Doctoral Dissertation
博士論文

Measurement of time-dependent CP asymmetries in $B^0 \to K^+K^-K^0_S$ decays with Dalitz-plot analysis at the Belle II experiment

(Belle II 実験における $B^0 \to K^+K^-K^0_S$ 崩壊過程の時間依存 CP 非対称度の Dalitz-Plot を用いた測定)

A Dissertation Submitted for the Degree of Doctor of Philosophy
July 2023
令和5年7月博士(理学)申請

Department of Physics, Graduate School of Science,
The University of Tokyo
東京大学大学院理学系研究科
物理学専攻

Ryohei Sugiura
杉浦　亮平

Abstract

$B^0 \to \phi K_S^0$ decay process is dominated by $b \to s\bar{s}s$ penguin transition, which is sensitive to the contribution from new physics. $B^0 \to \phi K_S^0$ decay is particularly suited for the search of new physics due to the small theoretical uncertainty and suppressed tree diagram contribution within Standard Model. However, the measurement of CP asymmetry for $B^0 \to \phi K_S^0$ is affected by other resonant and non-resonant decays due to the same final state $B^0 \to K^+K^-K_S^0$. In order to solve interference between decay processes and measure CP violating parameters of $B^0 \to \phi K_S^0$ decay precisely, a study of time-dependent CP asymmetries including multiple decay processes is carried out with Dalitz plot technique. $B^0 \to f_0 K_S^0$, $B^0 \to f_X K_S^0$, $B^0 \to \chi_{c0} K_S^0$ and non-resonant $B^0 \to K^+K^-K_S^0$ are taken into consideration to solve possible interference.

This thesis describes a study of time-dependent CP asymmetry using e^+e^- energy-asymmetric collision data including 387×10^6 $B^0\bar{B}^0$ pairs collected at Belle II experiment. The result is

$$\phi_1(\phi K_S^0) = 28.9 \pm 10.1(\text{stat}) \pm 1.7(\text{syst})^\circ$$
$$A_{CP}(\phi K_S^0) = 0.07 \pm 0.18(\text{stat}) \pm 0.04(\text{syst}).$$

The existence of CP violation is confirmed at 2.4σ significance, and the result is consistent with SM within 0.6σ.

Personal Contribution and Outline

The CP asymmetry measurement reported in this thesis is based on the contribution and effort from all the member of Belle II collaboration and the SuperKEKB accelerator group. The maintaining of the Belle II experiment and the development of analysis frameworks are done by the collaboration. The main personal contribution by the author is the development of the K_S^0 selection method called KsSelector as mentioned in Sec. 5.4.

The selection of K_S^0 through $K_S^0 \to \pi^+\pi^-$ is critical to good CP asymmetry measurement due to large number of mis-reconstructed K_S^0 candidates originating from continuum background evetns. The K_S^0 selection method based on multivariate machine-learning algorithm was developed previously from other collaborator, but the previous method suffers from the problem due to the correlation between discriminant calculated by K_S^0 selection method and K_S^0 mass itself. The author improved the K_S^0 selection method and offered this method as a common tool called KsSelector which can be use widely among Belle II collaboration. The author also analyzed the performance of K_S^0 selection method, and showed that KsSelector have no significant correlation between discriminant calculated by KsSelector and important observables related to K_S^0. The author also showed that the performance of KsSelector is also improved from the K_S^0 selection method that was previously adopted in the Belle experiment.

In this thesis, the CP violating mechanism in B meson system is described in Sec. 2. The Belle II experiment and Belle II detector are explained in Sec. 3. The Dalitz-plot analysis technique used for the analysis is described in Sec. 4. The method of reconstructing B meson candidates and calculating signal fraction is explained in Sec. 5 and Sec. 6. The detailed description for CP asymmetry measurement procedure and expected systematic errors is described in Sec. 7 and Sec. 8. The discussion and conclusion is given in Sec. 9 and Sec. 10.

Contents

1 Introduction

The current elementary particle physics is largely described by the Standard Model (SM), which provides a framework for describing electromagnetic, weak, and strong interactions. The SM describes three generations of quarks and leptons, gauge bosons that mediate the interactions, and the Higgs boson. Nevertheless, certain phenomena remain unexplained within the framework of the SM. One notable example is the observation of CP asymmetry in the universe, manifesting as the absence of antimatter, which cannot be entirely explained by the CP violation mechanism within the SM. Such discrepancies between the predictions of the SM and observations act as a motivation for the exploration of new physics beyond the SM. Consequently, diverse experiments have been conducted worldwide in search for new physics.

The first observation of CP violation in the K meson system was accomplished in 1964 [13]. Subsequently, in 1973, the Kobayashi-Maskawa theory emerged, establishing a framework for CP violation within the confines of the SM [27]. Starting data collection in 1999, the Belle experiment aimed to generate a substantial quantity of B meson pairs via energy-asymmetric e^+e^- collisions and observe CP violation within the B meson system. This experiment was conducted at the High Energy Accelerator Research Organization (KEK) situated in Tsukuba, Japan. The B meson pairs were produced using KEKB accelerator, while their decay processes were observed by the Belle detector. Consequently, the Belle experiment verified the existence of CP violation in the B meson system, and thus providing experimental evidence supporting the Kobayashi-Maskawa theory.

The Belle experiment, despite its success in confirming CP violation in the B meson system, did not yield any significant hints, beyond a 3-sigma level, of new physics beyond the SM. This limitation was primarily due to the large statistical uncertainties in the measurements. Based on these background, the Belle II experiment is planned as an upgrade to the original Belle experiment. The Belle II experiment involves the upgrade of the KEKB accelerator to the more powerful SuperKEKB accelerator. The primary objective of the Belle II experiment is to obtain a dataset with statistics approximately 50 times larger than that of the Belle experiment. By significantly increasing the amount of data collected, Belle II aims to enhance its sensitivity to potential new physics phenomena that lie beyond the SM.

In this thesis, the CP violating mechanism in B meson system is described in Sec. 2. The Belle II experiment and Belle II detector are explained in Sec. 3. The Dalitz-plot analysis technique used for the analysis is described in Sec. 4. The method of reconstructing B meson candidates and calculating signal fraction is explained in Sec. 5 and Sec. 6. The detailed description for CP asymmetry measurement procedure and expected systematic errors is described in Sec. 7 and Sec. 8. The discussion and conclusion is given in Sec. 9 and Sec. 10.

2 CP asymmetries in B Meson Decay

This section explains the mechanism of CP violation within the B meson system, accompanied by an explanation of the precise measurement of CP asymmetries in B meson decay. This measurement serves as a probing instrument for identifying potential new physics phenomena that lie beyond the scope of the Standard Model.

2.1 Unitarity Triangle

With mass eigenstates of quarks, a Lagrangian governing the charged-current weak interaction can be expressed as

$$\mathcal{L}_{\text{int}} = -\frac{g}{\sqrt{2}} \left(\bar{u_L}, \bar{c_L}, \bar{t_L} \right) \gamma^{\mu} V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} W_{\mu} + h.c. \tag{1}$$

where W_μ is the W-boson field, γ^μ is gamma matrices, g is the gauge coupling of $SU(2)_L$ gauge group, and V_{CKM} is the Cabbibo-Kobayashi-Maskawa (CKM) matrix.

Here we introduce weak interaction doublets of quarks denoted as $(u, d')^T$, $(c, s')^T$, and $(t, b')^T$. Here, u, c, and t represent the mass eigenstates of up-type quarks, while d', s', and b' correspond to the weak interaction doublet partners. The CKM matrix is a unitary matrix of dimensions 3 × 3, which characterizes the relationship between (d', s', b') and the mass eigenstates of down-type quarks (d, s, b). It can be represented by the following equation:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{2}$$

The CKM matrix, denoted as V_{CKM}, can be explicitly written as:

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{3}$$

It is possible to parametrize the CKM matrix using three rotation angles and six phases. However, five out of the six phases can be absorbed by rotating the individual left-handed quark fields. Consequently, the CKM matrix can be represented using three angles, namely θ_{12}, θ_{13}, and θ_{23}, along with one phase denoted as δ, as initially proposed by Kobayashi and Maskawa [27]. Thus, the CKM angles can be parametrized as follows:

$$V_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}s_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ -s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} . \tag{4}$$

where $c_{ij} = \cos(\theta_{ij})$, $s_{ij} = \sin(\theta_{ij})$, and δ is the complex phase.

In order to highlight the observed hierarchy in the mixing angles, namely $|V_{ub}| \ll |V_{cb}| \ll |V_{us}|$, the Wolfenstein parametrization [36] is commonly employed. Using the

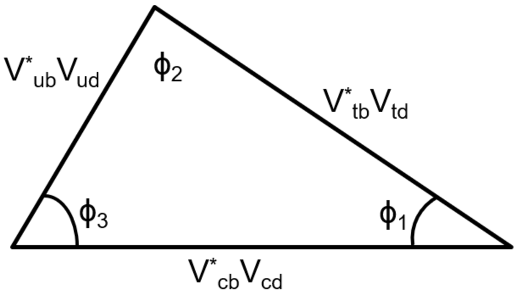


Figure 1: Unitarity triangle of the CKM matrix. The closed triangle can be drawn from the unitarity of V_{CKM}.

Wolfenstein parametrization, the CKM matrix can be expressed as:

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{5}$$

Here, the parameter $\lambda \simeq 0.22$, while the other parameters are of order $\mathcal{O}(1)$.

V_{CKM} is unitary matrix, and thus V_{CKM} satisfies

$$V^{\dagger}_{CKM} V_{\mathrm{CKM}} = 1. \tag{6}$$

This equation leads to

$$V^*_{ub}V_{ud} + V^*_{cb}V_{cd} + V^*_{tb}V_{td} = 0. \tag{7}$$

The terms $V^*_{ij}V_{ik}$ can be interpreted as vectors on the complex plane. Therefore, Eq. 7 represents the unitarity triangle on the complex plane, as depicted in Fig. 1.

Those ϕ_1, ϕ_2, and ϕ_3 can be represented using V_{ij} in CKM matrix as

$$\phi_1 = arg\left(-\frac{V^*_{cb}V_{cd}}{V^*_{tb}V_{td}}\right), \tag{8}$$

$$\phi_2 = arg\left(-\frac{V^*_{ub}V_{ud}}{V^*_{tb}V_{td}}\right), \tag{9}$$

$$\phi_3 = arg\left(-\frac{V^*_{cb}V_{cd}}{V^*_{ub}V_{ud}}\right) \tag{10}$$

and the unitarity of CKM matrix is expressed as a closure of the unitarity triangle.

The scaled unitarity triangle has three vertices located at $(0,0)$, $(1,0)$ and $(\overline{\rho}, \overline{\eta})$, where $(\overline{\rho}, \overline{\eta})$ is defined as

$$\overline{\rho} = \left(1 - \frac{\lambda^2}{2}\right)\rho, \qquad \overline{\eta} = \left(1 - \frac{\lambda^2}{2}\right)\eta. \tag{11}$$

The latest constraints of Wolfenstein parameters $(\overline{\rho}, \overline{\eta})$ from various experiments is shown in Fig. 2 [12].

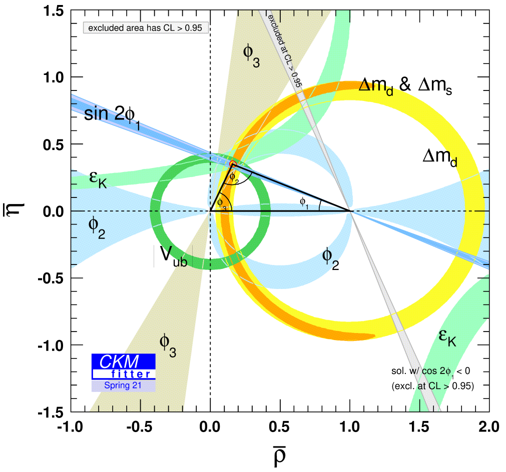


Figure 2: The constraints of Wolfenstein parameters $(\bar{\rho}, \bar{\eta})$ [12]. Each colored area represents the 95% CL area. The red hashed region represents the 68% CL region of the vertex of unitarity triangle.

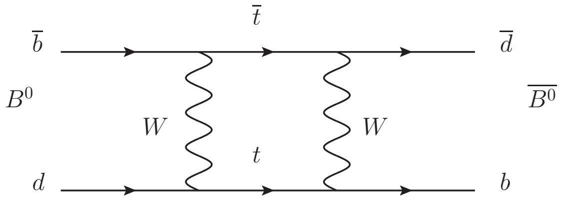


Figure 3: Feynman diagram for B^0-$\overline{B}^0$ mixing through box diagram, where neutral B mesons can change its flavor via loop diagram.

2.2 B meson decay and CP asymmetries

We describe the time evolution of the B-meson system, followed by an introduction to CP asymmetries within this system.

The B^0 meson and its antiparticle, $\bar{B^0}$, can mix through a weak interaction mediated by a loop diagram, as illustrated in Fig. 3. Consequently, the initial state of the B-meson system and its time evolution can be expressed as a superposition of the B^0 and $\bar{B^0}$ states:

$$|\psi(0)\rangle = a(0)\left|B^0\right\rangle + b(0)\left|\bar{B^0}\right\rangle, \tag{12}$$

$$|\psi(t)\rangle = a(t)\left|B^0\right\rangle + b(t)\left|\bar{B^0}\right\rangle + \sum_i c_i(t)\left|f_i\right\rangle \tag{13}$$

Here, f_i represents the final states to which the B-meson system can decay, and $a(t)$, $b(t)$, and $c_i(t)$ are coefficients that depend on time.

$|\psi(t)\rangle$ follows time-dependent Schrödinger equation as

$$i\hbar\frac{\partial}{\partial t}\psi(t) = \mathbf{H}\psi(t) \tag{14}$$

where $\psi(t) = (a(t), b(t))^T$. The matrix $\mathbf{H}$ is given as

$$\mathbf{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} \tag{15}$$

$$= \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{21} - \frac{i}{2}\Gamma_{21} & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix} \tag{16}$$

where $\mathbf{M}$ and $\mathbf{\Gamma}$ are 2×2 Hermitian matrices, and representing mass and decay respectively. With CPT invariance, the diagonal elements of $\mathbf{M}$ and $\mathbf{\Gamma}$ obeys the relation

$$M_{11} = M_{22}, \Gamma_{11} = \Gamma_{22} \tag{17}$$

The B^0 and $\bar{B}^0$ mesons are flavor eigenstates, and thus they can be expressed in terms of mass eigenstates as follows:

$$|B_L\rangle = p\left|B^0\right\rangle + q\left|\bar{B}^0\right\rangle \tag{18}$$

$$|B_H\rangle = p\left|B^0\right\rangle - q\left|\bar{B}^0\right\rangle \tag{19}$$

Here, p and q are complex parameters, and $|B_L\rangle$ and $|B_H\rangle$ represent the light and heavy mass eigenstates, respectively. The ratio between p and q can be determined by solving the mass eigenvalue of the system, and it is given by:

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{20}$$

The eigenvalues for the mass eigenstates are expressed in terms of the mass and decay width of B_H and B_L as

$$M_L = m_L - i\Gamma_L, M_H = m_H - i\Gamma_H, \tag{21}$$

where m_i and Γ_i (for $i = H, L$) represent the mass and decay width of each mass eigenstate. The time evolution of the mass eigenstates is given by

$$|B_L(t)\rangle = e^{-im_L t}e^{-\Gamma_L t/2}|B_L\rangle \tag{22}$$

$$|B_H(t)\rangle = e^{-im_H t}e^{-\Gamma_H t/2}|B_H\rangle. \tag{23}$$

By solving the time evolution equation in Eq. 15, the time evolution of the flavor eigenstates is obtained as

$$\left|B^0(t)\right\rangle = g_+(t)\left|B^0\right\rangle - \frac{q}{p}g_-(t)\left|\bar{B}^0\right\rangle \tag{24}$$

$$\left|\bar{B}^0(t)\right\rangle = g_+(t)\left|\bar{B}^0\right\rangle - \frac{p}{q}g_-(t)\left|B^0\right\rangle. \tag{25}$$

where

$$g_\pm = \frac{1}{2}(e^{-im_H t - \Gamma_H t/2} \pm e^{-im_L t - \Gamma_H t/2}). \tag{26}$$

Using this result, one can proceed to compute the decay amplitudes of B^0 and $\bar{B}^0$ transitioning into a CP eigenstate denoted as f_{CP}. The expressions for the time-dependent decay rates $A_{f_{CP}} = \langle f_{CP} \,|\, H \,|\, B^0 \rangle$ and $A_{\bar{f}_{CP}} = \left\langle f_{CP} \,\middle|\, H \,\middle|\, \bar{B}^0 \right\rangle$ are given as

$$\begin{aligned}
\Gamma(B^0(t) \to f_{CP}) &= |\left\langle f_{CP} \,|\, H \,|\, B^0 \right\rangle|^2 \\
&= e^{-\Gamma t} \Bigg[\left(|A_{f_{CP}}|^2 + |\frac{q}{p} A_{\bar{f}_{CP}}|^2 \right) \cosh(\frac{1}{2}\Delta\Gamma t) \\
&\quad + \left(|A_{f_{CP}}|^2 - |\frac{q}{p} A_{\bar{f}_{CP}}|^2 \right) \cos(\Delta m_d t) \\
&\quad + \quad 2\,\mathrm{Re}\left(\frac{q}{p} A^*_{f_{CP}} A_{\bar{f}_{CP}} \right) \sinh(\frac{1}{2}\Delta\Gamma t) \\
&\quad -2\,\mathrm{Im}\left(\frac{q}{p} A^*_{f_{CP}} A_{\bar{f}_{CP}} \right) \sin(\Delta m_d t) \Bigg],
\end{aligned} \tag{27}$$

$$\begin{aligned}
\Gamma(B^{\bar{0}}(t) \to f_{CP}) &= |\left\langle f_{CP} \,|\, H \,|\, \bar{B}^0 \right\rangle|^2 \\
&= e^{-\Gamma t} \Bigg[\left(|\frac{p}{q} A_{f_{CP}}|^2 + |A_{\bar{f}_{CP}}|^2 \right) \cosh(\frac{1}{2}\Delta\Gamma t) \\
&\quad + \left(|\frac{p}{q} A_{f_{CP}}|^2 - |A_{\bar{f}_{CP}}|^2 \right) \cos(\Delta m_d t) \\
&\quad + \quad 2\,\mathrm{Re}\left(\frac{p}{q} A_{f_{CP}} A^*_{\bar{f}_{CP}} \right) \sinh(\frac{1}{2}\Delta\Gamma t) \\
&\quad -2\,\mathrm{Im}\left(\frac{p}{q} A_{f_{CP}} A^*_{\bar{f}_{CP}} \right) \sin(\Delta m_d t) \Bigg],
\end{aligned} \tag{28}$$

where $\Delta\Gamma = \Gamma_H - \Gamma_L$ and $\Delta m_d = m_H - m_L$. Using these decay amplitudes and assuming $\Delta\Gamma/\Gamma \ll 1$ and $|q/p| = 1$, the time-dependent CP asymmetry for f_{CP} is given as

$$\begin{aligned}
A_{CP}(t) &= \frac{\Gamma(\bar{B}^0(t) \to f_{CP}) - \Gamma(B^0(t) \to f_{CP})}{\Gamma(\bar{B}^0(t) \to f_{CP}) + \Gamma(B^0(t) \to f_{CP})} \\
&= S_{CP} \sin \Delta m_d t + A_{CP} \cos \Delta m_d t
\end{aligned} \tag{29}$$

where we introduce CP asymmetry parameter S_{CP} and A_{CP} as

$$\lambda_{CP} = \frac{q}{p} \frac{\bar{A}_f}{A_f}, \tag{30}$$

$$S_{CP} = \frac{2\,\mathrm{Im}\,\lambda_{CP}}{1 + |\lambda_{CP}|^2} \tag{31}$$

$$A_{CP} = \frac{|\lambda_{CP}|^2 - 1}{|\lambda_{CP}|^2 + 1}. \tag{32}$$

The parameters S_{CP} and A_{CP} represent the mixing-induced and direct CP violation respectively. For the sake of convenience, the parameter λ_{CP}, which is independent of phase conventions, is introduced.

Mixing-induced CP violation occurs when the imaginary part of λ_{CP} takes non-zero values, indicating a disparity between the phase of $B^0 - \bar{B}^0$ mixing and the decay phase.

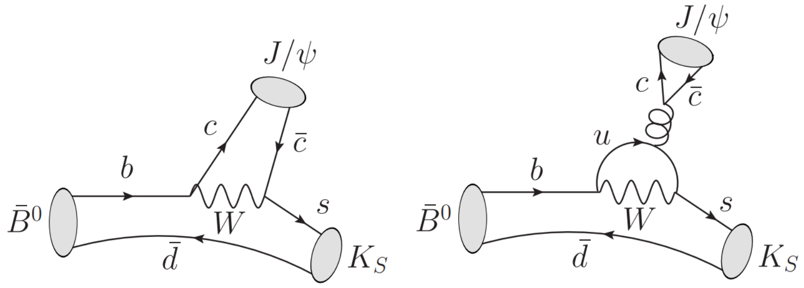


Figure 4: The diagrams contributing to the tree (left) and penguin (right) amplitude for $B^0 \to J/\psi K_S^0$ decay [6].

CP violation also manifests itself through the direct CP violation term A_{CP} when $|\lambda_{CP}|^2 \neq 1$. This form of CP violation arises when multiple diagrams contribute to the decay of B-mesons, each possessing distinct weak and strong phases.

2.3 Measurement of CP asymmetries

In this section, we discuss CP asymmetries for $b \to q\bar{q}s$ trainsition to measure $\sin 2\phi_1$ or ϕ_1.

2.3.1 Measurement of $\sin 2\phi_1$ in $b \to c\bar{c}s$ transition

The decay process $B^0 \to J/\psi K_S^0$ is suitable for the precise measurement of ϕ_1 due to its small theoretical uncertainty from SM prediction, substantial branching fraction, and low background yield.

The $B^0 \to J/\psi K_S^0$ decay occurs via the $b \to c\bar{c}s$ transition and is primarily governed by tree and penguin Feynman diagrams, as depicted on the left side of Fig. 4. The loop diagram, commonly referred to as the penguin diagram, makes a significantly smaller contribution compared to the tree diagram [11].

