$+0.07 \pm 0.10$	c^3_0	$+3.15 \pm 2.98$
$U^{+,Im}_{2+0}$	-0.16 ± 0.22	c^4_0	-4.19 ± 2.14
$U^{+,Im}_{3+0}$	-0.08 ± 0.12		
$U^{+,Re}_{-20}$	-0.34 ± 0.29		
$U^{+,Re}_{-30}$	$+0.04 \pm 0.11$		
$U^{+,Re}_{2-0}$	-0.05 ± 0.35		
$U^{+,Re}_{3-0}$	$+0.00 \pm 0.11$		
$U^{+,Im}_{-20}$	-0.17 ± 0.24		
$U^{+,Im}_{-30}$	-0.13 ± 0.11		
$U^{+,Im}_{2-0}$	-0.14 ± 0.27		
$U^{+,Im}_{3-0}$	$+0.03 \pm 0.12$		

the 10 complex valued model parameters to the measured quantities, c^i_κ. Since they are the differences between $|f_\kappa|^2|A^\kappa|^2 + |\overline{f}_\kappa|^2|\overline{A}^\kappa|^2$ and $(|A^\kappa|^2 + |\overline{A}^\kappa|^2)|f^{\rm avg}_\kappa|^2$ in the mass region of ${m_i}^2 < s_\kappa < {m_{i+1}}^2$, c^i_κ in the $\vec{x}_{\rm model}$ can be related as

$$c^i_\kappa = \frac{1}{{m_{i+1}}^2 - {m_i}^2}\int_{{m_i}^2}^{{m_{i+1}}^2} ds\ \frac{|A^\kappa|^2(|F^\kappa_{\pi\pi}(s)|^2 - |F^{\rm avg}_{\pi\pi}(s)|^2) + |\overline{A}^\kappa|^2(|\overline{F}^\kappa_{\pi\pi}(s)|^2 - |F^{\rm avg}_{\pi\pi}(s)|^2)}{|A^\kappa|^2 + |\overline{A}^\kappa|^2}\ , \tag{E.34}$$

where $F^{xx}_{\pi\pi}$ are defined in the same way as f^{xx}_κ.[3]

By minimizing the χ^2, we obtain the fitted value and 1σ allowed region of the 10 complex valued model parameters. Table E.2 show the fit result.

Table E.2: Fit result of 10 complex (20 real valued) lineshape deviation parameters.

$\Delta\beta^{T,{\rm Re}}_+$	$+0.11 \pm 0.19$	$\Delta\gamma^{T,{\rm Re}}_+$	$+0.09 \pm 0.10$
$\Delta\beta^{T,{\rm Im}}_+$	$+0.27 \pm 0.20$	$\Delta\gamma^{T,{\rm Im}}_+$	-0.07 ± 0.13
$\Delta\beta^{T,{\rm Re}}_-$	$+0.19 \pm 0.30$	$\Delta\gamma^{T,{\rm Re}}_-$	-0.01 ± 0.13
$\Delta\beta^{T,{\rm Im}}_-$	$+0.19 \pm 0.21$	$\Delta\gamma^{T,{\rm Im}}_-$	-0.10 ± 0.18
$\Delta\beta^{T,{\rm Re}}_0$	$+0.11 \pm 0.19$	$\Delta\gamma^{T,{\rm Re}}_0$	$+0.11 \pm 0.09$
$\Delta\beta^{T,{\rm Im}}_0$	$+0.45 \pm 0.16$	$\Delta\gamma^{T,{\rm Im}}_0$	-0.14 ± 0.08
$\Delta\beta^{P,{\rm Re}}_+$	-1.71 ± 1.29	$\Delta\gamma^{P,{\rm Re}}_+$	$+0.36 \pm 0.74$
$\Delta\beta^{P,{\rm Im}}_+$	-0.64 ± 1.51	$\Delta\gamma^{P,{\rm Im}}_+$	$+0.89 \pm 0.85$
$\Delta\beta^{P,{\rm Re}}_-$	$+2.41 \pm 1.00$	$\Delta\gamma^{P,{\rm Re}}_-$	-0.88 ± 0.67
$\Delta\beta^{P,{\rm Im}}_-$	-0.17 ± 1.17	$\Delta\gamma^{P,{\rm Im}}_-$	-0.84 ± 0.51

5 Third step: estimating the systematic errors

With the fitted 10 complex valued (20 real valued) parameters and $1\,\sigma$ errors, we generate random numbers, which correspond to the 10 complex parameters, taking account of the correlation. For each one of the parameter sets, we generate a toy MC sample and fit them; we generate 200 sets of the $\Delta\beta^T_{+,-,0}, \Delta\beta^P_{+,-}, \Delta\gamma^T_{+,-,0}, \Delta\gamma^P_{+,-}$ and we generate 20 k events corresponding to each set. In total, we obtain 200 toy MC samples, with different lineshapes for each sample. Each sample is fitted with the common lineshape assumption, where $f^{\rm avg}_\kappa$ is used as the common lineshape. We then make the residual distributions of the 26 coefficients of the time-dependent Dalitz plot fit.

In the calculation of the residuals, we do not simply calculate the residual using the original input parameters of the toy MC, but we *optimize* the parameters so that the residual get small. In the original formalism, the amplitudes are those of $\rho(770)$. For example, the amplitude of $B^0 \to \rho^+\pi^-$ is

$$A^+ f_+ = A^+(f^\rho_+ + \beta_+ f^{\rho'}_+ + \gamma_+ f^{\rho''}_+)\ . \tag{E.37}$$

[3] Explicitly,

$$F^{\rm avg}_{\pi\pi}(s) \equiv BW_{\rho(770)}(s) + \beta BW_{\rho(1450)}(s) + \gamma BW_{\rho(1700)}(s)\ , \tag{E.35}$$

which is analogous to the equation (E.6), and

$$^{(}\overline{)}F^\kappa_{\pi\pi}(s) \equiv F^{\rm avg}_{\pi\pi}(s) + \Delta^{(}\overline{)}\beta_\kappa BW_{\rho(1450)}(s) + \Delta^{(}\overline{)}\gamma_\kappa BW_{\rho(1700)}(s)\ , \tag{E.36}$$

which is analogous to the equation (E.5).

The underlying meaning of this formalism is that the decay amplitudes of $B^0 \to \rho^+\pi^-$, $B^0 \to \rho'^+\pi^-$, and $B^0 \to \rho''^+\pi^-$ are A^+, $A^+\beta_+$, and $A^+\gamma_+$, respectively. Comparison of the fit result with the input A^+, i.e. to calculate the residual with A^+ as an input, is to see if *the amplitude of $\rho(770)$ is correctly measured or not.* This leads to the overestimation of the residual in some cases, which include ours.