Utilizing the relationship between CKM matrix elements $V_{us}V_{ub}^*/V_{cs}V_{cb}^* \simeq \mathcal{O}(\lambda^2)$, the ratio of decay amplitudes can be expressed as follows:

$$\frac{\bar{A}(J/\psi K^0)}{A(J/\psi K^0)} = \frac{V_{cs}^* V_{cb}}{V_{cs} V_{cb}^*}. \tag{33}$$

In terms of q/p, combined with the explicit calculation of $\Gamma_{12}/M_{12} \simeq 10^{-2}$, an estimation for q/p can be obtained to be

$$\frac{q}{p} \simeq \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}. \tag{34}$$

Since the $B^0 \to J/\psi K_S^0$ includes K_S^0 in the final state, the K^0-$\bar{K}^0$ mixing should be taken into consideration in the same manner with B^0-$\bar{B}^0$, resulting in the expression

$$\frac{\bar{A}(J/\psi K_S^0)}{A(J/\psi K_S^0)} = \frac{V_{cd}^* V_{cs}}{V_{cd} V_{cs}^*} \frac{\bar{A}(J/\psi K^0)}{A(J/\psi K^0)} \tag{35}$$

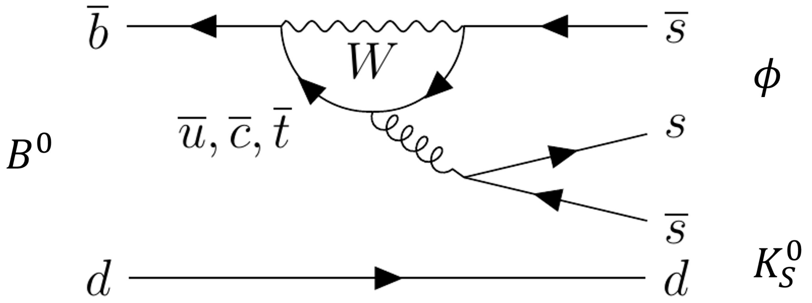


Figure 5: The example of diagrams contributing to the penguin amplitude for $B^0 \to \phi K_S^0$ decay.

Following these discussions, $\lambda_{J/\psi K_S^0}$ is calculated to be

$$\lambda_{J/\psi K_S^0} = -\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \frac{V_{cd}^* V_{cs}}{V_{cd} V_{cs}^*} \frac{V_{cs}^* V_{cb}}{V_{cs} V_{cb}^*} = -e^{2i\phi_1}. \tag{36}$$

where the negative sign incorporates the effect of $J/\psi K_S^0$ being a CP-odd eigenstate. Consequently, the CP-violating parameters are calculated as

$$A_{CP} = 0, S_{CP} = \sin 2\phi_1 \tag{37}$$

As indicated in Eq. 37, the unitarity angle $\sin 2\phi_1$ can be accurately determined through the measurement of time-dependent CP asymmetries in the $B^0 \to J/\psi K_S^0$ decay. However, an ambiguity between ϕ_1 and $\pi - \phi_1$ still persists.

2.3.2 Measurement of $\sin 2\phi_1$ in $b \to s\bar{s}s$ transition

In contrast to the $B^0 \to J/\psi K_S^0$ decay channel, the measurement of S_{CP} and A_{CP} in penguin-dominated decay channels involving $b \to q\bar{q}s$ transitions (where $q = u, d, s$) serves as a good probe for investigating contributions from new physics. The $B^0 \to \phi K_S^0$ decay occurs through a $b \to s\bar{s}s$ flavor-changing neutral current, proceeding via a second-order loop diagram, as depicted in Fig. 5. Owing to the suppression of the CKM tree-level diagram and the relatively small penguin-level diagram, it becomes feasible to detect the effects of new physics through contributions to the loop diagram in the penguin amplitude.

By neglecting the small contribution from the CKM-suppressed tree diagram, the decay amplitude of $B^0 \to \phi K_S^0$ can be expressed as

$$A_{\phi K_S^0} = \sum_{q=u,c,t} (V_{qb}^* V_{qs}) P_{\phi K_S^0}^q \tag{38}$$

where $P_{\phi K_S^0}^q$ represents the penguin amplitude involving the quarks u, c, and t in the loop with a W boson. Utilizing the unitarity of the CKM matrix, this expression can be transformed into the following form:

$$A_{\phi K_S^0} = V_{cb}^* V_{cs} \left((P_{\phi K_S^0}^c - P_{\phi K_S^0}^t) + \frac{V_{ub}^* V_{us}}{V_{cb}^* V_{cs}} (P_{\phi K_S^0}^u - P_{\phi K_S^0}^t) \right). \tag{39}$$

Based on the relation $|V_{ub}| \ll |V_{cb}| \ll |V_{us}| < |V_{cs}|$, the ratio of decay amplitudes can be approximated as

$$\frac{\bar{A}(\phi K_S^0)}{A(\phi K_S^0)} \simeq \frac{V_{cs}^* V_{cb}}{V_{cs} V_{cb}^*}, \tag{40}$$

which yields the same value as the calculation for $B^0 \to J/\psi K_S^0$. Consequently, the parameter $\lambda_{\phi K_S^0}$ and the CP asymmetry parameters for $B^0 \to \phi K_S^0$ can be determined in a similar manner to the $B^0 \to J/\psi K_S^0$ decay, resulting in the following expressions:

$$\lambda_{\phi K_S^0} \simeq -e^{2i\phi_1}. \tag{41}$$

$$A_{CP} = 0, S_{CP} = \sin 2\phi_1 \tag{42}$$

where the odd CP eigenvalue of ϕK_S^0 is taken into consideration.

2.4 Possible effect of new physics on time-dependent CP asymmetry

In the previous section, the contamination from the tree diagram was neglected in the calculation of CP asymmetries. However, the presence of the tree diagram and potential contributions from new physics can lead to a deviation between $\sin 2\phi_1$ and the measured value of S_{CP} [8]. This deviation can be expressed as follows:

$$-\eta_{CP} S_{CP} = \sin 2\phi_1 + \Delta S_{CP} \tag{43}$$

where ΔS_{CP} is given by

$$\Delta S_{CP} = 2\cos 2\phi_1 \sin\phi_3 |\epsilon_{uc}| \operatorname{Re} r_{CP}^T + \Delta S_{CP}^{NP}. \tag{44}$$

In the equation, $\epsilon_{uc} = V_{us}V_{ub}^*/V_{cs}V_{cb}^* = \mathcal{O}(\lambda^2)$, where λ represents a parameter associated with CKM matrix elements. The ratio r_{CP}^T denotes the ratio between penguin and tree diagram amplitudes. The term ΔS_{CP}^{NP} corresponds to a possible contribution from new physics. According to the calculation using QCD factorization [8], the predicted value for ΔS_{CP} in the ϕK_S^0 final state is $\Delta S_{CP} = 0.02^{+0.01}_{-0.01}$. This value is notably small compared to other charmless two-body final states. Currently, there is no tension between the predictions for ΔS_{CP} and experimental results. Furthermore, the theoretical uncertainties associated with ΔS_{CP} are much smaller than the experimental measurements. Therefore, the measurement of the time-dependent CP asymmetry in $B \to \phi K_S^0$ decay provides a theoretically clean analysis for investigating contributions from new physics.

The parametrization of the amplitudes for $B \to \phi K_S^0$ decay in the presence of new physics (NP) can be expressed as

$$A^{SM}(\phi K_S^0) = |A^{SM}|e^{i\delta_{SM}}, \qquad A^{NP}(\phi K_S^0) = |A^{NP}|e^{i\theta_{NP}}e^{i\delta_{NP}}, \tag{45}$$

$$\bar{A}^{SM}(\phi K_S^0) = |\bar{A}^{SM}|e^{i\delta_{SM}}, \qquad \bar{A}^{NP}(\phi K_S^0) = |\bar{A}^{NP}|e^{-i\theta_{NP}}e^{i\delta_{NP}}. \tag{46}$$

where δ_i (with i = SM or NP) represents the CP-conserving phase, and θ_{NP} is the CP-violating phase associated with new physics. Using this parameterization, the CP

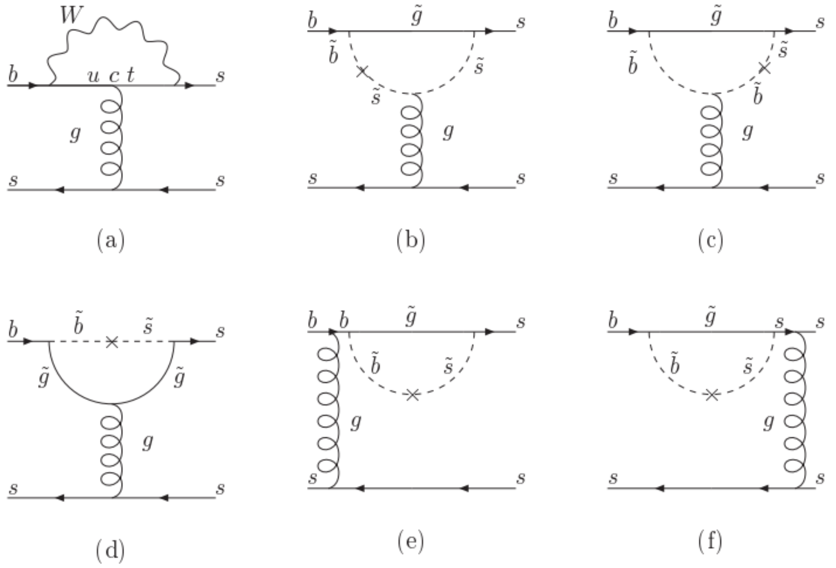


Figure 6: The standard model contribution (a) and gluino-down contributions (b) to (f) for $B \to \phi K_S^0$ decay [15]. The cross denotes mass insertion $(\delta^d_{AB})_{23}$ term in SUSY.

asymmetry for $B \to \phi K_S^0$ decay is given by

$$S_{CP} = \frac{\sin 2\phi_1 + 2\frac{|A_{NP}|}{|A_{SM}|}\cos\delta_{12}\sin(\theta_{NP}+2\phi_1) + \left(\frac{|A_{NP}|}{|A_{SM}|}\right)^2 \sin(2\theta_{NP}+2\phi_1)}{1 + 2\frac{|A_{NP}|}{|A_{SM}|}\cos\delta_{12}\cos\theta_{NP} + \left(\frac{|A_{NP}|}{|A_{SM}|}\right)^2}, \tag{47}$$

where $\delta_{12} = \delta_{SM} - \delta_{NP}$. By measuring the CP asymmetry S_{CP} in $B \to \phi K_S^0$ decay, it becomes possible to constrain the weak phase θ_{NP} associated with CP violation in the presence of new physics.

One of the NP candidates that affects the CP asymmetries in the decay of $B \to \phi K_S^0$ is the supersymmetric (SUSY) model [18]. The Fig. 6 presents candidates for the $b \to s\bar{s}s$ process through SUSY particles with the mass insertion approximation [15], wherein loop diagrams involving gluinos and squarks are considered. Within the basis adopted by mass insertion approximation, the coupling of fermions and sfermions to neutral gauginos leads to flavor violation due to the off-diagonal term of the sfermion mass matrix.

2.5 Measurement of CP violation in Belle II experiment

In the Belle II experiment, pairs of B mesons are coherently produced through the $\Upsilon(4S) \to B^0\bar{B}^0$ decays. The B mesons originating from the $\Upsilon(4S) \to B^0\bar{B}^0$ decay conserve the quantum number $C = -1$, establishing an initial state where B^0 and $\bar{B}^0$ are correlated with $C = -1$. B^0 is a pseudo-scalar particle, leading to an orbital angular momentum of $L = 1$ for the B meson pairs. Consequently, according to Bose-Einstein statistics, if one of the B mesons is observed as B^0 at time $t = 0$, the other B meson is expected to be $\bar{B}^0$; thus, the formation of B^0B^0 and $\bar{B}^0\bar{B}^0$ pairs is forbidden. The time evolution of B meson pairs can be represented as follows:

$$|B_{CP}B_{tag}(t_{CP}=0,t_{tag}=0)\rangle = \frac{1}{\sqrt{2}}\left(\left|B^0_{CP}\bar{B}^0{}_{tag}\right\rangle - \left|\bar{B}^0{}_{CP}B^0_{tag}\right\rangle\right). \tag{48}$$

Then by using Eq. 24 and Eq. 48 leads to the description of the time evolution of B meson system as

$$\begin{aligned}|B_{CP}B_{tag}(t_{CP},t_{tag})\rangle = &\frac{1}{\sqrt{2}}e^{-\Gamma(t_{CP}+t_{tag})/2} \qquad (49)\\ &[\cos\frac{\Delta m_d\Delta t}{2}\left(\left|B^0_{CP}\bar{B}^0{}_{tag}\right\rangle - \left|\bar{B}^0{}_{CP}B^0_{tag}\right\rangle\right) + \\ &i\sin\frac{\Delta m_d\Delta t}{2}\left(\frac{p}{q}\left|B^0_{CP}B^0_{tag}\right\rangle - \frac{q}{p}\left|\bar{B}^0{}_{CP}\bar{B}^0{}_{tag}\right\rangle\right)]\end{aligned} \tag{50}$$

where Δt is defined as $\Delta t = t_{CP} - t_{tag}$.

Here we examine the case where one B meson decays into the CP eigenstate f_{CP} while the other B meson decays into a flavor-specific state f_{tag}, allowing us to determine the flavor of the B meson. We assume that B^0 only decays into f_{tag} and $\bar{B}^0$ only decays into $\bar{f_{tag}}$. Under the assumption of no direct CP violation in B_{tag} decays, which implies $A_{f_{tag}} = \bar{A}_{\bar{f_{tag}}}$ and $|q/p| = 1$, the time-dependent decay rate for the states $|f_{CP}f_{tag}\rangle$ and $\left|f_{CP}\bar{f_{tag}}\right\rangle$ can be expressed as follows:

$$\begin{aligned}\frac{d\Gamma}{dt_{tag}dt_{CP}} \propto\ & e^{-\Gamma(t_{CP}+t_{tag})}[(|A_{f_{CP}}|^2 + |\bar{A}_{\bar{f}_{CP}}|^2) \qquad (51)\\ &- q_{tag}(|A_{f_{CP}}|^2 - |\bar{A}_{\bar{f}_{CP}}|^2)\cos(\Delta m_d\Delta t) + \\ &q_{tag}\,\mathrm{Im}(2\eta_{CP}\frac{q}{p}\bar{A}_{f_{CP}}A^*_{\bar{f}_{CP}})\sin(\Delta m_d\Delta t).\end{aligned}$$

Here, the value of q_{tag} represents the flavor of the tag-side B meson and can take on the values of ±1. Specifically, $q_{tag} = +1(-1)$ when the tag-side decays into f_{tag} ($\bar{f}_{tag}$), and η_{CP} denotes the eigenvalue of the CP-side decay channel. By integrating the decay rate over t_{CP} and t_{tag} and normalizing the overall decay width, the time-dependent decay rate, as a function of the decay time difference Δt, is given by:

$$\frac{d\Gamma}{d\Delta t} \propto e^{-\Gamma|\Delta t|}\left(1 + q_{tag}A_{CP}\cos(\Delta m_d\Delta t) + q_{tag}S_{CP}\sin(\Delta m_d\Delta t)\right) \tag{52}$$

Here, A_{CP} and S_{CP} are defined in the same manner as the notation used in Eq. 31. Therefore, the CP-violating parameters can be determined by measuring the asymmetry of the decay rate with respect to the decay time difference Δt, which can be measured precisely in the Belle II experiment.

2.6 Experimental results from previous analyses

Both Belle [31] and BaBar [29] previously measured CP asymmetries of $B \to \phi K^0_S$ decay using $B \to K^+K^-K^0_S$ Dalitz analysis. The Belle analysis obtained four solutions for

$\phi_1(B^0 \to \phi K_S^0)$ as

$$\phi_1 = (32.2 \pm 9.0 \pm 2.6 \pm 1.4)^\circ \tag{53}$$
$$\phi_1 = (26.2 \pm 8.8 \pm 2.7 \pm 1.2)^\circ \tag{54}$$
$$\phi_1 = (27.3 \pm 8.6 \pm 2.8 \pm 1.3)^\circ \tag{55}$$
$$\phi_1 = (24.3 \pm 8.0 \pm 2.9 \pm 5.2)^\circ \tag{56}$$

where errors are statistical, systematic, and Dalitz model uncertainties.

The BaBar analysis obtained single solution as

$$\phi_1 = (21 \pm 6 \pm 2)^\circ \tag{57}$$

where errors are statistical and systematic uncertainties.

The comparison of this analysis against those previous analyses is discussed in Sec. 9.4.

3 Belle II experiment

The Belle II experiment, conducted at KEK in Tsukuba-city, Japan, is a B-factory experiment that started operating from 2019. The SuperKEKB accelerator is responsible for accelerating electrons to 7 GeV and positrons to 4 GeV and have them colliding with high luminosity, where the center-of-mass energy of electrons and positrons is set to 10.58 GeV corresponding to the $\Upsilon(4S)$ resonance which subdominantly decays into B meson pairs. These high-energy beams collide at an interaction point where the Belle II detector is positioned.

3.1 SuperKEKB accelerator

The overview of SuperKEKB accelerator is shown in Fig. 7. In SuperKEKB, the energy of electron and positron is changed to 7.0 GeV and 4.0 GeV respectively. SuperKEKB consists of a linear accelerator with 600m length and a ring accelerator with the circumstance of 3km.

The electron and positron is accelerated to 7.0 GeV and 4.0 GeV in linear accelerator, and injected to electron storage ring called high energy ring (HER) and positron storage ring called low energy ring (LER). The electron and positron collide at one collision point on these rings. Belle II detector is positioned around the collision point. The center of mass energy is set to 10.58 GeV, which is the mass of $\Upsilon(4S)$.

The number of collision event per unit time R [event/sec] can be represented as

$$R = \mathcal{L}\sigma \tag{58}$$

where $\mathcal{L}$ and σ represents instantaneous luminosity $[m^{-2}s^{-1}]$ and interaction cross section [cm^{-2}] for target event respectively. In SuperKEKB accelerator, the luminosity is 40 times higher luminosity than that of KEKB accelerator.

The luminosity can be written as is in Eq. 59.

$$L = \frac{\gamma_\pm}{2er_e}\left(1 + \frac{\sigma^*_{y\mp}}{\sigma^*_{x\mp}}\right)\frac{I_\pm \xi_{y\pm}}{\beta^*_{y\pm}}, \tag{59}$$

where γ is lorentz factor, e is elemental charge, r_e is classical electron radius, I is beam current, $\sigma_{x,y}$ is beam size along each axis, and $\beta^*_{y\mp}$ is beta function along y axis. $\xi_{y\pm}$ represents beam-beam parameter, which means the effect of interaction between electron and positron beams. Each $+$ and $-$ represents positron and electron.

When we squeeze $\beta^*_{y\mp}$, hourglass effect and synchrobetatron resonance disturb the luminosity to be maximized. In order to avoid these difficulty, the new schematics called "nano beam scheme" is adopted. The concept of nano beam scheme is shown in Fig. 9. In nano beam scheme, the large crossing angle of electron and positron bunches leads to the smaller $\beta^*_{y\mp}$ in the interaction region. Together with the enforcement of beam current, the luminosity in SuperKEKB is expected to become 30 times larger than KEKB.

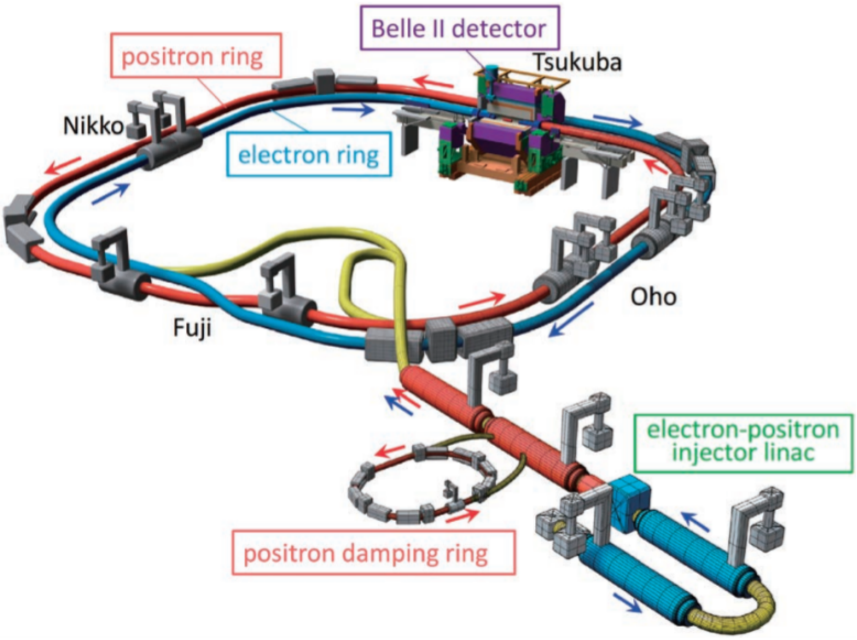


Figure 7: Overview of SuperKEKB accelerator [2].

		LER (e+)	HER (e-)	units
Beam Energy	E	4	7	GeV
Half Crossing Angle	ϕ	41.5		mrad
Horizontal Emittance	ε_x	3.2(2.7)	2.4(2.3)	nm
Emittance ratio	$\varepsilon_y/\varepsilon_x$	0.40	0.35	%
Beta Function at the IP	β_x^*/β_y^*	32 / 0.27	25 / 0.41	mm
Horizontal Beam Size	σ_x^*	10.2(10.1)	7.75(7.58)	μm
Vertical Beam Size	σ_y^*	59	59	nm
Betatron tune	ν_x/ν_y	45.530/45.570	58.529/52.570	
Momentum Compaction	α_c	2.74×10^{-4}	1.88×10^{-4}	
Energy Spread	σ_ε	$8.14(7.96) \times 10^{-4}$	$6.49(6.34) \times 10^{-4}$	
Beam Current	I	3.60	2.62	A
Number of Bunches/ring	n_b	2503		
Energy Loss/turn	U_0	2.15	2.50	MeV
Total Cavity Voltage	V_c	8.4	6.7	MV
Synchrotron Tune	ν_s	-0.0213	-0.0117	
Bunch Length	σ_z	6.0(4.9)	5.0(4.9)	mm
Beam-Beam Parameter	ξ_y	0.0900	0.0875	
Luminosity	L	8×10^{35}		cm^{-2}s^{-1}

Figure 8: The table of machine parameters of SuperKEKB accelerator.

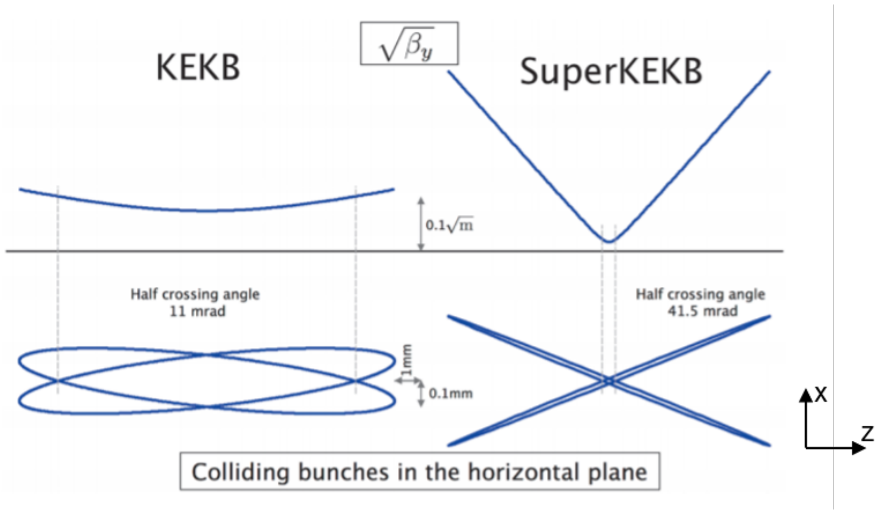


Figure 9: The concept of nano beam scheme. The left is the schematics of beams in Belle, and the right is the nano beam scheme in Belle II. The angle between the bunch of electron and positron is enhanced in nano beam scheme.