Figure E.3 shows an example for such cases. We suppose the left (right) figure is the lineshape of B^0 ($\overline{B}^0$) $\to \rho^0\pi^0$, and the purpose here is the measurement of direct CP violation in this process. The amplitudes of the decays can be written as

$$A(B^0 \to \rho^0\pi^0) = A^0 f_0 = A^0(f_0^\rho + \beta_0 f_0^{\rho'}) \,, \tag{E.38}$$

$$A(\overline{B}^0 \to \rho^0\pi^0) = \overline{A}^0 \overline{f}_0 = \overline{A}^0(f_0^\rho + \overline{\beta}_0 f_0^{\rho'}) \,, \tag{E.39}$$

where we ignore ρ'' for simplicity. Although the hight of ρ contributions (red line) are exactly the same in two figures (i.e., $|A^0| = |\overline{A}^0|$), what we really observe (green line), which is the sum of ρ and ρ' contributions in this case, is different due to the difference of ρ' contribution. Thus, applying a mass cut of $\sqrt{s} < 1.0$ for example, one will observe a direct CP violation in this situation. One may say, however, that this observation of direct CP violation is a bias due to the lineshape difference between B^0 decay and $\overline{B}^0$ decay, since $|A^0| = |\overline{A}^0|$, and there is no direct CP violation if one pick up the contribution from $\rho(770)$ alone. On the other hand, one can also say that this observation of direct CP violation is correct, since there is certainly difference between B^0 and $\overline{B}^0$ decay amplitudes once the contributions from ρ and ρ' are put together. This difference of two interpretations comes from two different standpoints: 1) to measure the amplitudes of $\rho(770)$ alone or 2) to measure the amplitudes in total without discriminating $\rho(770)$ and its radial excitations.

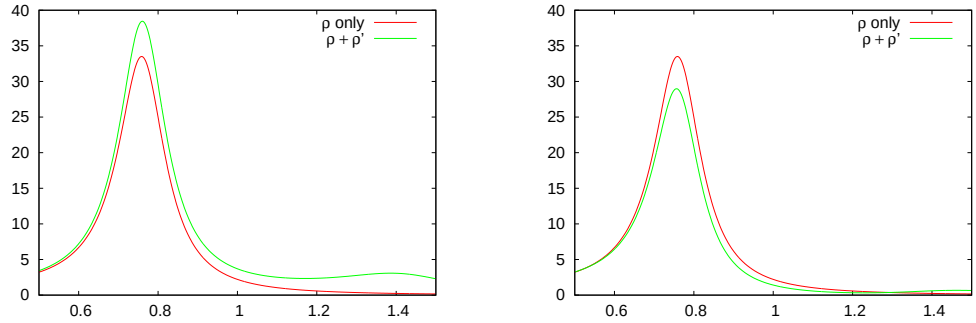


Figure E.3: Lineshapes with constructive interference between ρ and ρ' (left) and destructive interference (right). Here, only the phase of β is different for the two and $|\beta|$ is common.

The question is which standpoint we should take. Our standpoint is similar to the second one. This is because what we like to measure is ϕ_2, not the amplitude of ρ, and the weak phase of the higher resonance decay amplitudes is ϕ_2, being the same as the ρ. That is, we do not need to discriminate the ρ from ρ' and ρ'' since they all are described with ϕ_2 in the same way. In addition, we can expect the Isospin relation, on which we rely in putting constraint on ϕ_2, also holds for ρ' and ρ'' as well as ρ. Thus, in using Isospin relation, we do not have to discriminate ρ, ρ', and ρ''. To summerize, although we use ρ resonance region and there are ρ' and ρ'' contributions in this region, we do not have to distinguish them from ρ contribution.

The next question is how to calculate the residual in MC studies, or how to define the input amplitudes to be used for the residual calculation, when we are on this standpoint. Since we are only interested in ϕ_2, the definition of the other parameters are completely arbitrary. In

other words, two sets of the amplitudes are looked upon to be identical as long as they yield the same ϕ_2. Thus, as the amplitudes to be compared with the MC fit result to calculate the residual, we can choose ${}^{(}\overline{A}{}^{)\prime\kappa}$ that differs from amplitudes of $\rho(770)$, ${}^{(}\overline{A}{}^{)k}$, and close to the amplitudes of $\rho + \rho' + \rho''$ yielding the same ϕ_2 as ${}^{(}\overline{A}{}^{)k}$. To calculate the optimum ${}^{(}\overline{A}{}^{)\prime\kappa}$, we optimize the amplitudes T' and P' that parameterize ${}^{(}\overline{A}{}^{)\prime\kappa}$ as

$$A'^{\kappa} = e^{-i\phi_2}\, T'^{\kappa} + P'^{\kappa} \,, \tag{E.40}$$

$$\overline{A}'^{\kappa} = e^{+i\phi_2}\, T'^{\overline{\kappa}} + P'^{\overline{\kappa}} \,, \tag{E.41}$$

where ϕ_2 is the same as that of ${}^{(}\overline{A}{}^{)k}$. Here, we optimize the T' and P' so that the residual becomes minimum.[4] In this optimization, we minimize χ^2 defined as

$$\chi^2 \equiv \sum_i \frac{(x^i_{\rm fit} - x^i_{\rm model})^2}{{\sigma_i}^2} \,, \tag{E.43}$$

where $x^i_{\rm fit}$ is the 26 parameters fitted with MC, $x^i_{\rm model}$ is the 26 parameters calculated from T', P', and ϕ_2. As the σ_i, we use the statistical errors of our nominal time-dependent Dalitz plot fit.

The resultant amplitude is used for the residual calculation; the $x^i_{\rm fit} - x^i_{\rm model}$ in the fit result is defined to be the residual for each MC sample. The widths (root mean square) and the biases of the residual distributions are used as the systematic errors from the lineshape variation. Here, the number of events in each MC sample, 20 k, is large enough and the statistical fluctuation is negligible.

[4] Since we rely on Isospin relation in the ϕ_2 constraint, the P' have to satisfy the Isospin relation

$$P'^0 = -\frac{1}{2}(P'^+ + P'^-) \,. \tag{E.42}$$

This reduce the degree of freedom to be 10 real valued parameters. Further, since global normalization and phase are arbitrary, the number of parameters to be optimized is 8 in total.

Appendix F

Unbinned Extended Maximum Likelihood Fit

1 Formalism

In an usual unbinned maximum likelihood fit, the likelihood function is defined as

$$\mathcal{L}(\vec{p}) = \prod_i^N P(\vec{x}_i; \vec{p}) \ , \tag{F.1}$$

where $\vec{x}$, $\vec{p}$, i, and N are the set of event-by-event variables, parameters to be fitted, an index over the events, and the number of events, respectively. Here, $P(\vec{x}; \vec{p})$ is an event-by-event PDF normalized to be unity. In the unbinned extended maximum likelihood fit, we add another term to incorporate the Poisson distributed property of the number of events N:

$$\mathcal{L}(\vec{p}) = \frac{\nu(\vec{p})^N \, e^{-\nu(\vec{p})}}{N!} \prod_i^N P(\vec{x}; \vec{p}) \ , \tag{F.2}$$

where $\nu(\vec{p})$ is the estimated, or fitted, number of events.

2 Likelihood Function for Time-integrated ΔE-$M_{\rm bc}$ and Dalitz Plot Fit

As described in Sec. 3, the event-by-event PDF for the time-integrated ΔE-$M_{\rm bc}$ and Dalitz plot fit is different for the events in the signal region and the sideband region. Thus, we define the extended likelihood function separately for signal region ($\mathcal{L}_{\rm SR}$) and sideband region ($\mathcal{L}_{\rm SB}$) as

$$\mathcal{L}_{\rm SR} = \frac{{\nu_{\rm SR}}^{N_{\rm SR}} \, e^{-\nu_{\rm SR}}}{N_{\rm SR}!} \prod_i^{N_{\rm SR}} P_{\cancel{\Delta t}}^{\rm SR}(\vec{x}_i^{\rm SR}) \ , \tag{F.3}$$

$$\mathcal{L}_{\rm SB} = \frac{{\nu_{\rm SB}}^{N_{\rm SB}} \, e^{-\nu_{\rm SB}}}{N_{\rm SB}!} \prod_i^{N_{\rm SB}} \frac{P_{\cancel{\Delta t}}^{\rm SB}(\vec{x}_i^{\rm SB})}{\mathcal{N} - 1} \ , \tag{F.4}$$

with the event-by-event variables defined as

$$\vec{x}^{\rm SR} \equiv (\Delta E, M_{\rm bc}; m', \theta'; l) \ , \tag{F.5}$$

$$\vec{x}^{\mathrm{SB}} \equiv (\Delta E, M_{\mathrm{bc}}; l) \ . \tag{F.6}$$

Here, ν_{SR} and N_{SR} (ν_{SB} and N_{SB}) are the observed and estimated number of events in the signal (sideband) region, respectively; and i is an index over the events. While $P^{\mathrm{SR}}_{\not{\Delta t}}$ is normalized to be unity in the signal region by definition of $P(\Delta E, M_{\mathrm{bc}}; m', \theta'; \Delta t, q_{\mathrm{tag}}; l)$, $P^{\mathrm{SB}}_{\not{\Delta t}}$ is not normalized and thus divided by the normalization factor $\mathcal{N} - 1$, where $\mathcal{N}$ is the integration over the grand signal region defined as

$$\mathcal{N} \equiv \sum_{l} \iint_{\mathrm{GS}} d\Delta E \, dM_{\mathrm{bc}} \, P^{\mathrm{SB}}_{\not{\Delta t}}(\Delta E, M_{\mathrm{bc}}; l) \ . \tag{F.7}$$

Since the numbers of estimated events in the signal region and sideband region are to be proportional to the integral of the PDF in the corresponding regions, following relation is satisfied

$$\frac{\nu_{\mathrm{SR}}}{\nu_{\mathrm{SB}}} = \frac{\displaystyle\sum_{l} \iint_{\mathrm{SR}} d\Delta E \, dM_{\mathrm{bc}} \iint dm' d\theta' \, P^{\mathrm{SR}}_{\not{\Delta t}}(\Delta E, M_{\mathrm{bc}}; m', \theta'; l)}{\displaystyle\sum_{l} \iint_{\mathrm{SB}} d\Delta E \, dM_{\mathrm{bc}} \, P^{\mathrm{SB}}_{\not{\Delta t}}(\Delta E, M_{\mathrm{bc}}; l)} = \frac{1}{\mathcal{N} - 1} \ . \tag{F.8}$$

Consequently, the degree of freedom introduced in the "extended" part is only one, i.e., we set ν_{SR} to be a free parameter in the fit and calculate ν_{SB} by the relation of equation (F.8).

The total likelihood $\mathcal{L}$ is a product of the likelihood functions in the signal region and sideband region:

$$\mathcal{L} = \mathcal{L}_{\mathrm{SR}} \cdot \mathcal{L}_{\mathrm{SB}} \ . \tag{F.9}$$

This can be considered as a simultaneous fit of two likelihood fits, with some fit parameters in common.

Appendix G

Putting Limits on the Other $B^0 \to \pi^+\pi^-\pi^0$ Contributions

In this appendix, we describe how we put the upper limits on the contributions from $B^0 \to \pi^+\pi^-\pi^0$ decay processes other than $B^0 \to (\rho\pi)^0$. We assume contributions from $B^0 \to f_0(980)\pi^0$, $B^0 \to f_0(600)\pi^0$, $B^0 \to \omega\pi^0$, and $B^0 \to \pi^+\pi^-\pi^0$ non-resonant. As for $B^0 \to f_0(980)\pi^0$, $B^0 \to f_0(600)\pi^0$, and $B^0 \to \pi^+\pi^-\pi^0$ non-resonant, we put the limits using our data, since there is no measurement or the limits available as world average are poor. The upper limit for the $B^0 \to \omega\pi^0$ is obtained using the world average.

1 Scalar particles: $B^0 \to f_0(980)\pi^0$ and $B^0 \to f_0(600)\pi^0$

We perform the Dalitz plot fit to constrain the contributions from these decay channels. Assuming contributions from the $B^0 \to S\pi^0$, where S is $f_0(980)$ or $f_0(600)$, the amplitudes of Eq. (2.158) and (2.159) are rewritten as

$$A_{3\pi} = f_+A^+ + f_-A^- + f_0A^0 + f_SA^S \ , \quad \text{and} \tag{G.1}$$

$$\frac{q}{p}\overline{A}_{3\pi} = f_+\overline{A}^+ + f_-\overline{A}^- + f_0\overline{A}^0 + f_S\overline{A}^S \ , \tag{G.2}$$

Here A^S ($\overline{A}^S$) and f_S are the complex amplitude and the kinematical term of S, respectively. The f_S consists of the helicity distribution and the Breit-Wigner mass distribution as

$$f_S = T^S_{J=0}F^0_{\pi\pi,S} \ , \tag{G.3}$$

where

$$T^S_{J=0} = 1 \ , \tag{G.4}$$

and

$$F^0_{\pi\pi,S} = BW_S(s_0) \ . \tag{G.5}$$

Here we adopt a relativistic Brit-Wigner of $J = 0$ for the mass distribution. The time-integrated part of the decay width is

$$\begin{aligned} |A_{3\pi}|^2 + |\overline{A}_{3\pi}|^2 \quad = \quad & \sum_{\kappa\in\{+,-,0\}} |f_\kappa|^2U^+_\kappa + 2\sum_{\kappa<\sigma\in\{+,-,0\}} \left(\mathrm{Re}[f_\kappa f^*_\sigma]U^{+,\mathrm{Re}}_{\kappa\sigma} - \mathrm{Im}[f_\kappa f^*_\sigma]U^{+,\mathrm{Im}}_{\kappa\sigma}\right) \\ & + |f_S|^2U^+_S + 2\sum_{\kappa\in\{+,-,0\}} \left(\mathrm{Re}[f_S f^*_\kappa]U^{+,\mathrm{Re}}_{S\kappa} - \mathrm{Im}[f_S f^*_\kappa]U^{+,\mathrm{Im}}_{S\kappa}\right) \ , \end{aligned} \tag{G.6}$$

with 7 additional U^+_{Sx} parameters in the second line. Note that this is a general expression without any assumption on the CP violation of $B^0 \to S\pi^0$ and the relative phase between $B^0 \to S\pi^0$ and $B^0 \to (\rho\pi)^0$. Among them, U^+_S is the parameter that corresponds to the contribution (decay width) of $B^0 \to S\pi^0$ and other 6 parameters are nuisance parameters here. Fit to the data with this PDF yields the results of the tables G.1 and G.2.