3.2 Belle II detector

The Belle II detector is constructed in the vicinity of the collision point in SuperKEKB. It comprises 7 sub-detectors that measure energy, momentum, charged particle trajectory, and decay vertex of the produced particles. The overview and top view of the Belle II detector is shown in Fig. 10 and Fig. 11

3.2.1 Aerogel ring imaging Cherenkov counter (ARICH)

ARICH is located on the end cap of the Belle II detector and serves as an observer of Cherenkov light that comes from the aerogel radiator. The identification of particles is achieved by measuring the Cherenkov angle. A visual representation of the ARICH concept can be found in Fig. 12.

The aerogel radiator is comprised of multiple layers, each with a varying refractive index. This arrangement facilitates the focusing of photons within the aerogel radiator and subsequently eliminates the uncertainty that may arise from the emission point of Cherenkov photons within the radiator. The Cherenkov ring is collected by the Hybrid Avalanche Photon Detector (HAPD), which is positioned at a distance of 20cm from the aerogel radiator.

3.2.2 Central drift chamber (CDC)

The CDC, located exterior to the Silicon Vertex Detector (SVD), which is explained in Sec. 3.2.7, functions as a drift chamber and serves to measure the tracks of charged particles. It is filled with a gas mixture comprising of 50% Helium and 50% C_2H_6. Axial and stereo wires are arranged along the beam direction. As charged particles travel through the CDC, electrons are produced and subsequently collected by the anode wire.

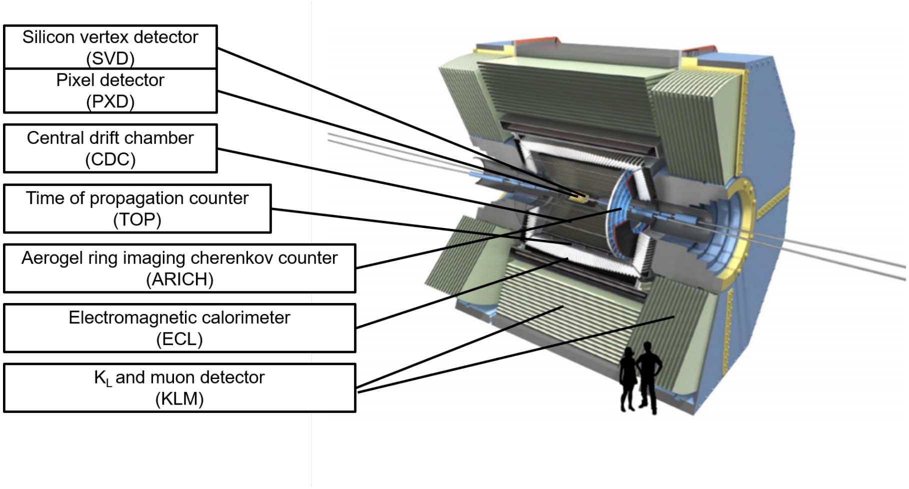


Figure 10: Overview of Belle II detector

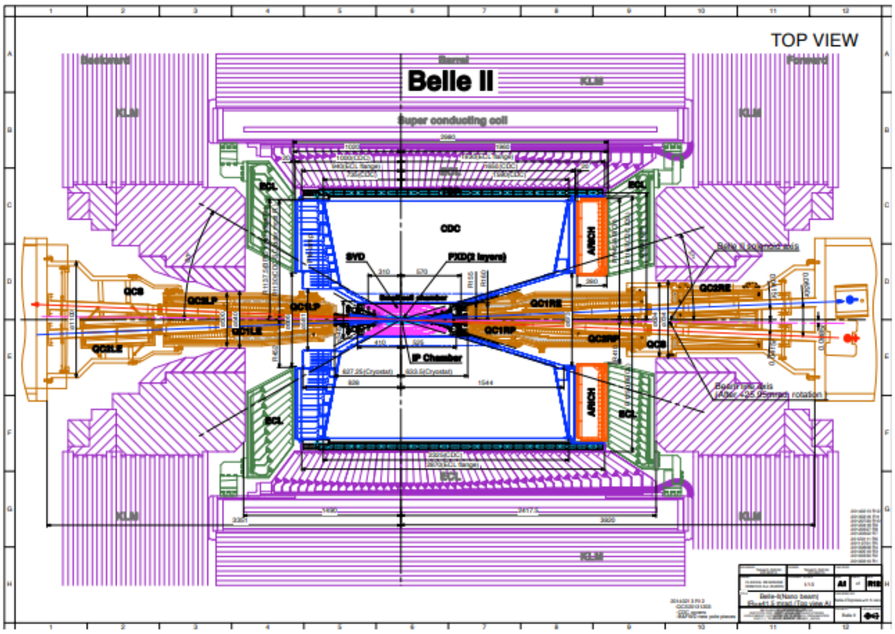


Figure 11: Top View of Belle II detector [6]

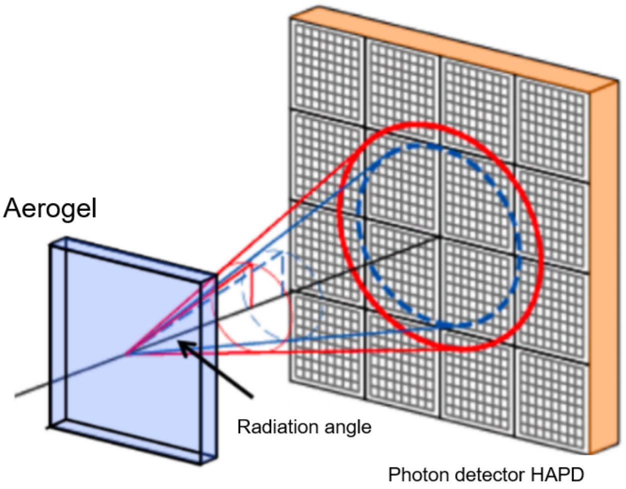


Figure 12: The concept of ARICH, where the radiation angle is dependent upon the mass and velocity of the particle.

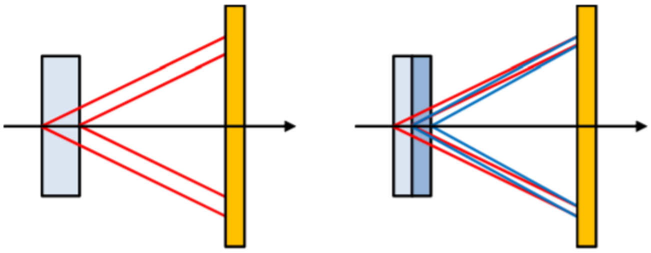


Figure 13: The composition of aerogel material. The left figure depicts a single layer of aerogel, while the right figure shows multiple layers of aerogel with varying refractive indices. The dependence of the radiation angle on the emission point of Cherenkov light is notably diminished in the multi-layer radiator.

In addition to facilitating the reconstruction of particle tracks, CDC also enables particle identification through the analysis of energy loss within its gas volume.

The wire placement configuration has been upgraded in Belle II with the goal of improving position resolution and managing the higher event rate and background. Specifically, the wire density in the inner layer has been increased. The information gathered by CDC is also utilized in the generation of triggers.

3.2.3 Electromagnetic calorimeter (ECL)

The Electromagnetic Calorimeter (ECL) is a detector designed to measure the energy of photons, electrons and hadrons, and is located in both end cap regions and barrel region. The energy and number of clusters detected by ECL are utilized in the generation of triggers.

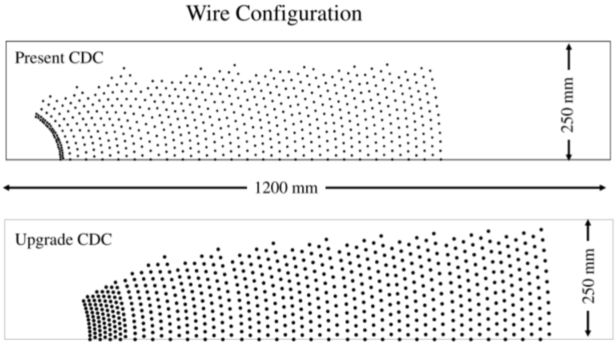


Figure 14: The wire configuration of CDC is illustrated in the provided figures, with the upper diagram corresponding to Belle and the lower diagram to Belle II. Notably, the wire density in Belle II CDC has been increased as compared to Belle.

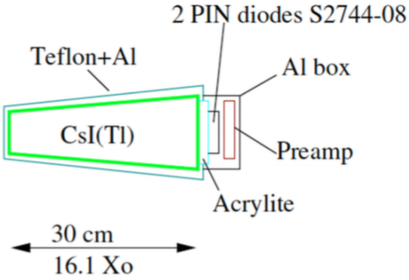


Figure 15: The unit of ECL with 2 pin diodes connected to CsI crystal.

In the Belle II experiment, the photon energy ranges from 20 MeV to 4 GeV. To achieve a high energy resolution, CsI(Tl) is implemented as the scintillation crystal owing to its high light output and short radiation length. Photons, electrons, and positrons deposit their energy in the scintillator and subsequently generate an electromagnetic shower, with the energy of the particles being measured by the scintillation light. ECL is capable of measuring luminosity via the detection of Bhabha scattering process. When combined with momentum information acquired from CDC, ECL can distinguish electrons and positrons from other charged particles by analyzing their respective energies.

The structure of an ECL unit is depicted in Fig. 15. The energy deposited in the crystal is collected by two pin photodiodes located behind the crystal and then transmitted to the electronics. The overall structure of ECL, as demonstrated in Fig. 16, has not been modified since the Belle experiment. However, the readout electronics have been updated in order to handle signal pileup that arises from the higher event rate and increased background.

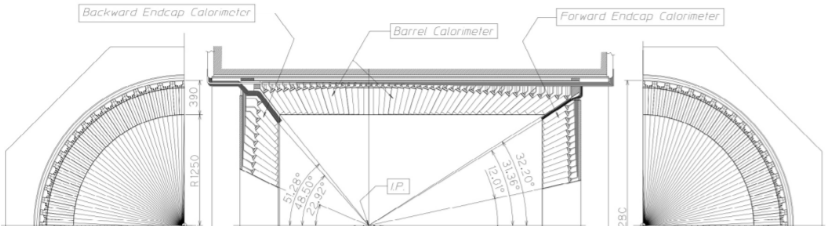


Figure 16: Overall configuration of ECL.

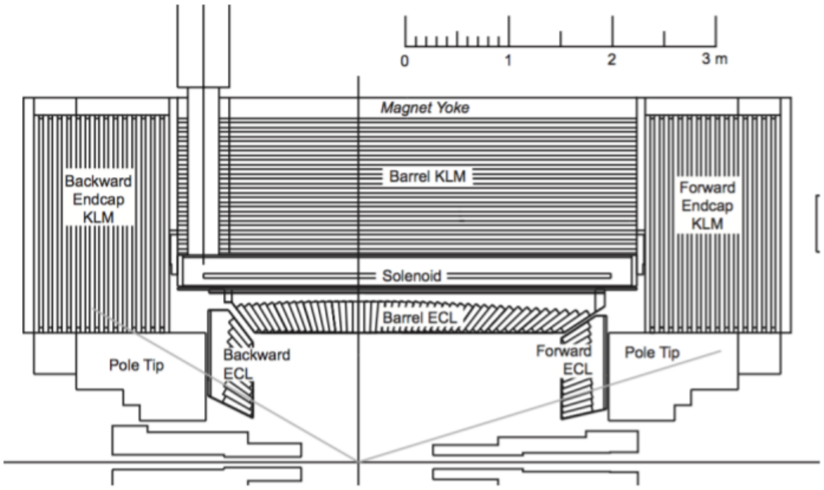


Figure 17: The configuration of KLM in the Belle II detector involving two distinct component, the Barrel KLM and the End-cap KLM respectively.

3.2.4 K$_L$ and muon detector (KLM)

The K_L^0 and μ detector, or KLM, functions to identify K_L^0 and μ particles from other particles. This detector is located externally to the ECL and is partitioned into two distinct regions: the endcap region KLM (EKLM) and the barrel region KLM (BKLM). KLM's coverage spans an angle ranging from 20° to 155° in total, relative to the beam axis. The overall layout of KLM is depicted in Fig. 17.

KLM is comprised of an alternating sandwich of iron plates and detector layers. The interaction of μ with the material occurs via electromagnetic interaction, while charged hadrons such as π experience the effects of strong interaction. Using this property, μ particles are identified through their penetration of the iron plates. On the other hand, K_L particles produce a hadronic shower within the iron plate layer, and are thus identified by observing the resultant particles from hadronic shower.

In the Belle experiment, glass-electrode resistive plate chambers (RPC) serve as the detector layer. However, these chambers possess a notable disadvantage in that they exhibit a large dead time of several seconds due to the long recovery period of the electric field following a discharge. In the Belle II experiment, a significant amount of neutrons generated by background processes within the electromagnetic shower produces a large

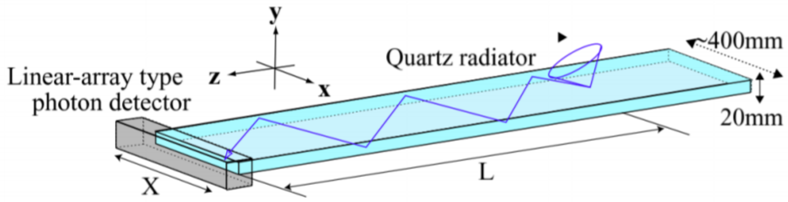


Figure 18: The structure of TOP. Cherenkov lights collected at the photon detector located at the side of the scintillator.

dead time, particularly in the endcap region and the innermost layers of the barrel region, thereby reducing the efficiency of particle detection.

In order to avoid this issue, a detector layer employing scintillator strips and wavelength-shifting fibers was developed and installed specifically for the endcap region and the two innermost layers of the barrel region. This detector layer utilizing scintillator strips has a lower dead time compared to the Belle experiment, while also displaying a heightened capacity to tolerate increased background levels within the Belle II experiment.

3.2.5 Time of propagation counter (TOP)

The Time of Propagation counter (TOP) is located in the barrel region outside the CDC and serves to identify charged particles through the use of Cherenkov light detection. Quartz bars serve as the radiator of Cherenkov light, with 16×2 microchannel plate PMTs (MCP-PMTs) attached at the side of each quartz bar, as depicted in Fig. 18.

The Cherenkov light produced is transmitted to the end of the scintillator, where it is detected by the MCP-PMT. Given that the angle of the Cherenkov light varies between different particle types, the time interval required for the Cherenkov light to reach the MCP-PMT differs across particles. Consequently, particle types are identified through the measurement of time-of-flight (TOF). The difference in propagation time amounts to approximately 200 ps, necessitating the MCP-PMT to possess a time resolution of less than 50 ps.

3.2.6 Pixel detector(PXD)

The Pixel Detector, positioned as the innermost sub-detector and covering an angle of $17^\circ < \theta < 150^\circ$ along the direction of the beam, serves to measure the tracks of charged particles and precisely reconstruct decay vertices, in combination with SVD.

The Pixel sensor is based on a pixelated DEPleted p-channel Field Effect Transistor (DEFPET) and comprises two layers of sensors surrounding the interaction point. The structure of DEFPET is shown in Fig. 19. When charged particles enter the depletion layer, the electron-hole pair produced is accumulated to the internal gate in Fig. 19 and then read out. Due to the large number of channels, it is impossible to read out all channels simultaneously. Thus, data is read out for each line of the DEFPET matrix by controlling the gate voltage of FET. It takes 100 ns to read out and clear the charge per

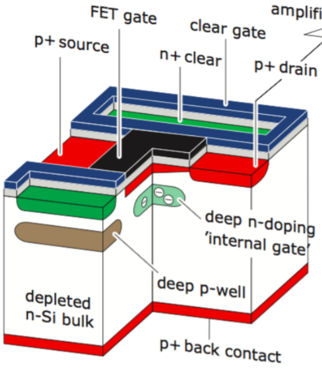


Figure 19: The structure of DEFPET adopted for PXD.

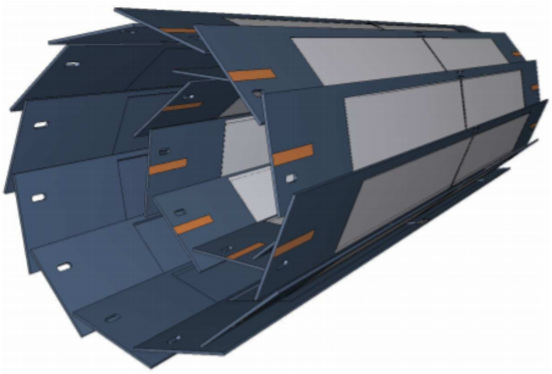


Figure 20: The overview of PXD, which consists of two layers of sensors.

line, and 20 μs to read out all the pixel channels. As the full data size of the PXD reaches 30 GB/s under 30 kHz trigger rate, it is not feasible to read out the entire PXD data. Therefore, the tracks of particles are reconstructed and extrapolated to the PXD region, and only the PXD data within the region of interest is read out in order to reduce the data volume of the PXD.

3.2.7 Silicon vertex detector(SVD)

The Silicon Vertex Detector (SVD) is positioned externally to the PXD, and is utilized to reconstruct the trajectory of charged particles. Comprising four layers of double-sided silicon strip detectors (DSSDs), as illustrated in Fig. 21, the SVD is configured with silicon strips on both sides of the sensor in such a way that the n-side strip is perpendicular to the beam direction. When charged particles pass through the n bulk, electron-hole pairs are generated, and the electron and hole are collected in the n-side and p-side, respectively, and then read out.

In order to achieve fast data readout, the SVD uses ASIC called APV25. In the Belle II experiment, we expect a 30 kHz trigger rate at a target luminosity of $\mathcal{L} = 6 \times 10^{35}$.

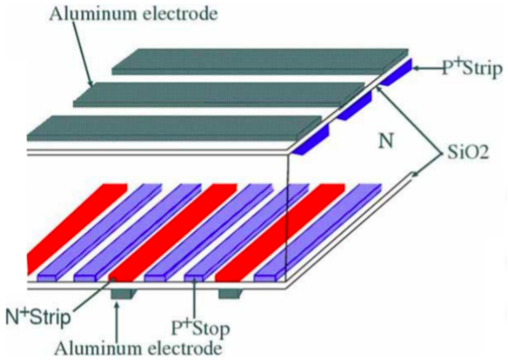


Figure 21: The structure of DSSD. Silicon strips are put on both sides of N bulk orthogonally.

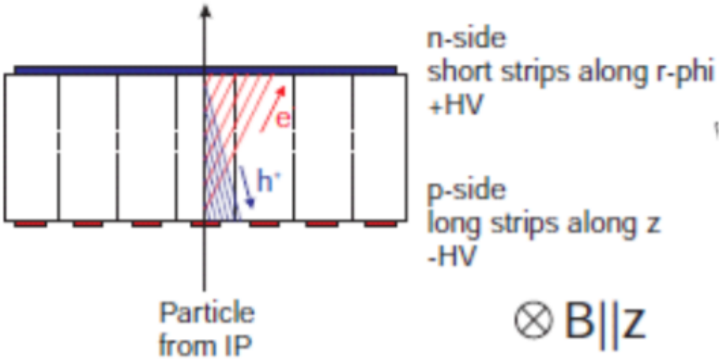


Figure 22: The detection mechanism of DSSD. The produced electron and hole are collected to silicon strips and read out by electrodes.

However, due to the limited buffer capacity of the APV25, the SVD is expected to have a 3.4% dead time at the 30 kHz trigger rate.

3.2.8 Trigger system (TRG)

The trigger system (TRG) fulfills the role of issuing triggers. In Belle II experiment, it is necessary to acquire data at a high trigger rate of up to 30 kHz due to the high luminosity of SuperKEKB. Consequently, the trigger system is required to achieve a high efficiency for physics event triggering, while simultaneously minimizing the number of events originating from beam background. The expected cross section and trigger rate for the dominant process, given a target luminosity of $\mathcal{L} = 6 \times 10^{35}\text{cm}^{-2}\text{s}^{-1}$, are presented in Table. 1. The trigger system is subject to the following requirements:

- Maximum trigger rate of 30kHz
- Trigger distribution latency under 5μs

Table 1: Table of expected trigger rate for each process under SuperKEKB target luminosity [35]. Due to the 30kHz upper limit of trigger rate, the trigger rate from beam background signals must be below 15kHz.

Process	σ (nb)	Rate (Hz)
$\Upsilon(4S)$	1.2	960
Continuum	2.8	2200
$\mu^+\mu_-$	0.8	640
$\tau^+\tau_-$	0.8	640
Bhabha	44	350
$\gamma\gamma$	2.4	19
Two photon	12	10000
Total	67	15000

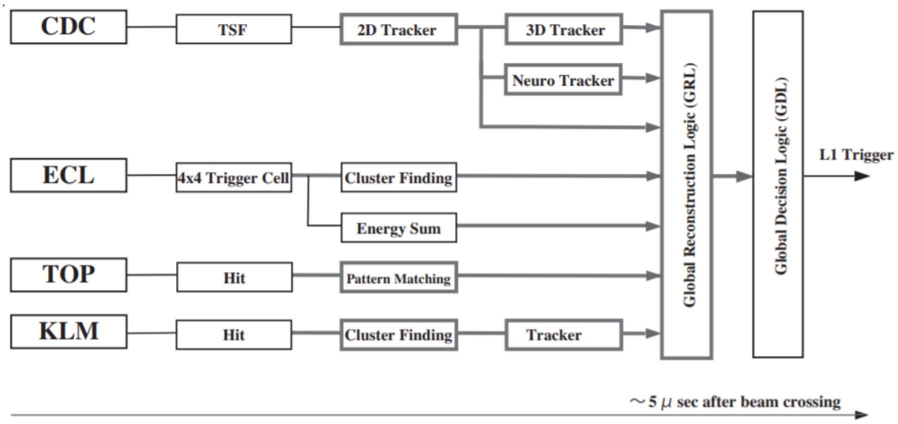


Figure 23: The overview of trigger system [22]. Level 1 trigger is generated using information from four sub-trigger systems.

- The timing resolution under 10ns

Fig. 23 provides an overview of the trigger system, which comprises four distinct sub-triggers: CDC, ECL, TOP, and KLM. Each sub-trigger system operates independently, and the information gathered from these four detectors is transmitted to the Global Decision Logic (GDL), where the final trigger signal determination takes place. The primary trigger decision for physics events relies on the sub-triggers from CDC and ECL. The TOP sub-trigger is primarily used to achieve better timing resolution for triggering, while the muon identification obtained from the KLM sub-trigger is employed in conjunction with the CDC and ECL sub-triggers.