Table G.1: Fit result for $B^0 \to f_0(980)\pi^0$.

Resonance parameters				
Mass (GeV)	Width(GeV)	Reference	U^+_S	$\mathcal{B}(f_0(980)\pi^0)/\mathcal{B}(\rho^0\pi^0)$
0.976	0.061	Belle [62]	$+0.5^{+1.1}_{-1.0}$	$0.03^{+0.08}_{-0.07}$

Table G.2: Fit result for $B^0 \to \sigma\pi^0$.

Resonance parameters				
Mass (GeV)	Width(GeV)	Reference	U^+_S	$Br(\sigma\pi^0)/Br(\rho^0\pi^0)$
0.541	0.504	BES2 [63]	$(+2.2^{+2.5}_{-2.3}) \times 10^2$	$+0.16^{+0.18}_{-0.16}$
0.513	0.335	CLEO [64]	$(+0.5^{+1.7}_{-1.5}) \times 10^2$	$+0.04^{+0.13}_{-0.12}$
0.478	0.324	E791 [65]	$(+0.7^{+2.4}_{-2.1}) \times 10^2$	$+0.04^{+0.13}_{-0.12}$

2 Vector particle: $B^0 \to \omega\pi^0$

In HFAG 2006, 90% C.L. upper limit of $\mathcal{B}(B^0 \to \omega\pi^0)$ is 1.2×10^{-6}. In PDG, the $Br(\omega \to \pi^+\pi^-) = 1.70 \pm 0.27\%$. Thus, 90% C.L. upper limit of $\mathcal{B}(B^0 \to \omega\pi^0, \omega \to \pi^+\pi^-)$ is $\sim 0.03 \times 10^{-6}$.

3 Phasespace: $B^0 \to \pi^+\pi^-\pi^0$ non-resonant

To avoid the $B^0 \to (\rho\pi)^0$ resonant contribution, We apply the following Dalitz plot cuts:

Region I (low purity region)

$$1.5 < \sqrt{s_+} \quad \text{and} \quad 1.5 < \sqrt{s_-} \quad \text{and} \quad 1.5 < \sqrt{s_0} \tag{G.7}$$

and

$$\sqrt{s_+} < 2.0 \quad \text{or} \quad \sqrt{s_-} < 2.0 \quad \text{or} \quad \sqrt{s_0} < 2.0 \tag{G.8}$$

Region II (high purity region)

$$2.0 < \sqrt{s_+} \quad \text{and} \quad 2.0 < \sqrt{s_-} \quad \text{and} \quad 2.0 < \sqrt{s_0}\,. \tag{G.9}$$

We also veto the following region in common for the "Region I" and "Region II"

$$0.44 < m' < 0.455\,, \tag{G.10}$$

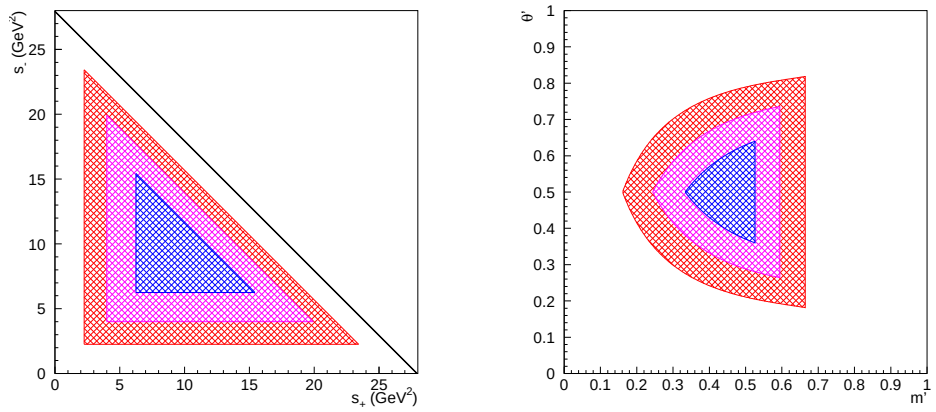


Figure G.1: Regions used for the non-resonant measurement in normal Dalitz Plot (left) and Square Dalitz Plot (right). "Region I" and "Region II" correspond to purple and blue regions, respectively. Red region is dominated by $B^0 \to DX$ and white region is dominated by resonant component; We do not use these regions.

to avoid a contribution from $B^0 \to J/\psi\pi^0$. Except for the Daltiz Plot cut, the event selection almost the same as the $B^0 \to (\rho\pi)^0$.[1] We perform ΔE-M_{bc} fits to the data and calculate the branching fraction assuming all of the $B^0 \to \pi^+\pi^-\pi^0$ are the contributions of non-resonant.

Fit results are shown Fig. G.2. Using the calculated efficiencies in Table G.3, we obtain the branching fractions in Table G.4. By simply symmetrizing the errors and averaging all the measurements, we obtain following value:

$$-0.15 \pm 0.62 \quad (\times 10^{-6}) \,. \tag{G.11}$$

Data Set	Region	Efficiency (%)
DS-I	Region I	2.21
DS-I	Region II	1.00
DS-II	Region I	2.32
DS-II	Region II	1.06

Table G.3: Efficiencies.

[1] Another difference, in addition to the cut in Dalitz plot, is best candidate selection. Since the region does not contain the kinematic region with low momentum π, the is virtually no contribution from SCF. Thus, we do not apply best candidate selection.