3.2.9 Data acquisition system (DAQ)

In Belle II experiment, the SuperKEKB accelerator exhibits an instantaneous luminosity that is 30 times higher than that of the KEKB accelerator in Belle. Due to the upgraded detectors in Belle II, the data size per event is expected to exceed 1 MB/s in total, given

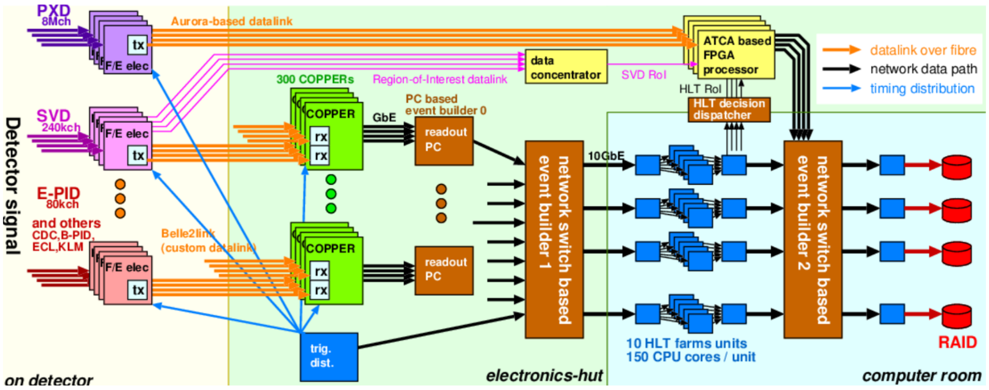


Figure 24: The overview of data acquisition system [21]

the target luminosity of $\mathcal{L} = 6 \times 10^{35}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. With an expected maximum trigger rate of 30 kHz, the online data processing necessitates handling a data flow of 30 GB/s.

The overall structure of the Belle II data acquisition system is illustrated in Fig. 24. A common clock signal, derived from the 508 MHz radio frequency of SuperKEKB, is utilized across all modules in the data acquisition system. This 508 MHz clock signal is divided by four, generating a common clock of 127 MHz. The trigger signal, originating from the GDL, is then distributed to all front-end electronics located near the Belle II detector. The acquired data is digitized by these front-end electronics and transmitted to modules which is responsible for receiving data from the front-end electronics, performing data formatting, and sending the data to a readout PC. A module called Common Pipelined Platform for Electronics Readout (COPPER) was used to be adopted for this purpose, but these COPPERs are currently replaced with a PCI-express-based high-speed readout modules known as PCIe40 [40]. Optical fiber connections are employed to link the front-end electronics with the PCIe40 modules, and the data from sub-detectors is transmitted using an original communication protocol named Belle2link.

Each readout PC undertakes partial event reconstruction using the data collected by the readout PC. The partially reconstructed data is then transmitted to the High-Level Trigger (HLT). The HLT is responsible for carrying out full event reconstruction and selecting which events to record. When combined with the PXD data, the final event reconstruction is accomplished, and the event data is stored in the storage system.

3.2.10 High level trigger

The High-Level Trigger (HLT) is a system designed to reconstruct events and generate software triggers. By employing software triggers, only essential events are selected, thereby reducing the data size that needs to be recorded in the storage system. Measured data, excluding the PXD, is transmitted to the HLT from the a network switch through a 10 Gbit connection. Each HLT node consists of a multi-core PC, and event data is stored in a sizable ring buffer and processed in parallel by multiple CPUs.

Within the HLT, particle tracks are reconstructed, enabling the selection of specific regions in the PXD to be read. These particle tracks are extrapolated to the PXD, defining regions of interest (ROIs). The ROI information is then transmitted to the PXD readout system. Only the PXD data within the ROIs is subsequently sent to EB2. EB2 receives data from both the HLT and the PXD readout system, facilitating event reconstruction using the entire Belle II detector. The collected data is ultimately transmitted to the storage system.

4 Dalitz-plot analysis

This section provides the motivation for the analysis of time-dependent CP violating parameters using Dalitz-plot technique. Additionally, the formalism of decay amplitude using Dalitz plane and square Dalitz plane are delivered.

4.1 Motivation and target of Dalitz-plot analysis

As explained in the section 2, the decay channel $B^0 \to \phi K_S^0$ is suited for the search of new physics through the measurement of CKM unitarity angle ϕ_1. However, the measurement of CP violation parameters for $B^0 \to \phi(\to K^+K^-)K_S^0$ is intricate due to the interference arising from other modes that decay to the same $B^0 \to K^+K^-K_S^0$ final state. To account for the effect of interference among the decay modes, the Dalitz-plot technique is employed in this analysis. Moreover, the utilization of Dalitz-plot technique allows us to measure ϕ_1 instead of $\sin 2\phi_1$ which is typically measured by a quasi-two-body approach, thereby solving the ambiguity between ϕ_1 and $\pi - \phi_1$.

The resonant and non-resonant decay channels to be considered in this analysis are presented in Table. 2, along with the respective parameters for each channel. The final state $K^+K^-K_S^0$ does not generally possess a definite CP eigenvalue, as the eigenvalue of the $K^+K^-K_S^0$ three-body system depends on the angular momentum l of the K^+K^- system. By employing the Dalitz-plot technique, the interference arising from CP-odd and CP-even final states can be effectively addressed, thereby enabling a more precise determination of ϕ_1 without the dilution caused by the final state's mixture of CP-odd and CP-even eigenstates.

Resonance	Parameters (MeV)	Line Shape
$\phi(1020)$	$M = 1019.445 \pm 0.020$ $\Gamma = 4.26 \pm 0.020$	RBW
$f_0(980)$	$M = 990 \pm 20$ $g_{\pi\pi} = 0.165 \pm 0.018$ $g_{KK}/g_{\pi\pi} = 4.21 \pm 0.33$	Flatt$\acute{e}$
f_X	$M = 1506 \pm 6$ $\Gamma = 112 \pm 9$	RBW
χ_{c_0}	$M = 3414.71 \pm 0.20$ $\Gamma = 10.8 \pm 0.6$	RBW
$(K^+K^-)_{NR}$	$\alpha = 0.14$	$e^{-\alpha s_0}$
$(K^+K_S^0)_{NR}$	$\alpha = 0.14$	$e^{-\alpha s_+}$
$(K^-K_S^0)_{NR}$	$\alpha = 0.14$	$e^{-\alpha s_-}$

Table 2: The list of all 7 decay channels included in this analysis and its physics parameters. The non-resonant component is divided into three component. The parameters for f_0 resonance are cited from Ref. [3], and the rest of parameters are cited from Ref. [31].

4.2 Modeling for decay amplitude

The expression for the decay rate of the $B^0 \to K^+K^-K^0_S$ three-body decay, involving pseudo-scalar mesons, can be calculated as follows:

$$\Gamma(B^0 \to K^+K^-K^0_S) = \frac{1}{(2\pi)^3}\frac{|A|^2}{32m_{B^0}}ds_+ds_-. \tag{60}$$

Here, A denotes the Lorentz invariant decay amplitude. The quantity s_+ and s_- represent the squared invariant masses of the $K^0_SK^+$ and $K^0_SK^-$ systems, respectively. To describe the amplitude A, we employ an isobar approximation, where A is modeled as the coherent sum of amplitudes for each decay channel, as expressed by:

$$A = \sum_i a'_i F_i \tag{61}$$

$$\bar{A} = \sum_i \bar{a}'_i \bar{F}_i \tag{62}$$

The amplitude F_i is dependent on the Dalitz plane and incorporates only strong dynamics, while a'_i is a complex coefficient including the influence of weak phases. The index i denotes each decay channel listed in Table. 2.

In the analysis of $B^0 \to K^+K^-K^0_S$, the Dalitz-plane dependent amplitude for the resonant decay process $B \to rK^0_S$, where r represents a resonance, can be expressed using four terms as follows:

$$F_i(L, s_+, s_-) = Z^L(\vec{p}, \vec{q}) \times X^B_L(|\vec{p}^*|) \times X^r_L(|\vec{q}|) \times R_r(s_+, s_-), \tag{63}$$

Here, L denotes the orbital angular momentum between K^0_S and the resonance. $\vec{p}^*$ represents the momentum of K^0_S in the rest frame of the B meson, while $\vec{p}$ and $\vec{q}$ represent the momenta of K^0_S and K^+, respectively, in the rest frame of the resonance. The terms Z, X, and R correspond to the angular distribution, the Blatt-Weisskopf barrier factor[10], and the line shape of the resonance, respectively.

The term Z represents the angular distribution, which is dependent on the orbital angular momentum L of the resonance r. It can be expressed as

$$Z^0(\vec{p}, \vec{q}) = 1, \tag{64}$$

$$Z^1(\vec{p}, \vec{q}) = -4\vec{p} \cdot \vec{q} \tag{65}$$

X^B_L and X^r_L represent the Blatt-Weisskopf barrier factors[10] for the production and decay of the resonance, respectively. The maximum angular momentum L is limited by $\vec{q}$ in a strong decay process, implying that slowly decaying particles have difficulty in generating angular momentum while still conserving the spin of the resonance. The Blatt-Weisskopf barrier factor is introduced to account for this effect and can be written as:

$$X_0 = 1 \tag{66}$$

$$X_1 = \sqrt{\frac{1+z_0}{1+z}} \tag{67}$$

Here, $z = (|\vec{q}|d)^2$, $z_0 = (|\vec{q_0}|d)^2$, and $\vec{q_0}$ represents the value of $\vec{q}$ when the invariant mass $m_{K^+K^-}$ is equal to the resonance mass m_r. The parameter d corresponds to the meson radius, which is fixed at 1 fm in this analysis.

The term R_r represents the form factor of the resonance, which varies for each decay channel. In this analysis, the Relativistic Breit-Wigner function [38] is employed for the $B \to \phi K_S^0$, $B \to f_X K_S^0$, and $B \to \chi_{c_0} K_S^0$ decay channels. The Breit-Wigner function is parameterized as:

$$RBW(s) = \frac{1}{M_0^2 - s - iM_0\Gamma(s)} \tag{68}$$

Here, M_0 denotes the nominal mass of the resonance, and s represents the invariant mass of the K^+K^- system. The mass-dependent width $\Gamma(s)$ is given by:

$$\Gamma(s) = \Gamma_0 \left(\frac{|\vec{q}|}{|\vec{q_0}|}\right)^{2L+1} \frac{m_0}{m} X_L^2(z), \tag{69}$$

where $m = \sqrt{s}$ denotes the invariant mass of K^+K^- system, and Γ_0 corresponds to the nominal width of the resonance.

The Flatté function[16] is used exclusively for the $B \to f_0 K_S^0$ decay channel, where the kinematic threshold for the creation of the resonance from the K^+K^- pair is very close to the mass of the resonance $f_0(890)$. In this analysis, the following parametrization is adopted to describe the f_0 lineshape:

$$Flatt\acute{e}(s) = \frac{1}{M_0^2 - s - i\left(g_{\pi\pi}\rho_{\pi\pi} + g_{KK}\rho_{KK}\right)} \tag{70}$$

$$\rho_{\pi\pi} = 2\frac{\sqrt{1 - 4m_\pi^2/s}}{\sqrt{s}} \tag{71}$$

$$\rho_{KK} = 2\frac{\sqrt{1 - 4m_K^2/s}}{\sqrt{s}}. \tag{72}$$

Here, $g_{\pi\pi}$ and g_{KK} denote the coupling constants. For this analysis, the values of these coupling constants are fixed to $g_{\pi\pi} = 0.165 \pm 0.018$ GeV$^2/c^4$ and $g_{KK}/g_{\pi\pi} = 4.21 \pm 0.33$ [3].

For the non-resonant component, an empirical exponentially decaying model is adopted, which parameterize the line-shape of the non-resonant component in terms of three components as

$$F_{NR} = F_{NR_0} + F_{NR_+} + F_{NR_-} \tag{73}$$

$$F_{NR_0} = e^{-\alpha s} \tag{74}$$

$$F_{NR_+} = e^{-\alpha s_+} \tag{75}$$

$$F_{NR_-} = e^{-\alpha s_-}. \tag{76}$$

The parameter α is fixed to 0.14, which is cited from a previous analysis[31].

4.3 Parametrization of the isobar model and CP asymmetry

The coefficients for each decay channel in the isobar approximation are parametrized as:

$$a_i' = a_i(1 + c_i)e^{i(b_i+d_i)}, \tag{77}$$

$$\bar{a}_i' = a_i(1 - c_i)e^{i(b_i-d_i)}. \tag{78}$$

In this notation, a_i, b_i, c_i, and d_i are real numbers representing the amplitude, phase, direct CP asymmetry, and mixing-induced CP asymmetry for each decay channel, respectively.

Using these parameters, the quasi-two-body expression of CP asymmetry for each decay channel can be calculated from the Dalitz parameters c_i and d_i. The direct CP asymmetry is given by:

$$A^i_{CP} = \frac{|\bar{a_i}'|^2 - |a_i'|^2}{|\bar{a_i}'|^2 + |a_i'|^2} = \frac{-2c_i}{1 + c_i^2}, \tag{79}$$

Furthermore, the CKM angle ϕ_1^i for each decay channel is determined as:

$$\phi_1(i) = \frac{arg(\bar{a_i}' a_i'^*)}{2} = d_i. \tag{80}$$

Here, the parameter d_i, which is one of the fitting parameters, is directly connected to the CKM unitarity angle. Along with the phase term b_i, the event distribution exhibits asymmetry between d_i and $90° - d_i$, and thus resolving the two-fold ambiguity in the quasi-two-body approach and enable us to measure ϕ_1 diectly.

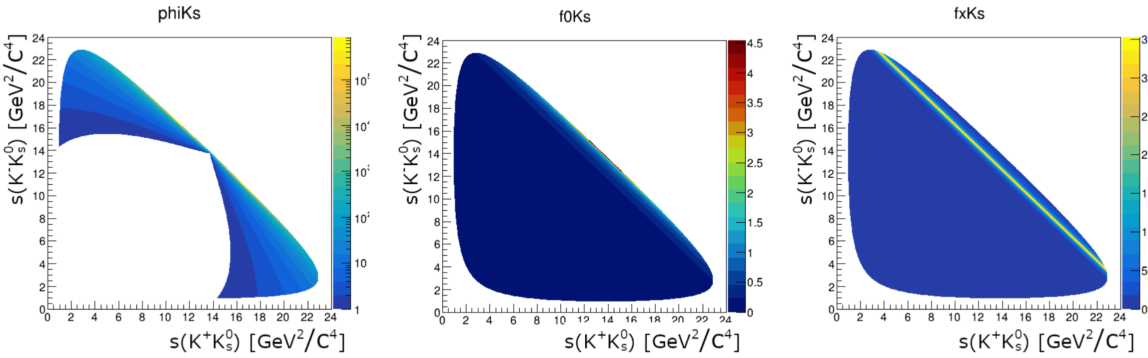
4.4 Dalitz Plot and square Dalitz Plot

The kinematics of a three-body decay can be described using two variables, as expressed in Eq. 60. Therefore, the kinematics of a three-body decay can be represented by the distribution of events over a two-dimensional plane. The parametrization that uses two independent variables as the invariant masses (s_+, s_-), is referred to as the Dalitz parametrization. By utilizing Eq. 63, the squared decay amplitudes for each channel, $|F_i(s_+, s_-)|^2$, can be visualized on the Dalitz plane, as illustrated in Fig. 25.

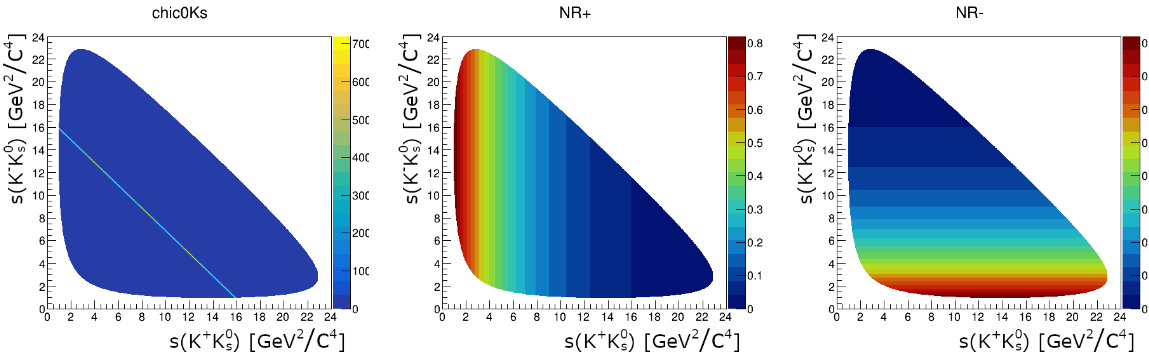
As depicted in Fig. 25, the event distribution for resonances such as ϕK_S^0 or $f_0 K_S^0$ exhibits a sharp peak near the boundaries of the three-body decay kinematics. Both signal and background events tend to concentrate in regions close to the kinematic boundary on the Dalitz plane. This characteristic poses challenges in modeling the background distribution and accounting for the reconstruction efficiency accurately. To address these challenges and enhance the regions of interest where $B^0 \to \phi K_S^0$ events populate, the square Dalitz plane [1] (m', θ') is employed to describe the event distribution in this analysis. The parametrization of the square Dalitz plane is given by:

$$m' \equiv \frac{1}{\pi}\arccos\left(2\frac{m - m_0^{min}}{m_0^{max} - m_0^{min}} - 1\right) \tag{81}$$

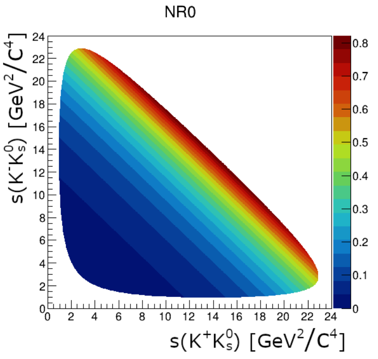
$$\theta' \equiv \frac{1}{\pi}\theta_0 \tag{82}$$



(a) ϕK_S^0 (b) $f_0 K_S^0$ (c) $f_X K_S^0$



(d) $\chi_{c0} K_S^0$ (e) $K^-(K^+K_S^0)_{NR}$ (f) $K^+(K^-K_S^0)_{NR}$



(g) $K_S^0(K^+K^-)_{NR}$

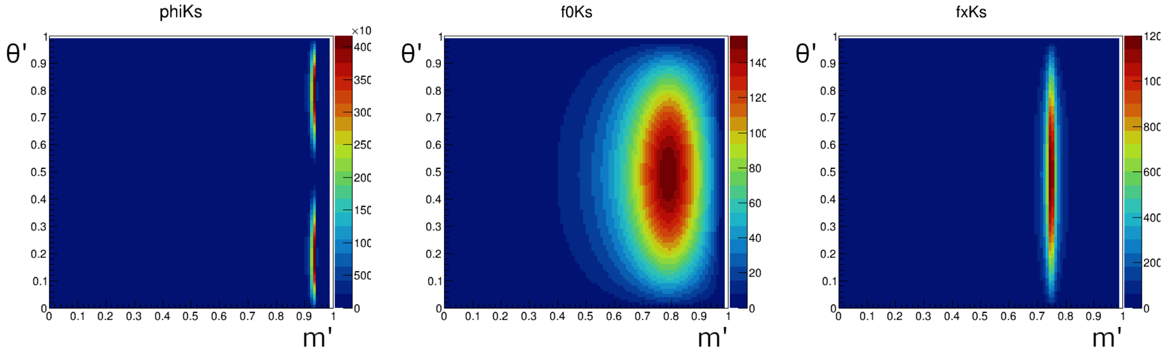
Figure 25: Distribution of $|F_i(m', \theta')|^2$ over Dalitz plane for each decay channels. Horizontal and vertical axes represents the squared invariant mass between K_S^0-K^+ and K_S^0-K^-, and the z axis represents $|F_i(s_+, s_-)|^2$.

Here, $m_0^{max} = m_{B^0} - m_{K_S^0}$ and $m_0^{min} = 2m_{K^\pm}$ represent the maximum and minimum kinematically allowed invariant masses of the K^+K^- system, respectively. The parameter θ_0 corresponds to the angle between K^+ and K_S^0 in the rest frame of K^+ and K^-.

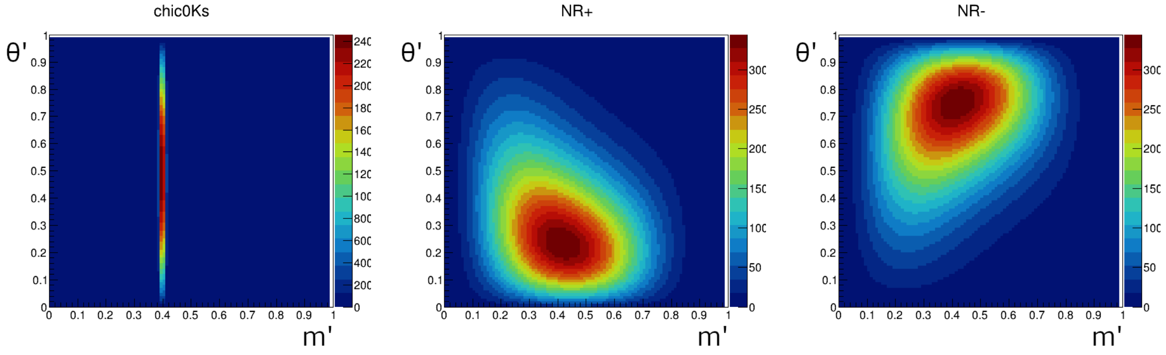
The range of both m' and θ' is defined to be within the interval $[0, 1]$. The Jacobian of the transformation, denoted as $\det J$, can be calculated as follows:

$$\det J = 4|\boldsymbol{p}^*_{\boldsymbol{K^+}}||\boldsymbol{p}^*_{\boldsymbol{K_S^0}}|\frac{\partial m_0}{\partial m'}\frac{\partial \cos\theta_0}{\partial \theta'} \tag{83}$$

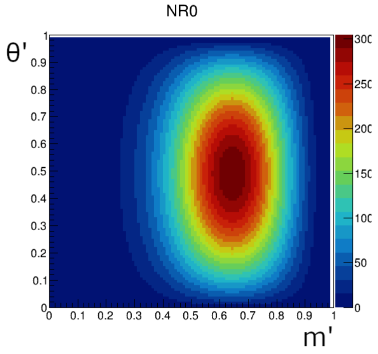
Here, $\boldsymbol{p}^*_{\boldsymbol{K^+}}$ and $\boldsymbol{p}^*_{\boldsymbol{K_S^0}}$ represent the momenta of K^+ and K_S^0 in the rest frame of K^+ and K^-. The distribution $|\det J||F_i(m', \theta')|^2$ is depicted in Fig. 26. By utilizing the square Dalitz plot to analyze the event distribution, the sharpness of the distribution is broadened, enabling a more precise handling of histogram probability density functions (PDFs) over the square Dalitz plane. Additionally, it make it easier to calculate the normalization of the PDF over the entire kinematically allowed region.



(a) ϕK_S^0 (b) $f_0 K_S^0$ (c) $f_X K_S^0$



(d) $\chi_{c0} K_S^0$ (e) $K^-(K^+K_S^0)_{NR}$ (f) $K^+(K^-K_S^0)_{NR}$



(g) $K_S^0(K^+K^-)_{NR}$

Figure 26: Distribution of $|detJ||F_i(m',\theta')|^2$ over square Dalitz plane for each decay channels. Horizontal and vertical axes represents the square Dalitz plot parameters m' and θ', and the z axis represents $|detJ||F_i(m',\theta')|^2$.

5 Event reconstruction and selection

In this section, we first deliver explanation for the flavor tagging and the measurement of decay time difference Δt, then provides strategy for event reconstruction and selection.

5.1 Flavor Tagging

The flavor tagger plays a crucial role in determining the flavor of the tag-side B meson, which is necessary to establish the flavor of the fully-reconstructed CP-side B meson. In the process of generating $B\bar{B}$ pairs at the $\Upsilon(4S)$ resonance, no additional particles are expected to present. Therefore, tracks and clusters that are not used in the full reconstruction of the CP-side B meson can be attributed to the tag-side B meson. Certain decay channels of B mesons exhibit flavor-specific final states, such as $\bar{B}^0 \to D^{*+}l^-\bar{\nu}_l$, where the information from the tracks and clusters on the tag side directly provides the flavor of the tag-side B meson.

The flavor tagger serves as a tool that utilizes an inclusive technique with a multivariate method to extract information about the tag-side flavor from flavor-specific signatures. It returns two values: q and r. The value of q is +1 or -1, indicating whether the flavor of the tag-side B meson is B^0 or $\bar{B}^0$, respectively. The parameter r represents the dilution factor, which is defined as $r = 1-2w$, where w is the wrong tag fraction. The wrong tag fraction corresponds to the fraction of incorrectly identified events among the total number of tagged events. A value of $r = 0$ implies that no flavor information is available for the tag-side B meson, while $r = 1$ signifies that the flavor determination is unambiguous.

The flavor tagging parameters are determined separately for each of the 7 bins defined by the values of the flavor tagger output $|qr|$. These 7 bins have edges located at:

$$|qr| = [0, 0.1, 0.25, 0.45, 0.6, 0.725, 0.875, 1]. \tag{84}$$

For each bin, the performance of flavor tagger is then represented as

$$\begin{aligned} w &= \frac{1}{2}(w_{B^0_{\rm tag}} + w_{\bar{B}^0{}_{\rm tag}}) \\ \Delta w &= (w_{B^0_{\rm tag}} - w_{\bar{B}^0{}_{\rm tag}}) \\ \mu &= \frac{\epsilon_{B^0_{\rm tag}} - \epsilon_{\bar{B}^0{}_{\rm tag}}}{\epsilon_{B^0_{\rm tag}} + \epsilon_{\bar{B}^0{}_{\rm tag}}} \end{aligned} \tag{85}$$

where $w_{B^0_{\rm tag}}$ and $w_{\bar{B}^0{}_{\rm tag}}$ are the wrong tag fraction and ϵ is the tagging efficiency defined as the fraction of events to which the flavor tagger can assign any flavor tag over the number of total events.

The flavor tagging parameters have been calibrated using 361.6 fb^{-1} of real data for the $B^0 \to D^{(*)-}\pi^-$ decay channels [20]. The calibration results, which determine the performance of the flavor tagger, are presented in Table. 3. These calibrated parameters are used in the construction of the CP asymmetry fitting PDF.

5.2 Δt measurement

The concept of the measurement of decay time difference Δt is shown in Fig. 27. The

qr-bin	$w_{B^0_{\mathrm{tag}}}$	$w_{\bar{B}^0{}_{\mathrm{tag}}}$	μ
[0, 0.1]	(0.4770 ± 0.0077)	(0.4838 ± 0.0076)	(− 0.0240 ± 0.0122)
[0.1, 0.25]	(0.4430 ± 0.0076)	(0.4049 ± 0.0078)	(0.0138 ± 0.0123)
[0.25, 0.45]	(0.3315 ± 0.0072)	(0.3504 ± 0.0073)	(− 0.0115 ± 0.0118)
[0.45, 0.6]	(0.2327 ± 0.0073)	(0.2397 ± 0.0076)	(0.0088 ± 0.0126)
[0.6, 0.725]	(0.1774 ± 0.0073)	(0.1577 ± 0.0077)	(0.0368 ± 0.0135)
[0.725, 0.875]	(0.1070 ± 0.0068)	(0.1075 ± 0.0068)	(− 0.0202 ± 0.0133)
[0.875, 1]	(0.0286 ± 0.0042)	(0.0262 ± 0.0042)	(− 0.0123 ± 0.0108)

Table 3: The parameters of flavor tagger calibrated for $B^0 \to D^{(*)-}\pi^-$ decay channels with 361.6fb^{-1} real data[20].

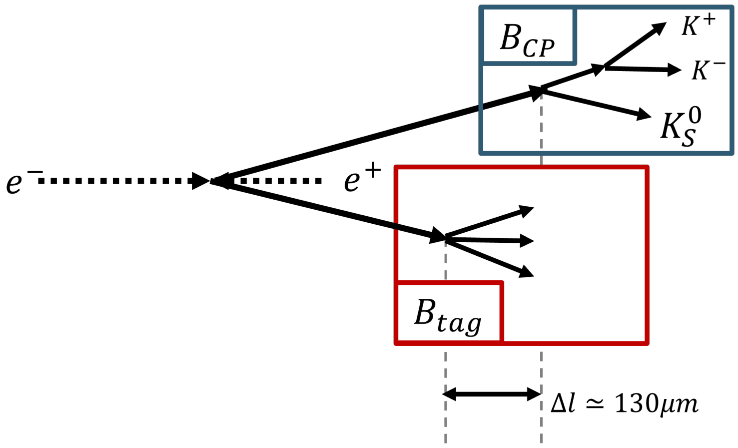


Figure 27: The concept of Δt measurement in Belle II experiment, where Δt is measured through the difference of the position of B meson decay vertices.

center-of-mass system of B meson pairs are boosted along beam axes with Lorentz boost factor $\beta\gamma = 0.286$ due to the asymmetric energy between electrons and positrons. This enables us to measure the difference of B meson decay time, which is typically around 1 ps and cannot be observed directly, through the difference in the decay vertex position of each B meson. A $B\bar{B}$ pair created by $\Upsilon(4S)$ decay is almost at rest in the center-of-mass frame, and thus we can approximate the boost factor of each B meson to be $\beta\gamma = 0.286$ as same as $B\bar{B}$ system. Then the Δt can be measured as

$$\Delta t = t_{\mathrm{CP}} - t_{\mathrm{tag}} = \frac{l_{\mathrm{CP}} - l_{\mathrm{tag}}}{\beta\gamma c} \tag{86}$$

where t_{CP} and t_{tag} denotes the decay time of CP-side and tag-side B meson, and l_{CP} and l_{tag} denotes the decay vertex position of CP-side B meson and tag-side B meson projected to the direction of boost vector.

5.3 Data set used for analysis

For the measurement of CP violating parameters in $B \to K^+K^-K^0_S$ decays, the electron-positron collision events at $\Upsilon(4S)$ resonance are collected by Belle II experiment. The collected events corresponds to integrated luminosity $361.6 \pm 3.9 \mathrm{fb}^{-1}$[23], and contains $(387 \pm 4) \times 10^6$ $B\bar{B}$ pairs. For analysis using Monte Carlo simulation, events corresponding to $1\,\mathrm{ab}^{-1}$ are generated for each event type of $u\bar{u}, d\bar{d}, s\bar{s}, c\bar{c}, B^+B^-$ and $B^0\bar{B}^0$ decays.

5.4 Event reconstruction and selection

In this section, the event reconstruction and selection criteria are explained. The list of the cut criteria is shown in Table. 7. In the analysis we are only concerned with hadronic events, and thus events originating from Bhabha scattering need to be vetoed. An event is identified as a Bhabha event by satisfying all the following condition:

- The number of good tracks with $p_t > 0.2$ GeV/c, $|d_0| < 2$ cm and $|z_0| < 4$ cm is two or more, where p_t is the transverse momentum of a track, d_0 is the distance between a track and interaction point in $r - \phi$ plane, and z_0 is the z coordinate of the closest point of a track against interaction point.
- The maximum opening angle between two charged tracks in the CMS frame of e^+e^- is larger than 2.88 radian.
- The number of tracks identified as electron is one or more.
- The number of tracks with $p^*/E_{\mathrm{beam}} > 0.35$, where p^* and E_{beam} are the momentum of a track and beam energy in the CMS frame of e^+e^-, is two or more.
- The total energy of ECL clusters is more than 4 GeV.

The Bhabha events satisfying those conditions are rejected, and in addition to that we require existence of three or more good tracks in order to identify a event to be hadronic event.

$K^{\pm}$ selection

For $K^{\pm}$ selection, the selection with $|d_0| < 0.5$cm and $|z_0| < 2.0$cm are applied. For particle identification, tight requirements are set for $K^{\pm}$ as kaonID > 0.6, where kaonID is the particle identification for each particle to be $K^{\pm}$, in order to reject possible peaking background from $B\bar{B}$ background due to mis-identification of $K^{\pm}$ from ohter particles, especially $\pi^{\pm}$. The kaonID is defined as

$$\mathrm{kaonID} = \frac{\mathcal{L}_K}{\sum_{i=e,\mu,\pi,K,\mathrm{proton},\mathrm{deuteron}} \mathcal{L}_i}, \tag{87}$$

where $\mathcal{L}_i$ is the likelihood of a particle to be each charged hadron and lepton [6]. The $K^{\pm}$ selection efficiency against the mis-identification rate of $\pi^{\pm}$ to $K^{\pm}$ is evaluated using MC dataset as shown in Fig. 28. For the cut with kaonID > 0.6, we have 76.8% selection efficiency with 1.3% mis-identification rate of $\pi^{\pm}$ to $K^{\pm}$.

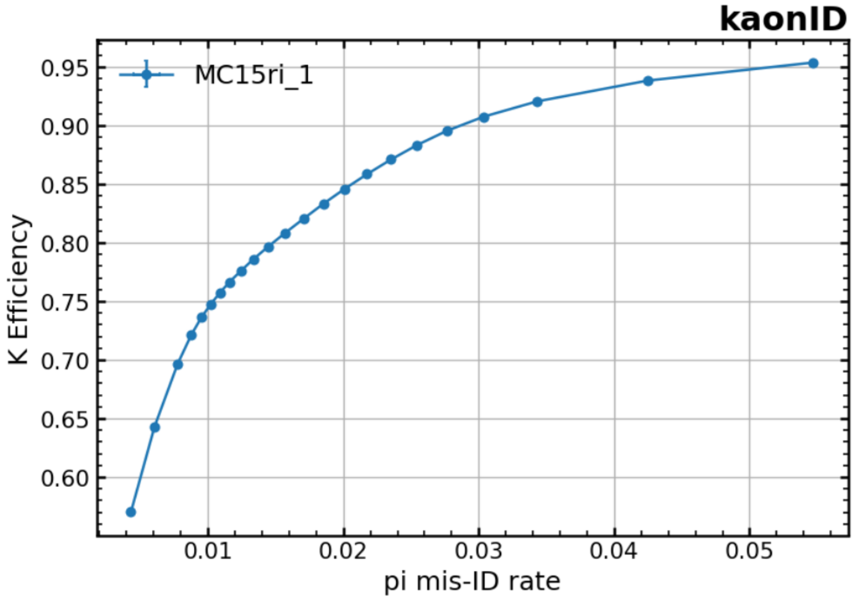


Figure 28: The $K^{\pm}$ selection efficiency against the mis-identification rate of $\pi^{\pm}$ to $K^{\pm}$ evaluated with MC simulation.

K_S^0 reconstruction, selection

K_S^0 is reconstructed from $K_S^0 \to \pi^+\pi^-$ decay with conditions of $0.485 < m_{K_S^0} < 0.515\,\mathrm{GeV}/c^2$ and the flight length of K_S^0 to be larger than 0.05 cm. Additional cut for K_S^0 candidates is applied based on a multivariate classification technique known as FastBDT[25]. In order to achieve higher performance for selecting K_S^0, mva-based selection method called KsSelector is developed. The KsSelector consists of two FastBDT: V0Selector and LambdaVeto. The V0Selector FastBDT discriminate long-lived true K_S^0 from mis-reconstructed fake K_S^0 which tend to originate from interaction point. The LambdaVeto FastBDT discriminate true K_S^0 from Λ using variables including different mass-hypothesis for K_S^0 daughter $\pi^{\pm}$. The full list of variables for those FastBDT training is listed in Table. 4 and Table. 5. The distributions of the multivariate analysis output for V0Selector and LambdaVeto in $1\,\mathrm{ab}^{-1}$ training MC dataset are shown in Fig. 29 and Fig. 30. The cut condition $\mathrm{mva(V0Selector)} > 0.90, \mathrm{mva(LambdaVeto)} > 0.11$ are then applied so that the figure of merit in terms of the number of trueK_S^0 against the number of remaining fake K_S^0 is maximized. As a result of those K_S^0 selection, we obtained 89.0 % signal K_S^0 selection efficiency with 0.45 % background K_S^0 efficiency for the selection using ksSelector.

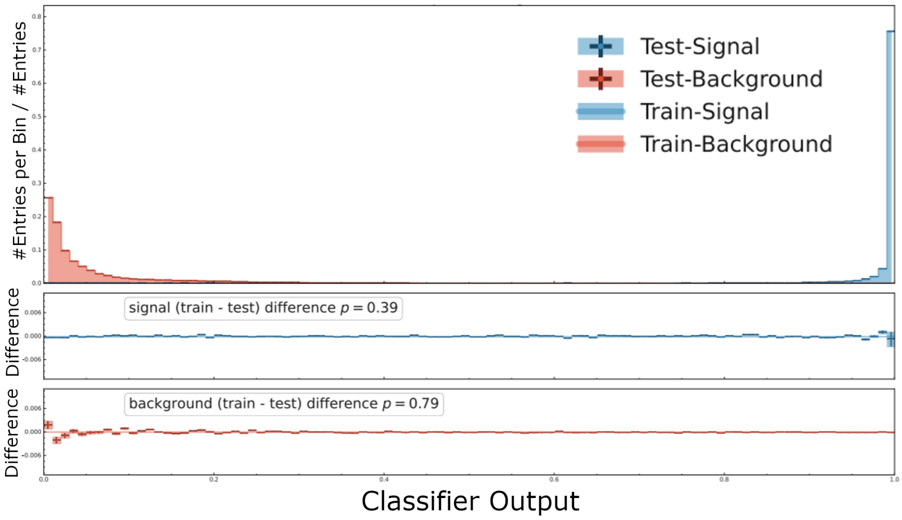


Figure 29: The output distribution of the V0Selector algorithm for true K_S^0 (Blue) and mis-reconstructed K_S^0 excluding Λ (Red).

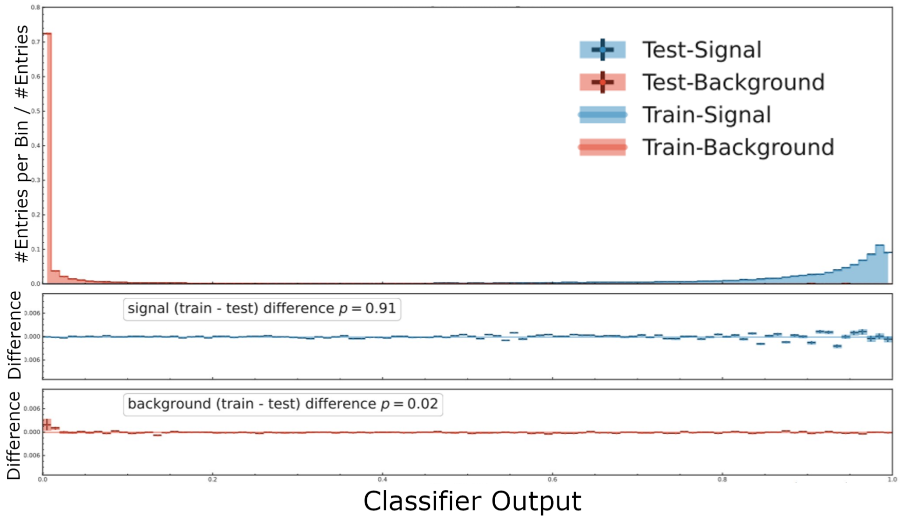


Figure 30: The output distribution of LambdaVeto algorithm for true K_S^0 (Blue) and Λ which is mis-reconstructed as K_S^0 (Red).

CP-side B^0 reconstruction and selection

CP-side B candidates, or B_{CP}, are reconstructed from $B \to K_S^0 K^+ K^-$ decay with vertex fitting method called TreeFitter [28], where B decay vertex is constrained to IP. Here we require tracks which are used for vertex fitting to have VXD hits sufficient for TreeFitter requirement. The decay vertex of a B_{CP} is constrained to the region extrapolated from interaction point region along the momentum of B_{CP}. The concept of vertex fitting for B_{CP} is shown in Fig. 31, where $p_{B_{CP}}$ is the momentum of B_{CP}. All the kinematics of

	Variable Description
cosVertexMomentum	The angle between K_S^0 vertex and K_S^0 momentum
FlightDistance	Flight Distance of K_S^0
significanceOfDistance	Significance of the distance between K_S^0 vertex and interaction point
cosHelicityAngleMomentum	The angle between momentum difference of $\pi^\pm$ in K_S^0 rest frame and K_S^0 momentum in lab frame
ImpactXY	The distance between K_S^0 vertex and IP in the xy plane
decayAngle(i) (i=0,1)	The angle between momentum of daughter $\pi^\pm$ and K_S^0 in K_S^0 rest frame
daughterAngleDiffInMother	The angle between π^+ and π^- in K_S^0 rest frame
daughtersDeltaZ	The difference of z position between π^+ and π^-
$dr_{\pi^\pm}$	Track parameter of π^+ and π^-
nSVDHits$_{\pi^\pm}$	The number of hits in SVD for π^+ and π^-
nCDCHits$_{\pi^\pm}$	The number of hits in CDC for π^+ and π^-

Table 4: Variables used for V0Selector algorithm, which mainly discriminate true K_S^0 from mis-reconstructed K_S^0 originating from IP.

	Variable Description
protonID$_{\pi^\pm}$	Particle identification of $\pi^\pm$ for proton
M_Λ	The mass of K_S^0 assuming proton mass hypothesis for π^+
$M_{\bar{\Lambda}}$	The mass of K_S^0 assuming proton mass hypothesis for π^-
$p_{\pi^\pm}$	The momentum for each $\pi^\pm$
cosTheta$_{\pi^\pm}$	The $\cos\theta$ for each $\pi^\pm$ defined as the polar angle of momentum

Table 5: Variables used for LambdaVeto algorithm, which mainly discriminate true K_S^0 from Λ.

daughter particles are updated using the fitting result for vertex position. CP-side B candidates satisfying $|\Delta E| < 0.2\,\mathrm{GeV}$ and $5.2 < M_{bc} < 5.3\,\mathrm{GeV}/c^2$ is then adopted, where M_{bc} and ΔE are defined as

$$M_{bc} = \sqrt{E_{\mathrm{beam}}^2 - p_B^{*2}}, \tag{88}$$

$$\Delta E = E_B^* - E_{\mathrm{beam}}. \tag{89}$$

Here E_{beam} and p_B^* is the beam energy and momentum of CP-side B meson in the center-of-mass frame, and E_B^* is the reconstructed B meson energy in the center-of-mass frame. Each candidate is ranked by B vertex quality χ^2 and used for selecting best B candidate when there exists multiple candidate for CP-side B meson within one event.

In this analysis, the region $|\Delta E| < 0.05\,\mathrm{GeV}$ and $5.27 < M_{bc} < 5.29\,\mathrm{GeV}/c^2$ is defined as a signal region, and the region $|\Delta E| < 0.2\,\mathrm{GeV}$ and $5.2 < M_{bc} < 5.26\,\mathrm{GeV}/c^2$ is defined

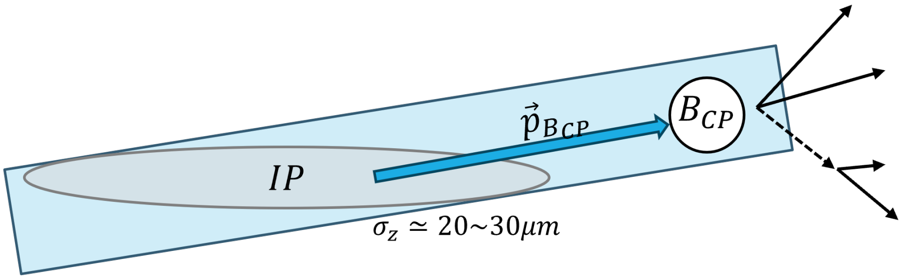


Figure 31: The concept of vertex fitting for the CP-side B candidates.

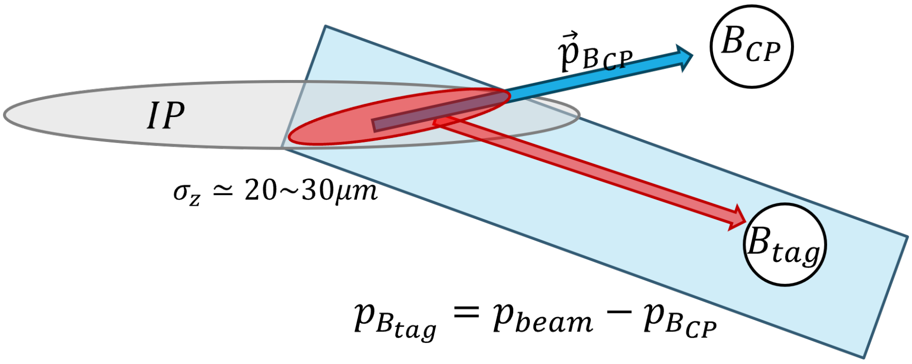


Figure 32: The concept of vertex fitting for the tag-side B candidates.

as a side-band region.

Tag-side B^0 reconstruction

Tag-side B meson, or B_{tag}, are reconstructed using remaining tracks and clusters which are not used for the reconstruction of CP-side B meson. The condition $|d_0| < 0.5\mathrm{cm}, |z_0| < 2.0\mathrm{cm}, p \geqq 0.05\,\mathrm{GeV}/c$ and the existence of hits in CDC and SVD are required for tracks to be used for vertex fitting of B_{tag}. The B_{tag} is then reconstructed using KFit algorithm with tube constraint. The concept of vertex fitting for B_{tag} is shown in Fig. 32. In the fitting with tube constraint, the B_{tag} vertex position is constraint to the tube region along the B_{tag} flight direction extrapolated from the intersection region of interaction point and B_{CP} vertex.

Additional tag-side vertex quality cuts are applied as $\sigma_{\Delta t} < 2\,\mathrm{ps}$, TagVpValue > 0 and TagVNDF > 0.5, where $\sigma_{\Delta t}$ is the error in the estimation of decay time difference between B_{CP} and B_{tag}, TagVpValue and TagVNDF is the p-value and the degree of freedom of the B_{tag} vertex fitting.

Flavor tagging

The flavor of B_{tag} is determined by identifying signatures through flavor-specific final states, as briefly mentioned in Sec. 5.1. In this section, we elaborate on the particle reconstruction process used for flavor tagging. To determie the flavor of B mesons, we employ multivariate algorithms utilizing a category-based flavor tagger [4] in this analysis.