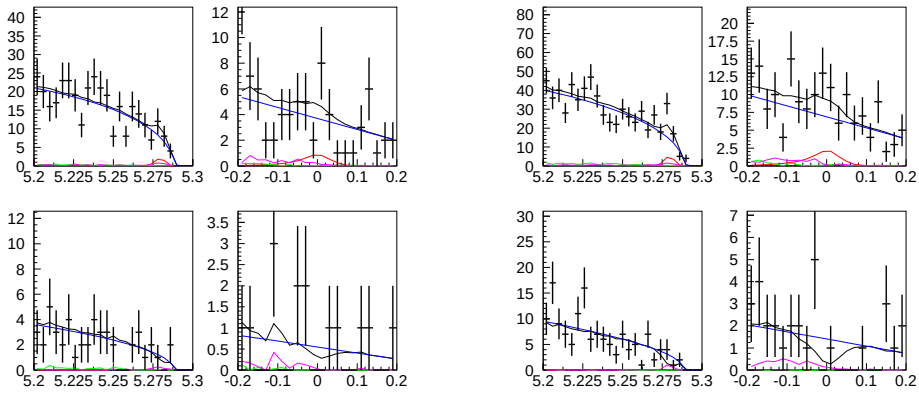


Figure G.2: ΔE-M_{bc} fit results of SVD1 (left) and SVD2 (right). Top and bottom plots correspond to the "Region I" and "Region II", respectively.

Data Set	Region	B.F. ($\times 10^{-6}$)
DS-I	Region I	$+1.11^{+1.38}_{-1.13}$
DS-I	Region II	$-0.76^{+1.31}_{-0.72}$
DS-II	Region I	$+1.41^{+1.20}_{-0.99}$
DS-II	Region II	$-1.89^{+1.03}_{-1.12}$

Table G.4: Branching fractions obtained by each fit.

Appendix H

Techniques to Estimate Systematic Errors

1 Statistical Uncertainty of Binned Histogram PDF

To estimate the statistical uncertainty of a binned histogram PDF, we use two techniques: a method using toy MC study and that using random division of the data sample. The former is more precise but more complex than the latter. We use the former for the Dalitz plot PDF of continuum events, since its systematic error is sizable. The latter is used for the Dalitz dependent signal efficiency and the ΔE-$M_{\rm bc}$ and Dalitz PDF for $B\overline{B}$ background.

Note that the systematic error due to the binning is taken into account as the fit bias. If there were a significant effect due to the binning, we would have seen significant fit bias in the MC fit; we do not find large bias and the small fit bias is included into the systematic error.

1-1 Method Using Toy MC

We model the (scaled) Dalitz PDF of continuum background $\mathcal{P}_{q\overline{q}}(m'_{\rm scale}, \theta')$ by a binned histogram using the events in the sideband region, as described in Sec. 3-2 of appendix A. Since the number of the sideband events is limited, the $\mathcal{P}_{q\overline{q}}(m'_{\rm scale}, \theta')$ has uncertainty due to the limited statistics. In this appendix, we describe the procedure to estimate the systematic error due to the uncertainty of $\mathcal{P}_{q\overline{q}}(m'_{\rm scale}, \theta')$.

In this method, we use toy MC corresponding to the sideband events. First, we generate 200 samples of pseudo experiments corresponding to the sideband data. Each pseudo experiment consists of 13161 events corresponding to DS-I and 25813 events corresponding DS-II, where the numbers of events are the same as data. Secondly, we prepare the Dalitz PDF of the continuum background using each of the pseudo experiment samples in the same way as we do for data; we obtain 200's of statistically fluctuated PDF's. Finally, we fit the data with the same condition as the nominal fit except for the continuum Dalitz PDF, for which we use a PDF based on the pseudo experiment instead of the nominal PDF based on data sideband. By fitting the data using the 200 PDF's we obtain the variations of the fitted parameters due to the statistical fluctuation of the sideband events. We quote these variations as the systematic errors due to the statistical fluctuation of the sideband events.

1-2 Method Using Random Division

In this method, we randomly divide the MC sample used to model the PDF into two subsamples. Using the subsamples, we make two PDF's in the same way as the nominal construction.

As shown in the following, the difference of the two fit results using these two PDF's is expected to distribute such that its width is twice of the systematic uncertainty to be estimated. We repeat the random division and fits, and make the distribution of the difference of the two fits. We assign the width of the distribution as the systematic error[1].

Provided the number of entries of the i-th bin of the histogram is N_i, its uncertainty is expected to be $\sqrt{N_i}$. When the number of entries of the i-th bin is N_i^1 in one of the subsamples, another subsample has $N_i^2 = N_i - N_i^1$ entries in the same bin. Assuming binomial distribution, the variance of the N_i^1 is

$$\left\langle (N_i^1 - N_i/2)^2 \right\rangle = \frac{N_i}{2} \, . \tag{H.1}$$

Thus, the variance of $N_i^1 - N_i^2$ $(= 2N_i^1 - N_i)$ is

$$\left\langle (N_i^1 - N_i^2)^2 \right\rangle = 2N_i \, . \tag{H.2}$$

Since the difference of each bin of the histogram PDF behaves as above, the difference of the fit result does in the same manner.

2 Systematic Error Sourced by Parameters with Large Correlation

There are parameters that have uncertainties and are fixed in the fit. To estimate systematic error due to them, we usually vary the parameters one-by-one by their $1\,\sigma$ uncertainties (Fig. H.1 left) and quote the difference of the fit result from the nominal one as the systematic error[2]. When the uncertainties of several parameters have sizable uncertainties, however, the procedure described above can be either an over or under estimation. The systematic error from the nominal lineshape parameters, β and γ, is the case, where four degrees of freedom $(|\beta|, \arg\beta, |\gamma|, \arg\gamma)$ are highly correlated with each other.

In such cases, we have to treat the systematic errors properly, as shown in Fig. H.1 right. With the variation of the systematic source parameters $\delta\vec{v}$ from the nominal value, e.g., $\delta\vec{v} = (\delta|\beta|, \delta\arg\beta, \delta|\gamma|, \delta\arg\gamma)^T$, and the error matrix for their uncertainties $\boldsymbol{E}$, the corresponding χ^2 is written as

$$\chi^2 = \delta\vec{v}^T \boldsymbol{E}^{-1} \delta\vec{v} \, . \tag{H.3}$$

Since $\boldsymbol{E}^{-1}$ is a real-valued symmetric matrix, it can be diagonalized with an orthogonal matrix $\boldsymbol{U}$. Note that the matrix $\boldsymbol{U}$ is uniquely determined except for the order of columns since the eigenvalues are non-zero and have no degeneration in general. Each column of the $\boldsymbol{U}$ is denoted by $\vec{e}_i$, where i is the index over the columns, and the eigenvalue of $\boldsymbol{E}^{-1}$ corresponding to $\vec{e}_i$ is λ_i. The systematic error is estimated by substituting $\delta\vec{v} = \pm\vec{e}_i/\sqrt{\lambda_i}$ for each i, which satisfies $\chi^2 = 1$. The fit result is compared with the nominal fit corresponding to $\delta\vec{v} = \vec{0}$ and the difference is quoted as a systematic error.

[1] This width corresponds to the twice of the expected uncertainty. As a convention, we assign twice of the expected uncertainty when it is sourced by MC statistics.

[2] Another standard method is to generate toy MC with varied parameters, fit the MC with the nominal parameters, and quote the bias of the fit result as systematic error. The treatment of the correlation described here can be applied for this method, too.

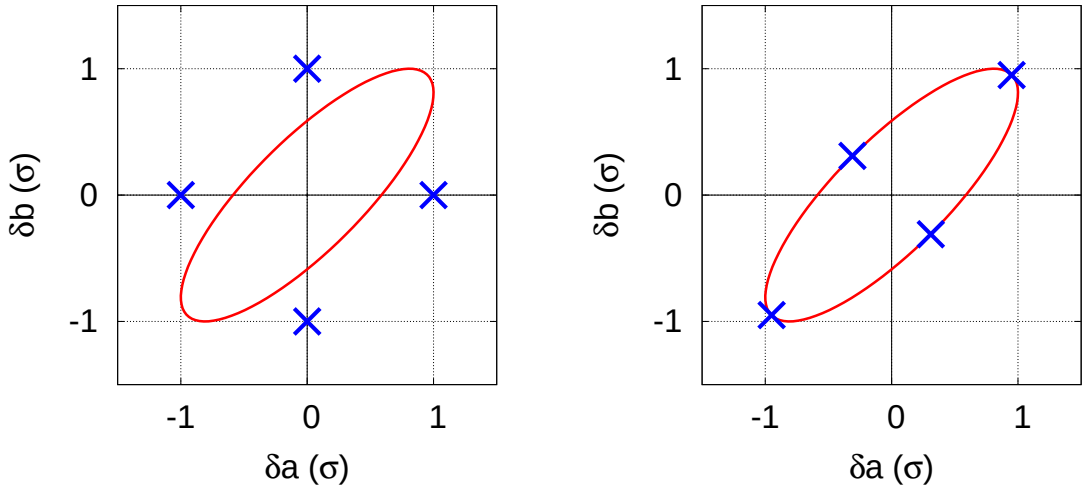


Figure H.1: Schematic figures of the correlation of the parameters a and b, which have uncertainties but fixed in the nominal fit. The ellipses correspond to the $1\,\sigma$ contour of the uncertainty of a and b; they are not circles but ellipses since a and b are correlated with each other. In the estimation of the systematic error, we usually vary a and b by $1\,\sigma$ separately, corresponding to the cross points in the left plot. To treat the correlation properly, however, we should estimate the systematic error using the cross points in the right plot.

Bibliography

[1] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2π Decay of the ${K_2}^0$ meson," *Phys. Rev. Lett.* **13** (1964) 138–140.

[2] N. Cabibbo, "Unitary Symmetry and Leptonic Decays," *Phys. Rev. Lett.* **10** (1963) 531–532.

[3] M. Kobayashi and T. Maskawa, "CP violation in the renormalizable theory of weak interaction," *Prog. Theor. Phys.* **49** (1973) 652–657.

[4] S. W. Herb *et al.*, "Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions," *Phys. Rev. Lett.* **39** (1977) 252–255.

[5] **CDF Collaboration**, F. Abe *et al.*, "Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab," *Phys. Rev. Lett.* **74** (1995) 2626–2631, `hep-ex/9503002`.

[6] **D0 Collaboration**, S. Abachi *et al.*, "Observation of the Top Quark," *Phys. Rev. Lett.* **74** (1995) 2632–2637, `hep-ex/9503003`.

[7] **KTeV Collaboration**, A. Alavi-Harati *et al.*, "Observation of Direct CP Violation in $K_{S,L} \to \pi\pi$ Decays," *Phys. Rev. Lett.* **83** (1999) 22, `hep-ex/9905060`.

[8] **NA48 Collaboration**, V. Fanti *et al.*, "A new measurement of direct CP violation in two pion decays of the neutral kaon," *Phys. Lett.* **B465** (1999) 335–348, `hep-ex/9909022`.

[9] **Belle Collaboration**, K. Abe *et al.*, "Observation of Large CP Violation in the Neutral B Meson System," *Phys. Rev. Lett.* **87** (2001) 091802, `hep-ex/0107061`.

[10] **BABAR Collaboration**, B. Aubert *et al.*, "Observation of CP Violation in the B^0 Meson System," *Phys. Rev. Lett.* **87** (2001) 091801, `hep-ex/0107013`.

[11] **CKMfitter Group**, "Results as of summer 2006," 2006. `http://www.slac.stanford.edu/xorg/ckmfitter/ckm_results_beauty2006.html`.

[12] A. B. Carter and A. I. Sanda, "CP Violation in Cascade Decays of B Mesons," *Phys. Rev. Lett.* **45** (1980) 952.

[13] A. B. Carter and A. I. Sanda, "CP Violation in B Meson Decays," *Phys. Rev. D* **23** (1981) 1567.

[14] I. I. Y. Bigi and A. I. Sanda, "Notes on the Observability of CP Violations in B Decays," *Nucl. Phys. B* **193** (1981) 85.

[15] **CDF Collaboration**, A. Abulencia *et al.*, "Observation of B_s^0-$\bar{B}_s^0$ Oscillations," `hep-ex/0609040`.

[16] M. Gronau and D. London, "Isospin analysis of CP asymmetries in B decays," *Phys. Rev. Lett.* **65** (1990) 3381–3384.

[17] H. J. Lipkin, Y. Nir, H. R. Quinn, and A. Snyder, "Penguin trapping with isospin analysis and CP asymmetries in B decays," *Phys. Rev.* **D44** (1991) 1454–1460.

[18] M. Gronau, "Elimination of penguin contributions to CP asymmetries in B decays through isospin analysis," *Phys. Lett.* **B265** (1991) 389–394.

[19] A. E. Snyder and H. R. Quinn, "Measuring CP asymmetry in $B \to \rho\pi$ decays without ambiguities," *Phys. Rev.* **D48** (1993) 2139–2144.

[20] S. Kurokawa and E. Kikutani, "Overview of the KEKB accelerators," *Nucl. Instrum. Meth.* **A499** (2003) 1–7. and other papers included in this volume.

[21] **Belle Collaboration**, A. Abashian *et al.*, "The Belle detector," *Nucl. Instrum. Meth.* **A479** (2002) 117–232.

[22] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa Matrix," *Phys. Rev. Lett.* **51** (1983) 1945.

[23] Throughout this thesis, B^0 and B^+ denotes B_d and B_u, respectively, unless otherwise stated.

[24] I. I. Y. Bigi and A. I. Sanda, *CP Violation*. No. 9 in Cambridge monographs on particle physics, nuclear physics, and cosmology. Cambridge University Press, 2000.

[25] J. Charles, *Désintégration des mésons B*. PhD thesis, 1999. `http://tel.archives-ouvertes.fr/documents/archives0/00/00/25/02/`.

[26] M. Gronau, O. F. Hernandez, D. London, and J. L. Rosner, "Electroweak penguins and two-body B decays," *Phys. Rev.* **D52** (1995) 6374–6382, `hep-ph/9504327`.