Categories	Targets for $\overline{B}^0$
Electron	e^-
Intermediate Electron	e^+
Muon	μ^-
Intermediate Muon	μ^+
Kinetic Lepton	ℓ^-
Intermediate Kinetic Lepton	ℓ^+
Kaon	K^-
Kaon-Pion	K^-, π^+
Slow Pion	π^+
Maximum p^*	ℓ^-, π^-
Fast-Slow-Correlated (FSC)	ℓ^-, π^+
Fast Hadron	π^-, K^-
Lambda	Λ

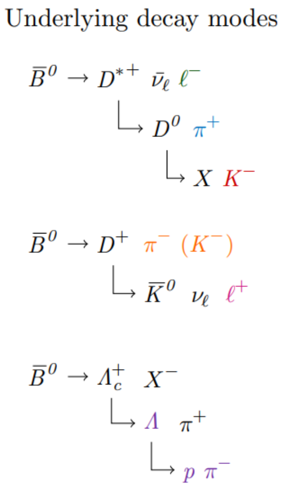


Figure 33: The list of tagging categories along with their corresponding targets (left) and the considered decay modes for each category (right) [4]. The parameter p^* represents the momentum in the center-of-mass frame, $l^\pm$ corresponds to charged leptons, and X denotes other potential particles involved in the decays.

This involves first categorizing the information related to the decay products of the B mesons and subsequently combining it to determine the flavor.

In the initial stage, kinematics, track hits, and Particle IDentification information (PID) are utilized as input values for each track. These are employed to identify particles among the candidates for the tag-side decay product, as listed in Fig. 33. For all categories except for the Maximum-p^* category, PID variables are employed. In the case of the Maximum-p^* category, PID variables are determined based on PID likelihoods for charged-particle hypotheses.

In the subsequent step, the category information obtained in the first step is integrated using the schematics presented in Fig. 34. The multivariate algorithm employed is based on FastBDT [25] classifiers. For each event, the decay products are categorized according to flavor signatures such as e, μ, K, π, and Λ, denoted by the green box in Fig. 34. The FBDT classifier calculates the qr value for each event, with the category information serving as input variables.

5.5 Charmed Veto

The contribution of the charmed decay of B mesons through the $b \to c$ transition adds to the background component originating from $B\bar{B}$ background. Consequently, it is necessary to reject this contribution in the analysis by applying cuts on the invariant mass of K^+K^-, $K^+K_S^0$, and $K^-K_S^0$. Table. 6 illustrates the charmed decays that are vetoed in this analysis through the application of the invariant mass cut. The $B^0 \to \chi_{c_0} K_S^0$ decay channel, however, is not vetoed in this analysis due to the substantial width of χ_{c_0}. Instead, it is included in the CP fitting procedure, with the CP asymmetry parameters

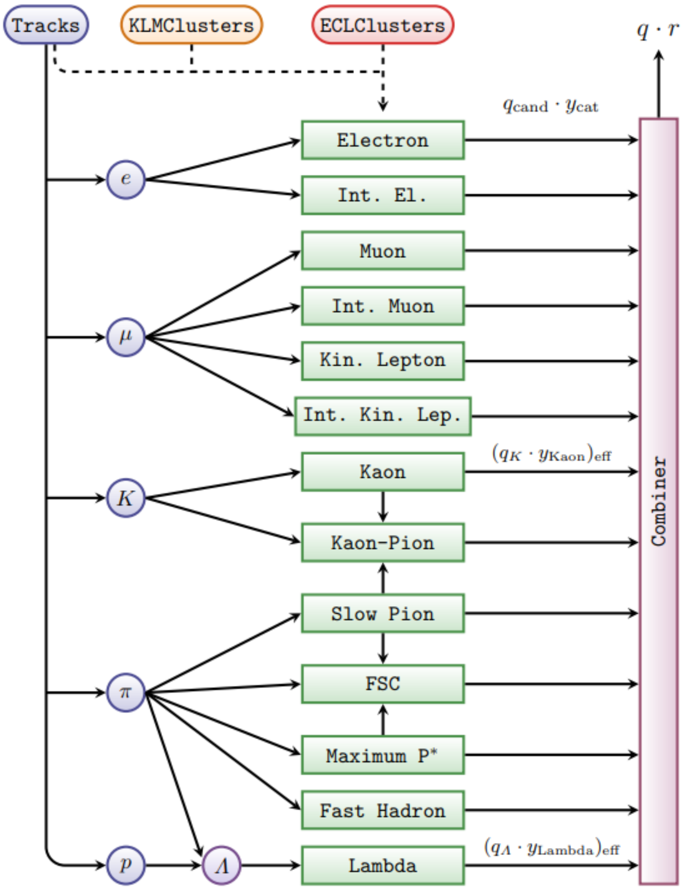


Figure 34: A schematic representation illustrating the second stage of the category-based flavor tagger [4].

of $B^0 \to \chi_{c_0} K_S^0$ fixed to the Particle Data Group (PDG) value [17] in the CP fitting. The potential $B\bar{B}$ background resulting from the mis-identification of B meson daughter particles is rejected by applying stringent kaonID criteria and therefore not vetoed in this analysis.

Mode	Vetoed Mass
$B^0 \to D^-(\to K_S^0 K^-)K^+$	$\lvert m_{K^\pm K_S^0} - m_{D^-}\rvert < 15\mathrm{MeV}/c^2$
$B^0 \to D_s^-(\to K_S^0 K^-)K^+$	$\lvert m_{K^\pm K_S^0} - m_{D_s^-}\rvert < 15\mathrm{MeV}/c^2$
$B^0 \to \bar{D}^0(\to K^+ K^-)K_S^0$	$\lvert m_{K^+K^-} - m_{D^0}\rvert < 15\mathrm{MeV}/c^2$

Table 6: Charmed decays to be vetoed in this analysis by applying cut to the invariant mass between two of the B meson daughters.

5.6 Continuum suppression

In order to reject contribution from $q\bar{q}$ background events, FastBDT is trained using variables representing event shape. The following variables are used for training :

- KSFW variables [9] : R2,M_{miss}^2, E_t, $H_{ij}^{\mathrm{so}}(i = 0, 1, 2 : j = 0, 2, 4)$, $H_i^{\mathrm{oo}}(i = 0, 1, 2, 3, 4)$. The definition of these variables are given in Appendix B.
- cosTBTO : Cosine of the angle between B_{CP} and B_{tag} thrust axes
- cosTBz : Cosine of the angle between B_{CP} thrust axis and z axis
- thrustOm : the magnitude of B_{tag} daughters thrust

The thrust axis is defined as a unit vector **t** maximizing the thrust magnitude T defined as

$$T = \frac{\sum_i |\boldsymbol{p_i} \cdot \boldsymbol{t}|}{\sum_i \boldsymbol{p_i}}, \tag{90}$$

where $\boldsymbol{p_i}$ is a momentum of a daughter track of B_{CP} or B_{tag}, and i denotes all daughter tracks of B_{CP} or B_{tag}. The FastBDT is trained with a sample consists of 20000 true samples from signal MC Dalitz events and 40000 false samples coming from $q\bar{q}$ continuum event. The checking of over-training is checked using independent sample than the sample used for FastBDT training. The result of training and over-training check is shown in Fig. 35 and Fig. 36. The cut threshold for the FastBDT output $\mathcal{O}_{\mathrm{CS}}$ are calculated to be 0.55 as a value which maximize $n_{\mathrm{sig}}/\sqrt{n_{\mathrm{sig}} + n_{\mathrm{bkg}}}$.

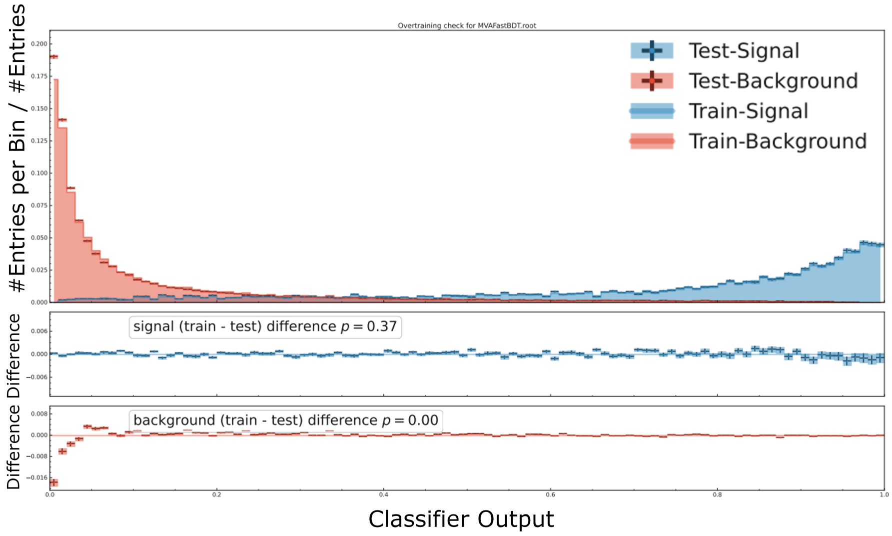


Figure 35: The result of continuum suppression FastBDT along with over training check and the distribution of multivariate output for signal event (Blue) and $q\bar{q}$ background events (Red).

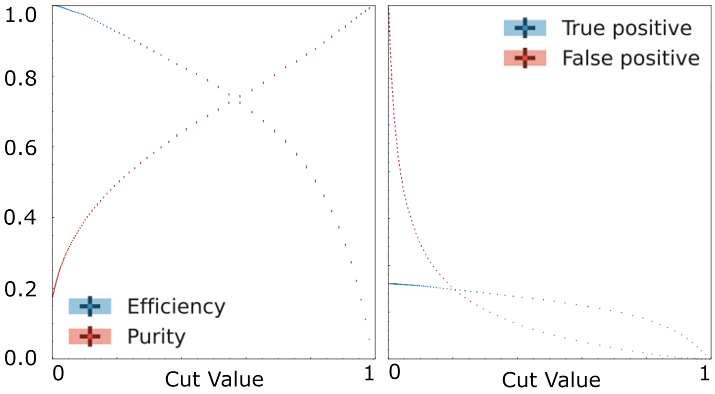


Figure 36: The plot of efficiency and purity of the continuum suppression FastBDT (left), and the true/false positive rates (right).

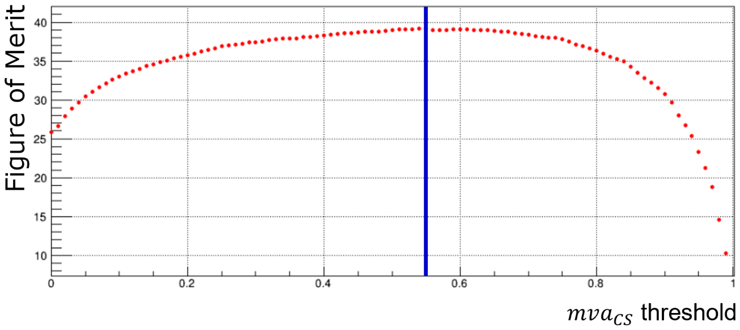


Figure 37: The relation between the threshold of continuum suppression FastBDT output and the figure of merit defined as $\mathrm{FoM} = n_{\mathrm{sig}}/\sqrt{n_{\mathrm{sig}} + n_{\mathrm{bkg}}}$.

5.7 Summary of event selection and reconstruction

The summary of all the cut criteria is shown in Table. 7. The breakdown of selection efficiency for each cut criteria are shown in Table. 8. The remaining number of event after all the cut in this section is listed in Table. 9, and the selection efficiency for signal events is estimated to be 14.25 ± 0.04 %. The tracking efficiency of charged particle including detector acceptance is expected to be $0.85 \sim 0.9$ for high momentum region with $p_T > 1\,\mathrm{GeV}$ and $0.7 \sim 0.8$ for low momentum region with $p_T < 1\,\mathrm{GeV}$ [39], where p_T is the transverse momentum of a track. We need all the daughter particles of B_{CP} to be reconstructed, leading to the dominant contribution to the signal reconstruction efficiency for the pre-selection step in Table. 9.

Pre-selection	
$K^\pm$	$\lvert d_0\rvert < 0.5\text{cm}$ $\lvert z_0\rvert < 2.0\text{cm}$ kaonID > 0.1
K_S^0	$0.485 < m_{K_S^0} < 0.515\,\text{GeV}/c^2$ flightLength $> 0.05\text{cm}$
B_{CP}	$5.2 < M_{bc} < 5.3\,\text{GeV}/c^2$ $\lvert\Delta E\rvert < 0.2\,\text{GeV}$ $\chi^2 > 0$
B_{tag}	TagVpValue > 0 TagVNDF > 0.5
Selection	
Vertex quality	$\sigma_{\Delta t} < 2\,\text{ps}$
ksSelector	$\text{mva}_{\text{V0Selector}} > 0.90$ $\text{mva}_{\text{LambdaVeto}} > 0.11$
$K^\pm$	kaonID > 0.6
Charmed Veto	$\lvert m_{K^\pm K_S^0} - m_{D^-}\rvert < 15\text{MeV}/c^2$ $\lvert m_{K^\pm K_S^0} - m_{D_S^-}\rvert < 15\text{MeV}/c^2$ $\lvert m_{K^+K^-} - m_{D^0}\rvert < 15\text{MeV}/c^2$
Best Candidate Selection	largest χ^2
Continuum Suppression	$\mathcal{O} > 0.55$

Table 7: List of selection criteria applied for the reconstruction of event candidates in this analysis.

	Signal Dalitz MC	Background MC	Real Data
Pre-Selection	100.00	100.00	100.00
$\sigma_{\Delta t} < 2ps$	98.23 (98.23)	98.42 (98.42)	98.35 (98.35)
K_S^0 Selection	94.81 (96.52)	82.74 (84.07)	78.13 (79.44)
kaonID > 0.6	78.95 (83.27)	53.93 (65.18)	45.65 (58.43)
Charmed Veto	75.99 (96.25)	52.10 (96.61)	44.13 (96.67)
Best Candidate Selection	74.75 (98.37)	50.71 (97.33)	43.18 (97.85)
Continuum Suppression	58.90 (78.80)	6.13 (12.09)	7.37 (17.07)
Signal Region Cut	58.09 (98.62)	0.23 (3.75)	0.52 (7.06)

Table 8: The breakdown of the selection efficiency for each cutting criteria for signal MC, generic background MC and real data in the unit of %. The number in a bracket represents the selection efficiency for each cut criteria.

Selection Criteria	$1\,\mathrm{ab}^{-1}$ background MC	1M signal Dalitz MC
PreSelection	763043	245229
All the cuts	1755	142454

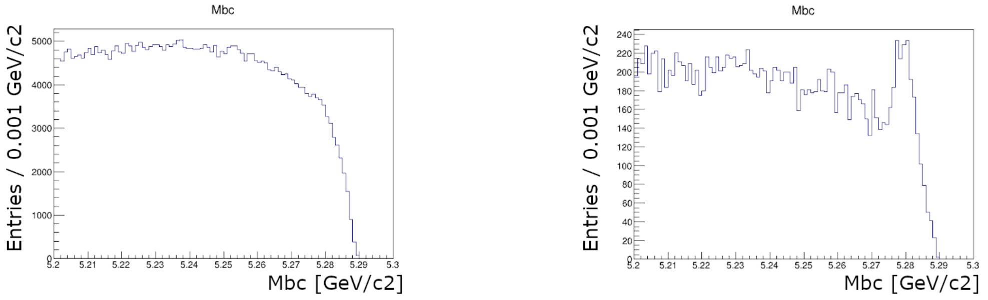
Table 9: The remaining number of event after all the event selection for $1\,\mathrm{ab}^{-1}$ background MC and 1M signal Dalitz MC. The criteria for Pre-selection is shown in Table. 7.

6 Signal Extraction Fitting

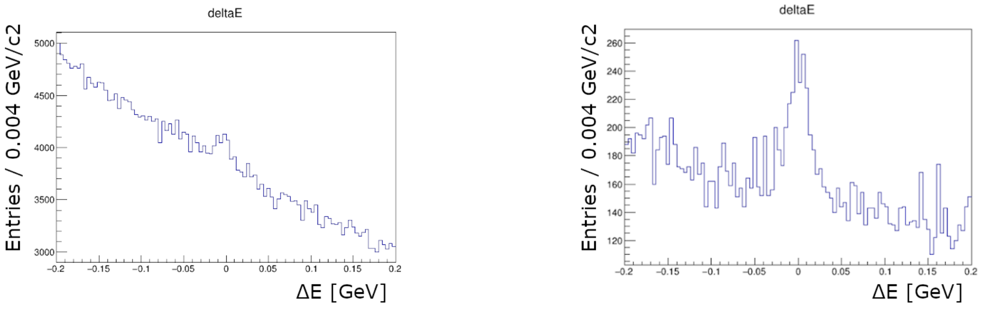
After all the event selection and reconstruction explained in Section.5, there still exists background events originating from $q\bar{q}$ or $B\bar{B}$ decays in addition to signal events. In order to calculate expected number of yield for each $q\bar{q}$, $B\bar{B}$ and signal event, the signal extraction fitting is performed. In this section, the method to estimate signal yield and event-by-event fraction is explained, which is used in the CP fitting procedure later in Sec. 7 .

6.1 Method to extract event fraction

In order to obtain event-by-event fraction and the yield of signal, $q\bar{q}$ and $B\bar{B}$ events, simultaneous fitting to M_{bc}, ΔE and qr-bin are performed. The event distributions for the M_{bc} and ΔE before or after event selection are shown in Fig. 38. The simultaneous fit is then executed by extended maximum likelihood fit with the following PDF,



(a) M_{bc} before selection (b)



(c) (d) ΔE after selection

Figure 38: Distribution of M_{bc} or ΔE for $361.6\,\mathrm{fb}^{-1}$ real data. The left figures represents distribution with only pre-selections, and the right figures represents distribution with all cuts including continuum suppression.

$$\mathcal{P}^{\rm ext}_{\rm total}(\vec{x}, l) = N_{\rm sig} f^{{\rm ext},l}_{\rm sig} \mathcal{P}^{\rm ext}_{\rm sig}(\vec{x}) + N_{qq} f^{{\rm ext},l}_{q\bar{q}} \mathcal{P}^{\rm ext}_{q\bar{q}}(\vec{x}) + N_{B\bar{B}} f^{{\rm ext},l}_{B\bar{B}} \mathcal{P}^{\rm ext}_{B\bar{B}}(\vec{x}) \tag{91}$$

where $\vec{x} = (M_{bc}, \Delta E)$, l is the qr-bin determined by flavor tagging, $N_{\rm sig}, N_{q\bar{q}}, N_{B\bar{B}}$ is the total number of yield for each event type, and $f^{{\rm ext},l}_{\rm sig}, f^{{\rm ext},l}_{q\bar{q}}, f^{{\rm ext},l}_{B\bar{B}}$ is the fraction of the number of event in qr-bin l for each event type. The individual yield of each event type for qr-bin l is then given as

$$N_x^l = N_x f_x^{\mathrm{ext},l} \tag{92}$$

where $x = \mathrm{sig}, q\bar{q}$, or $B\bar{B}$. The PDF modelling of $(M_{bc}, \Delta E)$ is common for all the qr-bin [0-6] in this analysis.

6.2 Models of PDF

In this subsection, PDF modelling for each event type are described. The summary of the $(M_{bc}, \Delta E)$ modeling is listed in Table. 10

	M_{bc}	ΔE
Signal	Double Gaussian	Double Gaussian
$q\bar{q}$	Argus	First Order Polynomial
$B\bar{B}$	Argus+Gaussian	Second Order Polynomial

Table 10: PDF models of signal, $q\bar{q}$ and $B\bar{B}$, which is used for signal extraction fit.

Signal events

For signal event, PDF is defined as

$$\begin{aligned}\mathcal{P}_{\mathrm{sig}}^{\mathrm{ext}}(\vec{x}) = {}& (G_{main}^{M_{bc}}(M_{bc}; \mu_{main}^{M_{bc}}, \sigma_{main}^{M_{bc}}) + f_{tail}^{M_{bc}} G_{tail}^{M_{bc}}(M_{bc}; \mu_{tail}^{M_{bc}}, \sigma_{tail}^{M_{bc}})) \\ & \times (G_{main}^{\Delta E}(\Delta E; \mu_{main}^{\Delta E}, \sigma_{main}^{\Delta E}) + f_{tail}^{\Delta E} G_{tail}^{\Delta E}(\Delta E; \mu_{tail}^{\Delta E}, \sigma_{tail}^{\Delta E}))\end{aligned} \tag{93}$$

where both M_{bc} and ΔE are modeled as a double Gaussian. The parameters for tail fraction and tail Gaussian are fixed to the values obtained by fitting to 1M signal only Dalitz MC samples. The main Gaussian parameters are obtained in the simultaneous fitting to the real data.

$q\bar{q}$ background events

For $q\bar{q}$ background events, PDF is defined as

$$\mathcal{P}_{q\bar{q}}^{\mathrm{ext}}(\vec{x}) = Argus(M_{bc}; m_{q\bar{q}}, k_{q\bar{q}}) \times Pol_1(\Delta E; p1_{q\bar{q}}) \tag{94}$$

where M_{bc} is modeled as an Argus function [5] and ΔE is modeled as a first order polynomial. The Argus function is defined as

$$Argus(m; m_0, k) = \mathcal{N} \cdot m \cdot \sqrt{1 - \left(\frac{m}{m_o}\right)^2} \cdot \exp\left(k \cdot \left(1 - \left(\frac{m}{m_o}\right)^2\right)\right). \tag{95}$$

The shape parameters for $q\bar{q}$ background are determined by the simultaneous fitting to the real data.

$B\bar{B}$ background events

For $B\bar{B}$ background events, PDF is defined as

$$\mathcal{P}^{\text{ext}}_{B\bar{B}}(\vec{x}) = (Argus(M_{bc}; m_{B\bar{B}}, k_{B\bar{B}}) + f_G^{B\bar{B}} G^{M_{bc}}_{B\bar{B}}(M_{bc}; \mu^{M_{bc}}_{B\bar{B}}, \sigma^{M_{bc}}_{B\bar{B}})) \times Pol_2(\Delta E; p1_{B\bar{B}}, p2_{B\bar{B}}) \tag{96}$$

where M_{bc} is modeled as the sum of Argus function and Gaussian, due to the peaking $B\bar{B}$ background component. ΔE distribution is modeled as a second order polynomial. The parameters for $B\bar{B}$ background shapes are fixed to the value obtained by fitting to $1\,\text{ab}^{-1}$ MC dataset due to the small statistics of $B\bar{B}$ background events in the real data. In addition to the PDF parameters, the event fraction $f^{\text{ext},l}_{B\bar{B}}$ is fixed to the value obtained from 1ab^{-1} background MC dataset.

PDF parameters from fitting to MC samples

In order to determine part of PDF parameters, fitting to $(M_{bc}, \Delta E)$ distribution with 1M signal MC samples and $1\,\text{ab}^{-1}$ background MC are performed. Here free parameters to be fitted is $N_{sig,q\bar{q},B\bar{B}}$, $f^{\text{ext},l}_{sig,q\bar{q}}$ in addition to $\mathcal{P}^{\text{ext}}_{sig,q\bar{q}}$ parameters. The comparison between $(M_{bc}, \Delta E)$ distribution and fitted PDF for all qr-bin region is shown in Fig. 40, and the fitted PDF parameters are listed in Table. 11.

6.3 Validation of signal extraction fit with Toy MC

The signal extraction fitting is then performed using unbinned maximum likelihood by minimizing log-likelihood defined as

$$-2\log\mathcal{L} = -2\sum_i \log(P^{\text{ext}}_{\text{total}}(\vec{x})) \tag{97}$$

The fitting procedure of signal fraction is validated using Toy MC study. 150 Toy MC datasets are generated based on the PDF described above. The mean and sigma of the pull distribution of the fit result of $f^{\text{ext},l}_{\text{sig}}$ and $f^{\text{ext},l}_{q\bar{q}}$ from nominal values are shown in Fig. 39. For all the event fractions to be fitted in the simultaneous fitting, the pull distribution of the fit result are consistent with normal distribution.

6.4 Signal extraction fit to real data

The simultaneous fitting are performed to the real dataset with 361.6fb^{-1}, where part of the PDF parameters is fixed to the value obtained by fitting to MC dataset as mentioned in Sec. 6.2. The result of the fitting parameters and parameters fixed from MC is shown in Table. 12. The yield of each event type is shown in Table. 13 and Table. 14. The comparison between event distribution and fitted PDF are shown in Fig. 41, Fig. 42, and Fig. 43. The comparison of the event distribution in signal region and the PDF shapes projected to signal region is shown in Fig. 44. As seen in Fig. 44, no bias between event distribution and projected PDF was observed in the signal region, and thus event fractions in the signal region is correctly estimated by fitting to all the events including side-band region.

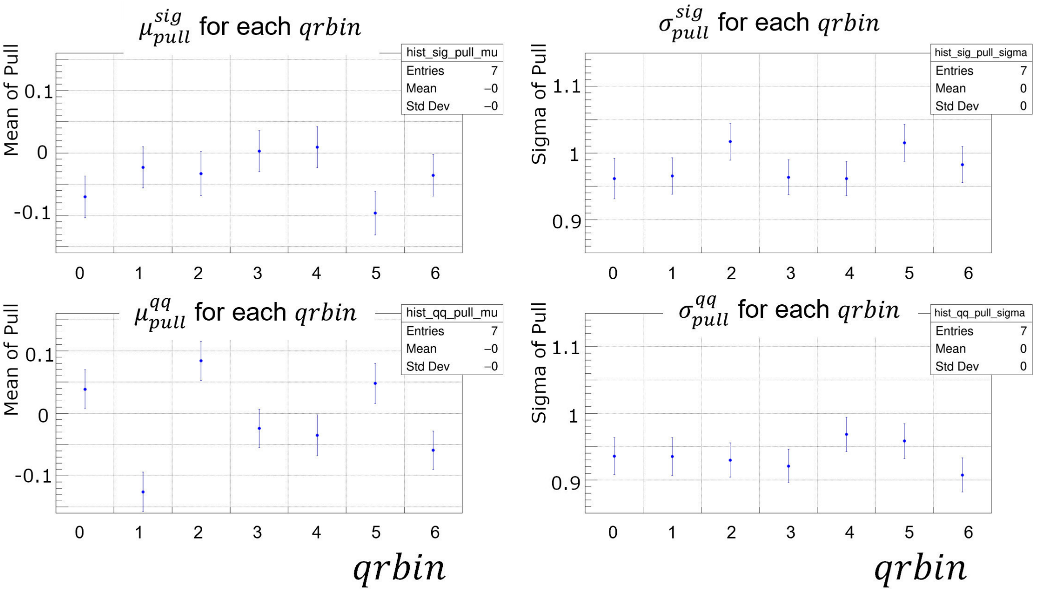
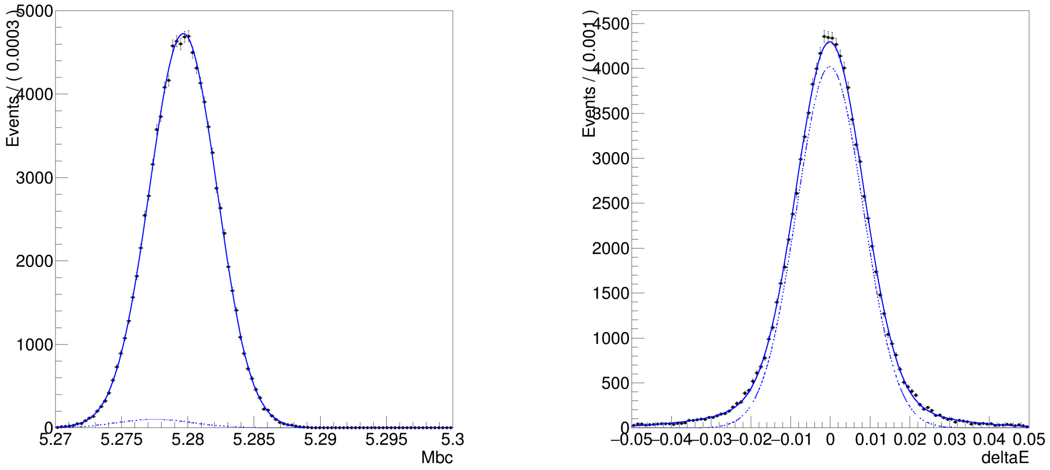
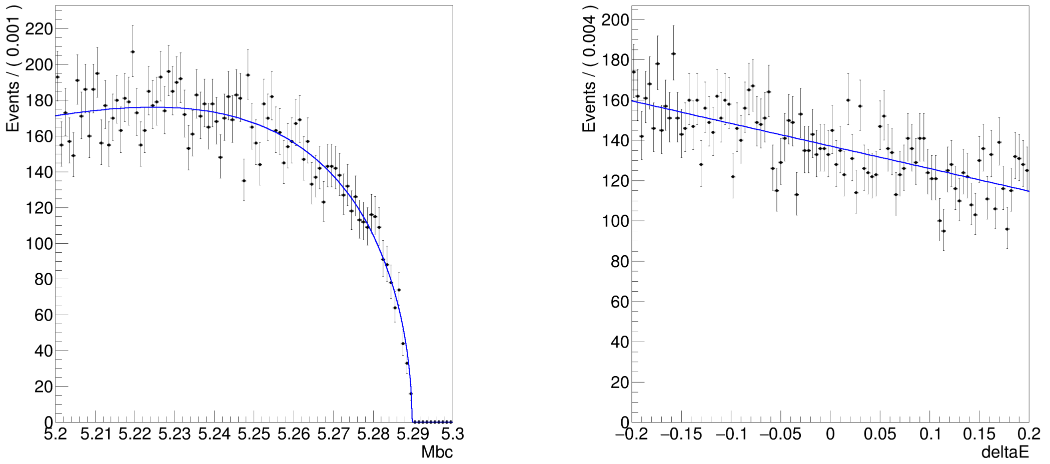


Figure 39: The result of the linearity check of signal extraction fitting procedure using Toy MC study. The left (right) figures shows the mean (σ) of the pull distribution of the fitted event fraction $f_{\rm sig}^{{\rm ext},l}$ and $f_{q\bar{q}}^{{\rm ext},l}$ obtained by Gaussian fitting to the pull distribution.



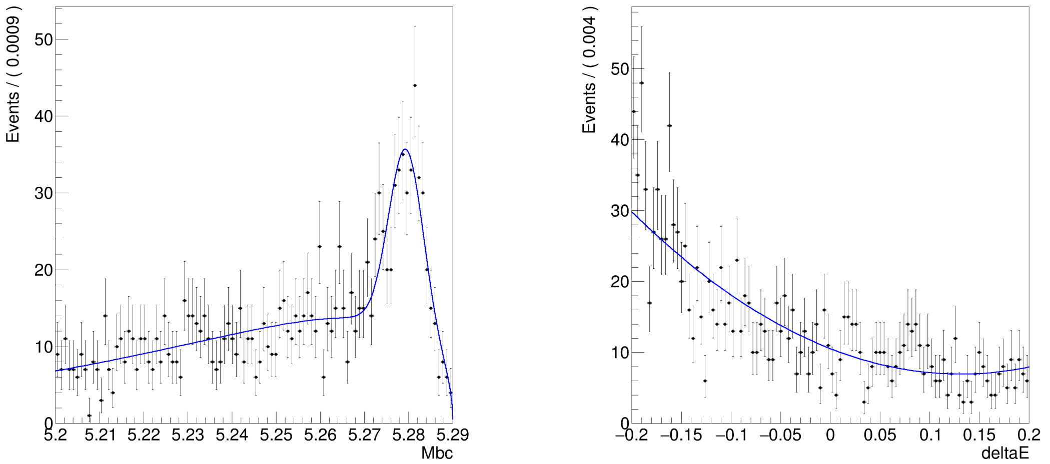
(a) M_{bc} for signal event

(b) ΔE for signal event



(c) M_{bc} for $q\bar{q}$ event

(d) ΔE for $q\bar{q}$ event



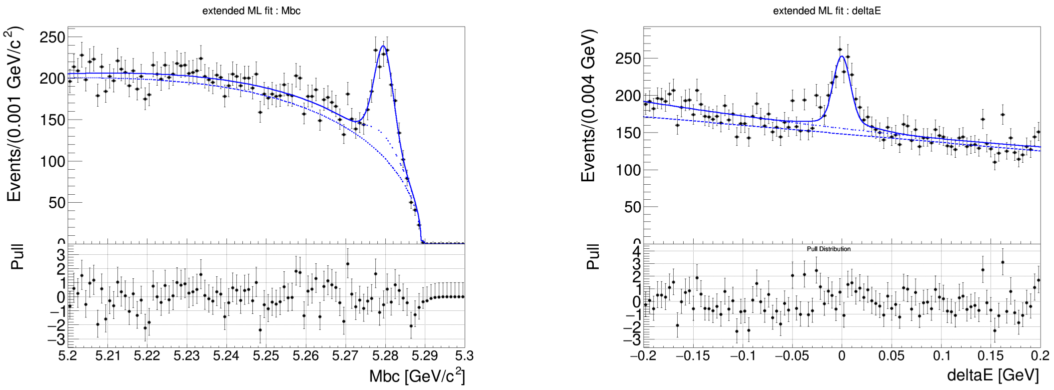
(e) M_{bc} for $B\bar{B}$ event

(f) ΔE for $B\bar{B}$ event

Figure 40: Distribution of M_{bc} and ΔE for MC datasets with fitted PDF for signal, $q\bar{q}$ and $B\bar{B}$ event respectively. The solid line represents fitted PDF, and the dotted line in signal events represents one of Gaussian component among the double Gaussian.

Parameter	Fit result
Signal Event	
$\mu_{main}^{M_{bc}}$	5.2797 ± 0.0000 GeV/c^2
$\sigma_{main}^{M_{bc}}$	0.0025 ± 0.0000 GeV/c^2
$\mu_{tail}^{M_{bc}}$	5.2775 ± 0.0030 GeV/c^2
$\sigma_{tail}^{M_{bc}}$	0.0029 ± 0.0005 GeV/c^2
$f_{tail}^{M_{bc}}$	0.2403 ± 0.0500
$\mu_{main}^{\Delta E}$	-0.0013 ± 0.0002 GeV
$\sigma_{main}^{\Delta E}$	0.0214 ± 0.0004 GeV
$\mu_{tail}^{\Delta E}$	-0.0010 ± 0.0003 GeV
$\sigma_{tail}^{\Delta E}$	0.0245 ± 0.0004 GeV
$f_{tail}^{\Delta E}$	0.2710 ± 0.0420
$q\bar{q}$ Event	
$m_{q\bar{q}}$	5.2898 ± 0.0005 GeV/c^2
$k_{q\bar{q}}$	-19.8969 ± 0.9557
$p1_{q\bar{q}}$	-0.8163 ± 0.0730
$B\bar{B}$ Event	
$\mu_{B\bar{B}}^{M_{bc}}$	5.27945 ± 0.0004 GeV/c^2
$\sigma_{B\bar{B}}^{M_{bc}}$	0.0039 ± 0.0005 GeV/c^2
$k_{B\bar{B}}$	-54.4695 ± 4.9340
$p1_{B\bar{B}}$	-5.19385 ± 0.3704
$p2_{B\bar{B}}$	19.8175 ± 3.4388 (GeV/c^2)$^{-1}$
f_G^{BB}	0.8086 ± 0.0236

Table 11: The list of parameters describing PDF for signal extraction fitting, calibrated with $1\,\mathrm{ab}^{-1}$ MC dataset for $q\bar{q}$ and $B\bar{B}$ background, and 1M signal MC for signal event.



(a) M_{bc} qr-bin 0

(b) ΔE qr-bin 0

Figure 41: Distribution of M_{bc} and ΔE summed over all qr-bin and including both signal region and side-band region. The dot, dashed and solid line represents $q\bar{q}$,$B\bar{B}$, and signal yield respectively. The bottom plot represents pull of the event distribution from fitted PDF.

Parameter	Fit result
Signal Event	
$\mu_{main}^{M_{bc}}$	5.2797±0.0001 GeV/c^2
$\sigma_{main}^{M_{bc}}$	0.0024±0.0001 GeV/c^2
$\mu_{tail}^{M_{bc}}$	5.2775 GeV/c^2
$\sigma_{tail}^{M_{bc}}$	0.0029 GeV/c^2
$f_{tail}^{M_{bc}}$	0.2403
$\mu_{main}^{\Delta E}$	0.0027±0.0053 GeV
$\sigma_{main}^{\Delta E}$	0.0289±0.0049 GeV
$\mu_{tail}^{\Delta E}$	-0.0010 GeV
$\sigma_{tail}^{\Delta E}$	0.0245 GeV
$f_{tail}^{\Delta E}$	0.2710
$q\bar{q}$ Event	
$m_{q\bar{q}}$	5.2889±0.0005 GeV/c^2
$k_{q\bar{q}}$	-15.3462±1.2870
$p1_{q\bar{q}}$	-0.7792±0.0840
$B\bar{B}$ Event	
$\mu_{B\bar{B}}^{M_{bc}}$	5.27945 GeV/c^2
$\sigma_{B\bar{B}}^{M_{bc}}$	0.0039 GeV/c^2
$k_{B\bar{B}}$	-54.4695
$p1_{B\bar{B}}$	-5.19385
$p2_{B\bar{B}}$	19.8175$(\mathrm{GeV}/c^2)^{-1}$
f_G^{BB}	0.8086

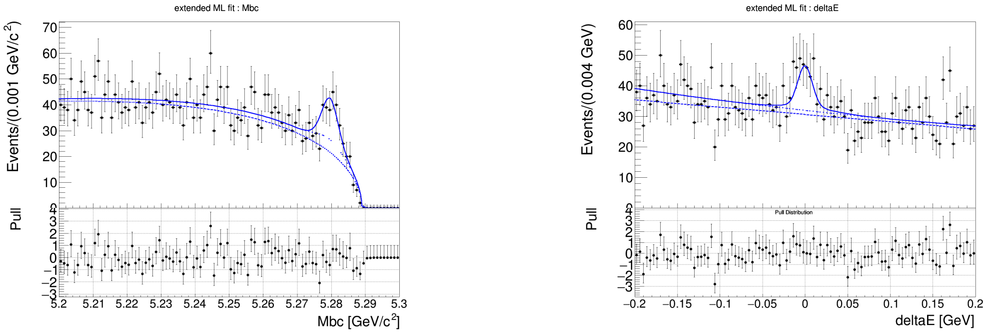
Table 12: The list of shape parameters used for signal extraction fit obtained by simultaneous fitting to the 361.6 fb^{-1} real data. The errors is statistics. Values without statistical error is fixed to the value calibrated using MC sample as listed in Table. 11.

All Region	Signal	$q\bar{q}$	$B\bar{B}$
qr-bin 0	95.1±11.7	3077.5±57.1	154.1±37.2
qr-bin 1	82.5±10.8	3018.2±56.4	136.3±32.9
qr-bin 2	106.6±12.2	2887.4±55.2	137.6±33.2
qr-bin 3	90.9±11.1	2220.2±48.5	110.2±26.6
qr-bin 4	75.3±10.2	1813.5±43.9	95.5±23.1
qr-bin 5	63.4±9.3	1365.7±38.3	91.7±22.2
qr-bin 6	119.3±11.6	513.9±25.2	123.5±29.9
Total	633.2±29.2	14896.5±126.0	848.9±205.1

Table 13: The number of event yields for each event type obtained by signal extraction fit to 361.6 fb^{-1} real data. This table includes both signal region and side-band region.

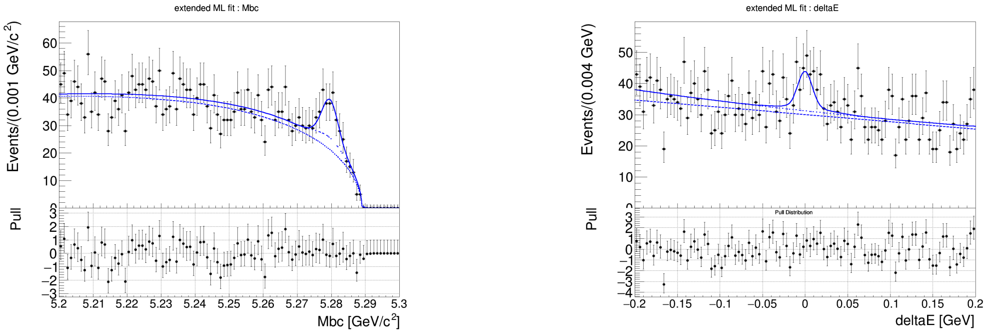
Signal Region	Signal	$q\bar{q}$	$B\bar{B}$
qr-bin 0	92.9±11.4	98.3±1.9	11.7±2.8
qr-bin 1	80.6±10.6	97.8±1.8	10.4±2.4
qr-bin 2	104.1±11.9	93.6±1.8	10.5±2.5
qr-bin 3	88.8±10.8	72.0±1.6	8.5±2.0
qr-bin 4	73.6±10.0	58.8±1.4	7.3±1.7
qr-bin 5	61.9±9.1	44.3±1.2	7.0±1.6
qr-bin 6	118.6±11.3	16.7±0.8	9.4±2.2
Total	620.5±28.5	481.4±4.1	64.8±15.2
Expected from MC	708.5	600.0	35.1

Table 14: The number of event yields for each event type obtained by signal extraction fit to 361.6 fb^{-1} real data. This table includes only signal region.



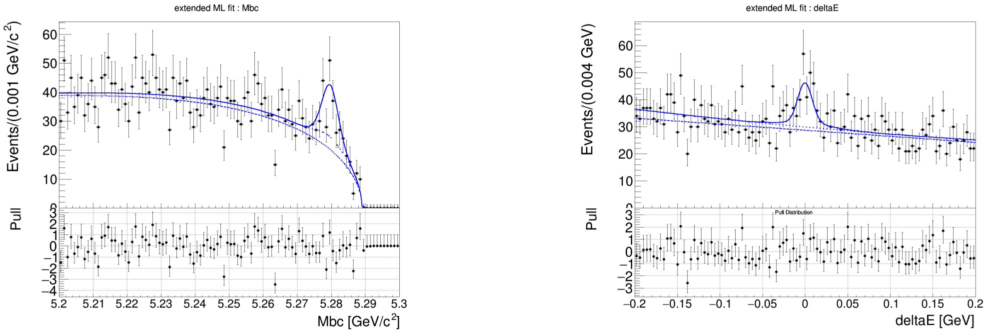
(a) M_{bc} qr-bin 0

(b) ΔE qr-bin 0



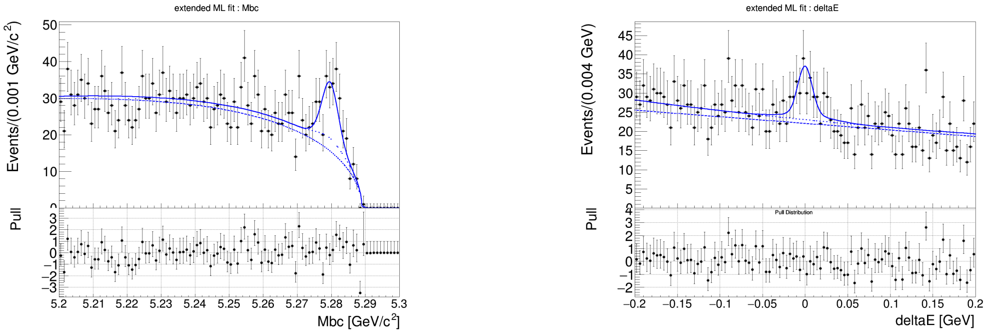
(c) M_{bc} qr-bin 1

(d) ΔE qr-bin 1



(e) M_{bc} qr-bin 2

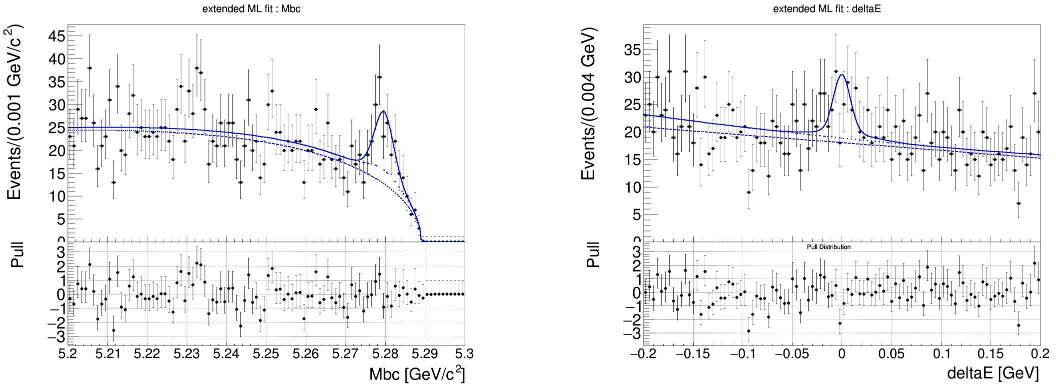
(f) ΔE qr-bin 2



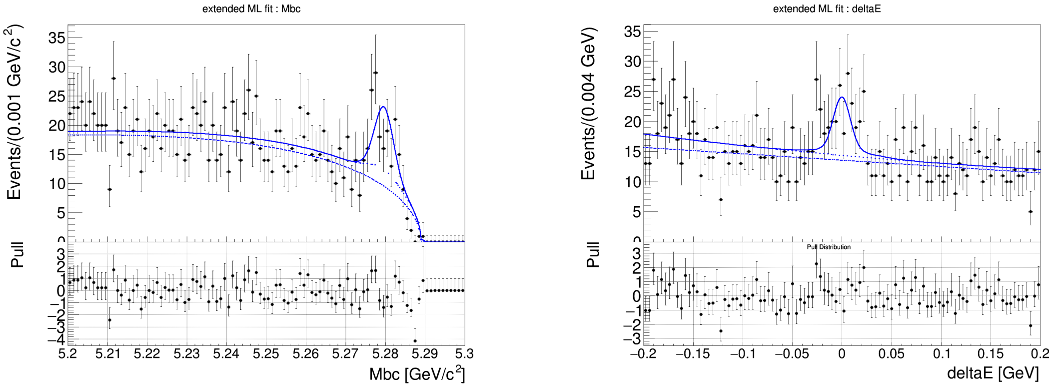
(g) M_{bc} qr-bin 3

(h) ΔE qr-bin 3

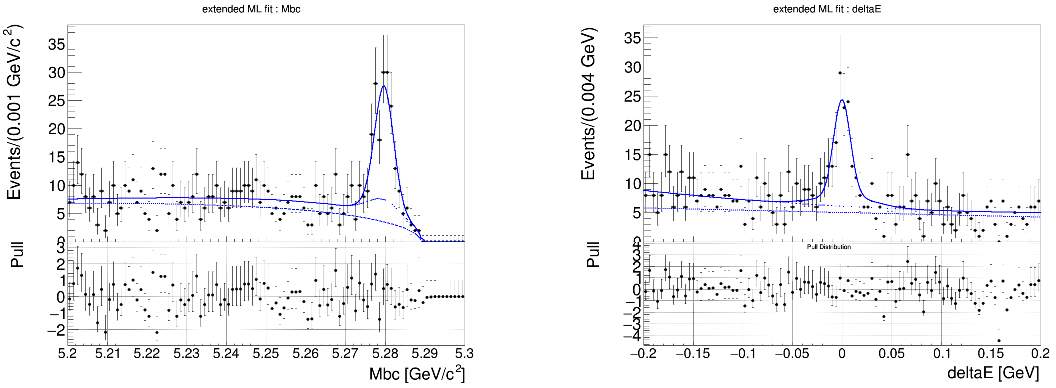
Figure 42: Distribution of M_{bc} and ΔE for qr-bin 0-3 including both signal region and side-band region. The dot, dashed and solid line represents $q\bar{q}$, $B\bar{B}$, and signal yield respectively. The bottom plot represents pull of the event distribution from fitted PDF.