[27] **Belle Collaboration**, G. Alimonti *et al.*, "The BELLE silicon vertex detector," *Nucl. Instrum. Meth.* **A453** (2000) 71–77.

[28] E. Nygard, P. Aspell, P. Jarron, P. Weilhammer, and K. Yoshioka, "CMOS low noise amplifier for microstrip readout: Design and results," *Nucl. Instrum. Meth.* **A301** (1991) 506–516.

[29] O. Toker, S. Masciocchi, E. Nygard, A. Rudge, and P. Weilhammer, "VIKING: A CMOS low noise monolithic 128-channel front end for Si strip detector readout," *Nucl. Instrum. Meth.* **A340** (1994) 572–579.

[30] M. Yokoyama *et al.*, "Radiation-hardness of VA1 with sub-micron process technology," *IEEE Trans. Nucl. Sci.* **48** (2001) 440–443.

[31] Y. Ushiroda, "Belle silicon vertex detectors," *Nucl. Instrum. Meth.* **A511** (2003) 6–10.

[32] H. Hirano *et al.*, "A high resolution cylindrical drift chamber for the KEK B-factory," *Nucl. Instrum. Meth.* **A455** (2000) 294–304.

[33] M. Akatsu *et al.*, "Cathode image readout in the BELLE central drift chamber," *Nucl. Instrum. Meth.* **A454** (2000) 322–333.

[34] T. Iijima *et al.*, "Aerogel Cherenkov counter for the BELLE detector," *Nucl. Instrum. Meth.* **A453** (2000) 321–325.

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

[36] H. Kichimi *et al.*, "The BELLE TOF system," *Nucl. Instrum. Meth.* **A453** (2000) 315–320.

[37] H. Ikeda *et al.*, "A detailed test of the CsI(Tl) calorimeter for BELLE with photon beams of energy between 20-MeV and 5.4-GeV," *Nucl. Instrum. Meth.* **A441** (2000) 401–426.

[38] **Belle Collaboration**, A. Abashian *et al.*, "The K_L/μ detector subsystem for the BELLE experiment at the KEK B-factory," *Nucl. Instrum. Meth.* **A449** (2000) 112–124.

[39] R. Akhmetshin *et al.*, "Survey of the properties of BGO crystals for the extreme forward calorimeter at BELLE," *Nucl. Instrum. Meth.* **A455** (2000) 324–328.

[40] Y. Makida *et al.*, Adv. Cryog. Eng. **37**, 401 (1991); **A43**, 221 (1998); H. Mukai *et al.*, in *Proceedings of the 15th International Conference on Magnet Technology (MT-15), Beijing, 1997*, edited by L. Lin, G. Shen, and L. Yan (Science Press, Beijing, 1998), p. 1044.

[41] Y. Ushiroda *et al.*, "Development of the central trigger system for the BELLE detector at the KEK B-factory," *Nucl. Instrum. Meth.* **A438** (1999) 460–471.

[42] **Belle KID Group**, H. Hamasaki *et al.*, "Kaon Identification in Belle." Belle Note 321 (unpublished).

[43] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, "Electron identification in Belle," *Nucl. Instrum. Meth.* **A485** (2002) 490–503, `hep-ex/0108044`.

[44] A. Abashian *et al.*, "Muon identification in the Belle experiment at KEKB," *Nucl. Instrum. Meth.* **A491** (2002) 69–82.

[45] H. Kakuno *et al.*, "Neutral B flavor tagging for the measurement of mixing-induced CP violation at Belle," *Nucl. Instrum. Meth.* **A533** (2004) 516–531, `hep-ex/0403022`.

[46] H. Tajima *et al.*, "Proper-time resolution function for measurement of time evolution of B mesons at the KEK B-factory," *Nucl. Instrum. Meth.* **A533** (2004) 370–386, `hep-ex/0301026`.

[47] **BABAR Collaboration**, B. Aubert *et al.*, "Measurement of CP-violating asymmetries in $B^0 \to (\rho\pi)^0$ using a time-dependent Dalitz plot analysis," `hep-ex/0408099`.

[48] **BaBar Collaboration**, B. Aubert *et al.*, "An amplitude analysis of the decay $B^\pm \to \pi^\pm\pi^\pm\pi^\mp$," *Phys. Rev.* **D72** (2005) 052002.

[49] H. R. Quinn and J. P. Silva, "The use of early data on b –¿ rho pi decays," *Phys. Rev. D* **62** (2000) 054002, `hep-ph/0001290`.

[50] N. Abe, *Measurement of a CP violation parameter ϕ_2 through B meson decay into $\pi^+\pi^-$*. PhD thesis, 2003.

[51] G. J. Gounaris and J. J. Sakurai, "Finite width corrections to the vector meson dominance prediction for $\rho \to e^+e^-$," *Phys. Rev. Lett.* **21** (1968) 244.

[52] **Particle Data Group**, S. Eidelman *et al.*, "Review of particle physics," *Phys. Lett.* **B592** (2004) 1. and 2005 partial update (http://pdg.lbl.gov).

[53] A. Kusaka, "Study of π^0 efficiency difference between MC and real data depending on the π^0 momentum." Belle Note 894 (unpublished).

[54] **Heavy Flavor Averaging Group (HFAG)**, "Averages of b-hadron properties at the end of 2005," `hep-ex/0603003`.

[55] **Heavy Flavor Averaging Group (HFAG)**, "Online update of Summer 2006," 2006. `http://www.slac.stanford.edu/xorg/hfag`.

[56] O. Long, M. Baak, R. N. Cahn, and D. Kirkby, "Impact of tag-side interference on time dependent CP asymmetry measurements using coherent $B^0\overline{B}^0$ pairs," *Phys. Rev.* **D68** (2003) 034010, `hep-ex/0303030`.

[57] **Belle Collaboration**, J. Dragic *et al.*, "Measurement of branching fraction and direct CP asymmetry in $B^0 \to \rho^0\pi^0$ decays," *Phys. Rev.* **D73** (2006) 111105.

[58] **BABAR Collaboration**, B. Aubert *et al.*, "Measurement of branching fractions and charge asymmetries in $B^\pm \to \rho^\pm\pi^0$ and $B^\pm \to \rho^0\pi^\pm$ decays, and search for $B^0 \to \rho^0\pi^0$," *Phys. Rev. Lett.* **93** (2004) 051802, `hep-ex/0311049`.

[59] **BABAR**, B. Aubert *et al.*, "Measurement of cp-violating asymmetries in $B^0 \to (\rho\pi)^0$ using a time-dependent dalitz plot analysis," `hep-ex/0608002`.

[60] **CKMfitter Group**, J. Charles *et al.*, "CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories," *Eur. Phys. J.* **C41** (2005) 1–131, `hep-ph/0406184`.

[61] **ARGUS Collaboration**, H. Albrecht *et al.*, "Search for Hadronic $b \to u$ Decays," *Phys. Lett.* **B241** (1990) 278.

[62] **Belle Collaboration**, A. Garmash *et al.*, "Dalitz analysis of the three-body charmless decays $B^+ \to K^+\pi^+\pi^-$ and $B^+ \to K^+K^+K^-$," *Phys. Rev.* **D71** (2005) 092003, `hep-ex/0412066`.

[63] **BES Collaboration**, M. Ablikim *et al.*, "The sigma pole in $J/\psi \to \omega\pi^+\pi^-$," *Phys. Lett.* **B598** (2004) 149–158, `hep-ex/0406038`.

[64] **CLEO Collaboration**, H. Muramatsu *et al.*, "Dalitz analysis of $D^0 \to K^0_S\pi^+\pi^-$," *Phys. Rev. Lett.* **89** (2002) 251802, `hep-ex/0207067`.

[65] **E791 Collaboration**, E. M. Aitala *et al.*, "Experimental evidence for a light and broad scalar resonance in $D^+ \to \pi^-\pi^+\pi^+$ decay," *Phys. Rev. Lett.* **86** (2001) 770–774, `hep-ex/0007028`.

List of Notations

Symbol	Defined in	Description
$\boldsymbol{V}$	Page 11	CKM matrix
λ_C	Page 12	Cabbibo angle
A, ρ, η	Page 12	Parameters of Wolfenstein parameterization
ϕ_1, ϕ_2, ϕ_3	Page 13	Angles of CKM triangle
p, q	Page 15	Coefficients that define the mass eigenstates of neutral B
Γ	Page 17	The averaged decay width of the neutral B meson
Δm_d	Page 18	The mass difference between the two mass eigenstates of the neutral B meson
Δt	Page 20	Proper time difference between the decays of two B mesons produced in pair
$B_{\rm tag}$	Page 20	One of the B's that is used to tag the B^0 flavor
$q_{\rm tag}$	Page 20	Flavor charge of the $B_{\rm tag}$
$\overset{(-)}{A}{}^{\kappa}$, $A^{\pm 0}$, $A^{0\pm}$	Page 27	Complex decay amplitudes of $B^0(\overline{B}^0) \to (\rho\pi)^0$ and $B^\pm \to \rho^\pm\pi^0$, $\rho^0\pi^\pm$
s_+, s_-, s_0	Page 28	Dalitz plot variables
$\overset{(-)}{A}{}_{3\pi}$	Page 28	Dalitz dependent complex amplitude of $B^0(\overline{B}^0) \to \pi^+\pi^-\pi^0$ decays
f_+, f_-, f_0	Page 28, 81	Complex functions in the Dalitz plot that describe the kinematics of $B^0 \to \rho^+\pi^-$, $\rho^-\pi^+$, and $\rho^0\pi^0$
$(\beta\gamma)_{\Upsilon(4S)}$	Page 30	Boost factor of the $\Upsilon(4S)$ with respect to the laboratory frame
$\mathcal{R}_{K/\pi}$	Page 51	Likelihood ratio to discriminate kaons from pions
$\mathcal{R}_{e/\overline{e}}$	Page 54	Electron ID likelihood ratio
$\mathcal{R}_{\mu/\pi,K}$	Page 54	Muon ID likelihood ratio
$q_{\rm tag}$, r	Page 59	Result returned from the flavor tagging algorithm. $q_{\rm tag}$ is the determined flavor of the $B_{\rm tag}$ and r is the quality of the determination
l	Page 60	Index over the regions of flavor tagging quality
w^l_+, w^l_-	Page 60	Wrong tag fraction for B^0 and $\overline{B}^0$ in each flavor tagging quality region
w_l, Δw_l	Page 60	Wrong tag fraction and wrong tag fraction difference
m', θ'	Page 65	Square Dalitz plot variables
$\boldsymbol{J}$	Page 65	Jacobian for the transformation from usual Dalitz plot variables to the square Dalitz plot variables

Symbol	Defined in	Description
$U_\kappa^\pm$, I_κ, $U_{\kappa\sigma}^{\pm,\mathrm{Re(Im)}}$, $I_{\kappa\sigma}^{\mathrm{Re(Im)}}$	Page 67	Parameters to be fitted in the time-dependent Dalitz plot analysis
ΔE	Page 71	Energy difference
M_{bc}	Page 71	Beam constrained mass
KLR	Page 71	Likelihood ratio to discriminate the continuum background, calculated by combining the modified super Fox-Wolfram and the $\cos\theta_B$
$\mathcal{R}_{BC}$	Page 76	Likelihood ratio used for the best candidate selection
θ_+, θ_-, θ_0	Page 81	Helicity angles of ρ^+, ρ^-, and ρ^0 (Note that these symbols are used in another meaning in a different context)
β, γ	Page 82	Relative amplitudes of $\rho(1450)$ and $\rho(1700)$
$\langle\overline{\beta}\rangle_\kappa$, $\langle\overline{\gamma}\rangle_\kappa$	Page 82	Relative amplitudes of $\rho(1450)$ and $\rho(1700)$ with possible mode dependence
s_{CR}, s_{NR}	Page 87, 124	Dalitz plot variables that describe the dependence of the ΔE-M_{bc} PDF of SCF well
θ_+, θ_-	Page 143	Phases related to the strong interaction
$\widetilde{A}^\kappa$, $\widetilde{A}^{-0}$, $\widetilde{A}^{0-}$	Page 142	$e^{-2i\phi_2}\overline{A}^\kappa$, $e^{-2i\phi_2}A^{-0}$, $e^{-2i\phi_2}A^{0-}$
$\mathcal{C}$, $\Delta\mathcal{C}$, $\mathcal{S}$, $\Delta\mathcal{S}$, $\mathcal{A}_{\rho\pi}^{CP}$	Page 103, 23	CP-violation parameters of the quasi-two-body process $B^0 \to \rho^\pm\pi^\mp$
$\mathcal{A}_{\rho\pi}^{+-}$, $\mathcal{A}_{\rho\pi}^{-+}$	Page 104, 23	Direct CP-violation parameters of the quasi-two-body process $B^0 \to \rho^\pm\pi^\mp$
$\mathcal{A}_{\rho^0\pi^0}$, $\mathcal{S}_{\rho^0\pi^0}$	Page 106	CP-violation parameters of the quasi-two-body process $B^0 \to \rho^0\pi^0$

List of Abbreviations

Name	Defined in	Description
SM	Page 6	Standard Model
CKM	Page 6	Cabbibo-Kobayashi-Maskawa
UT	Page 13	Unitarity Triangle
IP	Page 30	Interaction point
DS-I, DS-II	Page 32	Data sets with old and new silicon vertex detector configurations
PDF	Page 51	Probability Density Function
SDP	Page 65	Square Dalitz Plot
SCF	Page 72	Self Cross Feed
CR, NR	Page 86	Charged track Replaced and Neutral track Replaced (SCF's)