(a) M_{bc} qr-bin 4 (b) ΔE qr-bin 4

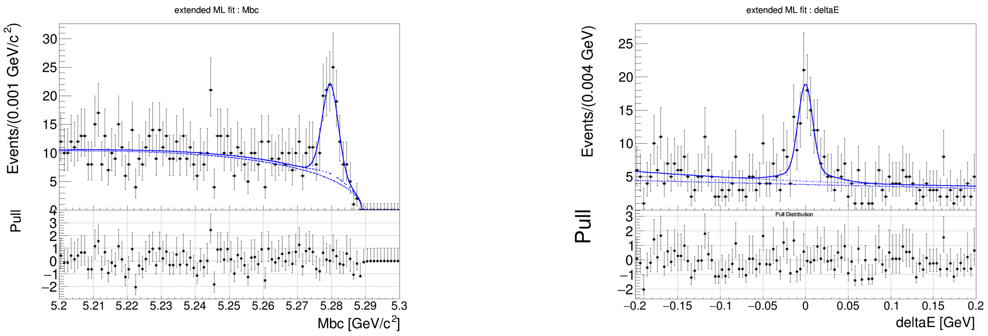


(c) M_{bc} qr-bin 5 (d) ΔE qr-bin 5

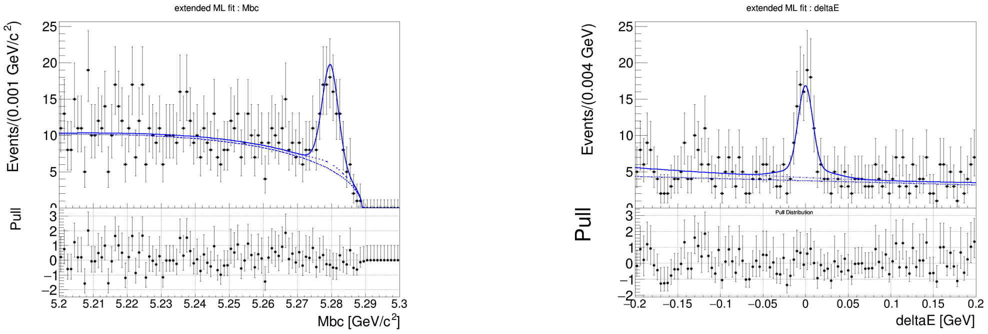


(e) M_{bc} qr-bin 6 (f) ΔE qr-bin 6

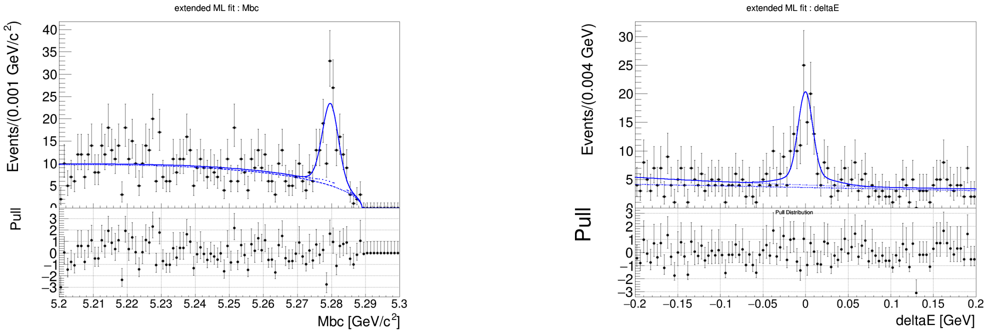
Figure 43: Distribution of M_{bc} and ΔE for qr-bin 4-6 including both signal region and side-band region. The dot, dashed and solid line represents $q\bar{q}$, $B\bar{B}$, and signal yield respectively. The bottom plot represents pull of the event distribution from fitted PDF.



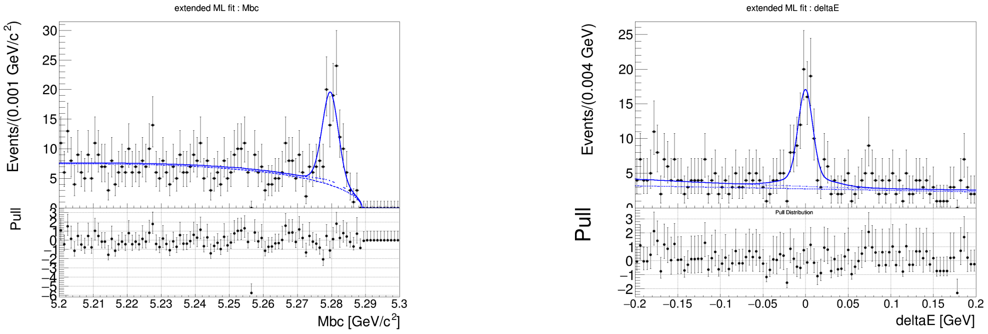
(a) M_{bc} qr-bin 0 (b) ΔE qr-bin 0



(c) M_{bc} qr-bin 1 (d) ΔE qr-bin 1

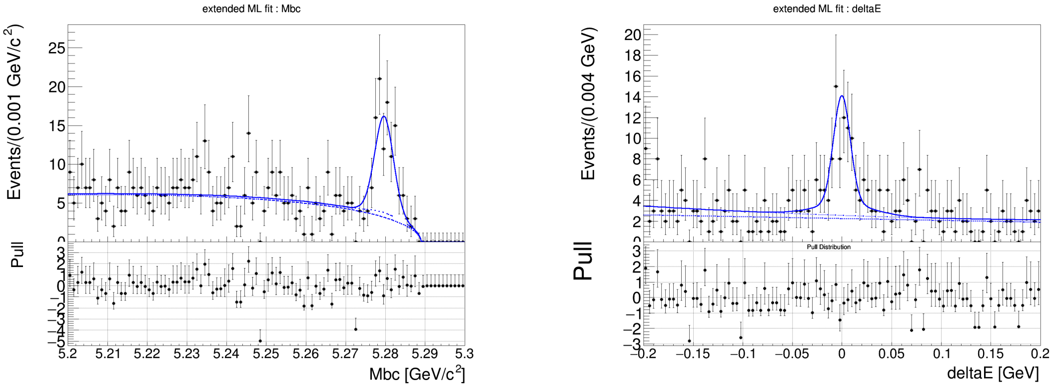


(e) M_{bc} qr-bin 2 (f) ΔE qr-bin 2

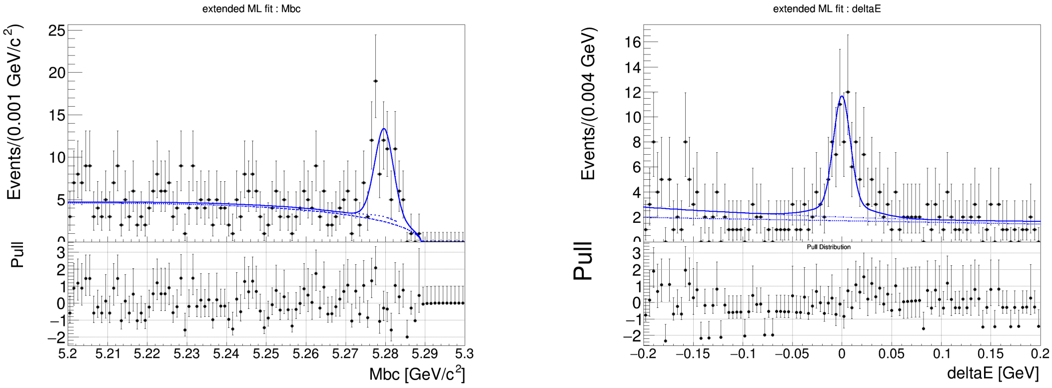


(g) M_{bc} qr-bin 3 (h) ΔE qr-bin 3

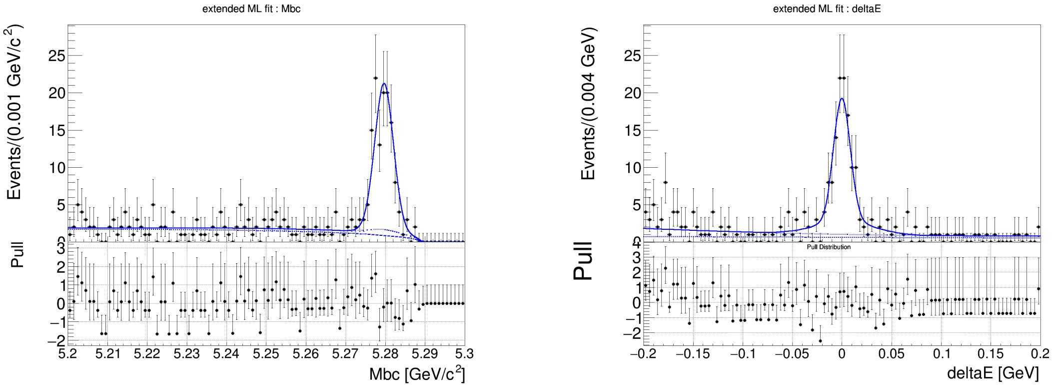
Figure 44: Distribution of M_{bc} and ΔE for qr-bin 0-3 projected to signal region. The dot, dashed and solid line represents $q\bar{q}$, $B\bar{B}$, and signal yield respectively. The bottom plot represents pull of the event distribution from PDF.



(a) M_{bc} qr-bin 4 (b) ΔE qr-bin 4



(c) M_{bc} qr-bin 5 (d) ΔE qr-bin 5



(e) M_{bc} qr-bin 6 (f) ΔE qr-bin 6

Figure 45: Distribution of M_{bc} and ΔE for qr-bin 4-6 projected to signal region. The dot, dashed and solid line represents $q\bar{q}$, $B\bar{B}$, and signal yield respectively. The bottom plot represents pull of the event distribution from PDF.

7 CP Fitting

In this section, we first provide the modelling of Δt distribution for each signal, $q\bar{q}$ and $B\bar{B}$ event type, which is critical for the measurement of CP asymmetries through Δt distribution. Then the full PDF to extract CP violating parameters is introduced.

7.1 Δt modeling

Signal event

For signal resolution function, we adopt the Δt resolution function which is developed and calibrated for $B^0 \to D^{(*)-}\pi^-$ and $B^0 \to J/\psi K_S^0$ decay channels [20]. The resolution function consists of a double Gaussian with exponentially decaying tails and outlier Gaussian, which is written as

$$\mathcal{R}_{\rm sig}(\delta\Delta t;\sigma_{\Delta t}) = (1-f_{\rm OL})\mathcal{R}_{\rm core}(\delta\Delta t;\sigma_{\Delta t}) + f_{\rm OL}\mathcal{R}_{\rm OL}(\delta\Delta t;\sigma_{\Delta t}) \tag{98}$$

where $\delta\Delta t$ is defined as the difference between measured Δt and true $\Delta t'$ defined as $\delta\Delta t \equiv \Delta t - \Delta t'$, and $f_{\rm OL}$ is the fraction of outlier. $\mathcal{R}_{\rm core}$ is expanded to

$$\begin{aligned}\mathcal{R}_{\rm core}(\delta\Delta t;\sigma_{\Delta t}) =& (1-f_{\rm tail})G(\delta\Delta t;\mu_{\rm main}\sigma_{\Delta t}, s_{\rm main}\sigma_{\Delta t})\\ &+ (1-f_{\rm exp})f_{\rm tail}G(\delta\Delta t;\mu_{\rm tail}\sigma_{\Delta t}, s_{\rm tail}\sigma_{\Delta t})\\ &+ f_{\rm exp}f_{\rm tail}G(\delta\Delta t;\mu_{\rm tail}\sigma_{\Delta t}, s_{\rm tail}\sigma_{\Delta t})\\ &\otimes (1-f_{\rm R})\exp_{-}(\delta\Delta t/(c\sigma_{\Delta t})) + f_{\rm R}\exp_{+}(-\delta\Delta t/(c\sigma_{\Delta t}))\end{aligned} \tag{99}$$

where $\sigma_{\Delta t}$ is the uncertainty of Δt measurement from vertex fitting quality and $\exp_+(x) = \exp(x)$ where $x > 0$, and similar for $\exp_-(x)$. The tail fraction $f_{\rm tail}$ is a function of $\sigma_{\Delta t}$ as

$$f_{\rm tail}(\sigma_{\Delta t}; f_{\rm max}, f_\mu, f_\sigma) = \frac{1}{2}f_{\rm max}\left(\operatorname{erf}\left(\frac{f_\mu}{f_\sigma\sqrt{2}}\right) - \operatorname{erf}\left(\frac{f_\mu - \sigma_{\Delta t}}{f_\sigma\sqrt{2}}\right)\right), \tag{100}$$

where the error function erf(x) is defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt. \tag{101}$$

The outlier of the resolution function $\mathcal{R}_{\rm OL}(\delta\Delta t;\sigma_{\Delta t})$ is described as a Gaussian with mean 0 and width independent of $\sigma_{\Delta t}$ as

$$\mathcal{R}_{\rm OL}(\delta\Delta t;\sigma_{\Delta t}) = G(\delta\Delta t; 0, \sigma_{\rm OL}) \tag{102}$$

All the parameters for the resolution function are common for qr-bin = $[0, 5]$ region, expect for the qr-bin = 6 region where c, $f_{\rm exp}$, f_R, $\mu_{\rm main}$ and $\mu_{\rm tail}$ are allowed to be independent from parameters for qr-bin = $[0, 5]$ region. The list of resolution function parameters calibrated for $B^0 \to D^{(*)-}\pi^-$ decay channels with $1\,\mathrm{ab}^{-1}$ MC dataset and $361.6\,\mathrm{fb}^{-1}$ real data is shown in Table. 15.

Parameter	MC15ri [1 ab^{-1}]	Data [361.6 fb^{-1}]
μ_{main}	(-0.1561 ± 0.0271)	(-0.0658 ± 0.0389)
μ^{6}_{main}	(-0.0756 ± 0.0586)	(-0.0263 ± 0.0854)
μ_{tail}	(-0.7940 ± 0.1036)	(-0.7604 ± 0.1887)
μ^{6}_{tail}	(-0.7280 ± 0.1975)	(-0.5369 ± 0.3238)
σ_{main}	(1.1875 ± 0.0335)	(1.0588 ± 0.0544)
σ_{tail}	(2.4129 ± 0.1318)	(2.4328 ± 0.2466)
f_{max}	(0.3074 ± 0.0251)	(0.2868 ± 0.0432)
c	4.2767	4.2767
c^{6}	3.9239	3.9239
f_R	0.2478	0.2478
f^{6}_R	0.2454	0.2454
f_{exp}	0.2097	0.2097
f^{6}_{exp}	0.2216	0.2216
f_μ	0.2361.6	0.2361.6
f_σ	0.0934	0.0934
f_{OL}	0.0001	0.0001

Table 15: Resolution function parameters calibrated for $B^0 \to D^{(*)-}\pi^-$ decay channel [20]. The fitting parameters without uncertainty is fixed from the fitting to Δt residual fit.

$q\bar{q}$ background event

For $q\bar{q}$ background event, the following function is used to describe Δt distribution :

$$\begin{aligned} \mathcal{P}^{q\bar{q}}(\Delta t) = &((1-f_\delta)\frac{1}{2\tau}\exp(-\frac{|\Delta t-\mu|}{\tau}) + \delta(\Delta t-\mu)) \\ &\otimes ((1-f_{\text{tail}})G(\Delta t; 0, s_{\text{main}}\delta\Delta t) + f_{\text{tail}}G(\Delta t; 0, s_{\text{tail}}\delta\Delta t)) \end{aligned} \tag{103}$$

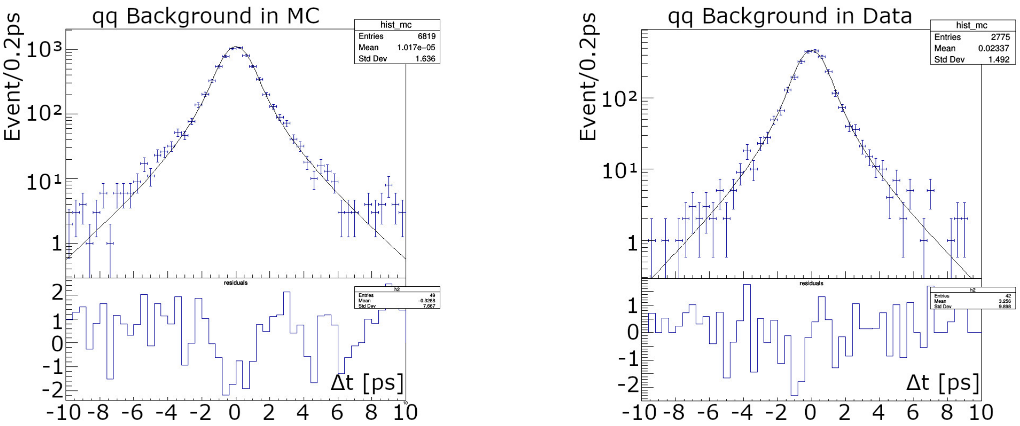
where the double Gaussian is convoluted with the sum of exponentially decaying term and delta function. In this parametrization, $f_\delta, f_{\text{tail}}, \mu, \tau, s_{\text{main}}$ and s_{tail} is floated, and fitted to $1\,\text{ab}^{-1}$ MC or Δt side-band region of $361.6\,\text{fb}^{-1}$ real data. The comparison of fitted $\mathcal{P}^{q\bar{q}}(\Delta t)$ against event distribution for MC and real data are shown in Fig. 46. The fitting result of 6 parameters are shown in Table. 16 and Table. 17.

$B\bar{B}$ background event

For $B\bar{B}$ background event, the same resolution function as signal events is adopted, but with smaller lifetime as

$$\mathcal{P}^{B\bar{B}}(\Delta t) = \frac{1}{2\tau}\exp(-\frac{|\Delta t|}{\tau}) \otimes \mathcal{R}_{\text{sig}}(\delta\Delta t; \sigma_{\Delta t}) \tag{104}$$

where τ is a parameter to be fitted to $1\,\text{ab}^{-1}$ MC. Due to low statistics in data, the Δt modeling for $B\bar{B}$ background is fixed to the value obtained by fitting to MC. The fitting



(a) MC

(b) Data

Figure 46: Fit result of Δt distributions to $q\bar{q}$ continuum events for $1\,\mathrm{ab}^{-1}$ MC dataset (left) and $361.6\,\mathrm{fb}^{-1}$ side-band data(right). The upper figures shows the comparison between event distribution and fitted PDF, and bottom figures shows the pull distrubiton of event distribution from PDF line shape.

	Fit result
f_δ	0.973 ± 0.004
f_{tail}	0.190 ± 0.011
μ [ps]	-0.006 ± 0.003
τ [ps]	2.164 ± 0.141
σ_{main} [ps]	1.243 ± 0.014
σ_{tail} [ps]	3.265 ± 0.111

Table 16: The parameters of Δt distribution function for $q\bar{q}$ background events calibrated for $1\,\mathrm{ab}^{-1}$ MC15 data sample.

to MC dataset is done for $B^0\bar{B}^0$ mixed MC and B^+B^- charged MC respectively, and the fitting parameter τ is obtained to be $\tau_{\mathrm{mixed}} = 1.29 \pm 0.03\,\mathrm{ps}$ and $\tau_{\mathrm{charged}} = 1.31 \pm 0.03\,\mathrm{ps}$ respectively as shown in Fig. 47.

7.2 Lifetime fitting

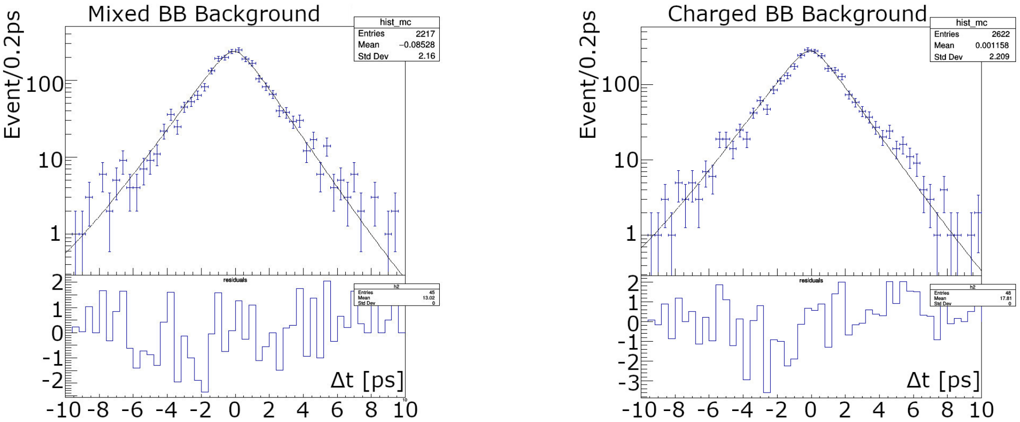
For the lifetime fitting including background events, PDF for maximum likelihood is expressed as

$$\mathcal{P}(\Delta t; \sigma_{\Delta t}, l) = f^l_{\mathrm{sig}} \mathcal{R}_{\mathrm{sig}}(\Delta t; \sigma_{\Delta t}) \otimes \frac{1}{2\tau_{B^0}} \exp(-\frac{|\Delta t|}{\tau_{B^0}}) + f^l_{q\bar{q}} P^{q\bar{q}}(\Delta t) + f^l_{B\bar{B}} P^{B\bar{B}}(\Delta t) \quad (105)$$

where τ_{B^0} is the fitting parameters representing the lifetime of B^0, l is the qr-bin for each event, and $f^l_x(x = \mathrm{sig}, q\bar{q}, B\bar{B})$ is the event fraction obtained by signal extraction fit. The

	Fit result
f_δ	0.935 ± 0.007
f_{tail}	0.160 ± 0.017
μ [ps]	-0.004 ± 0.003
τ [ps]	1.896 ± 0.104
σ_{main} [ps]	1.210 ± 0.019
σ_{tail} [ps]	3.086 ± 0.201

Table 17: The parameters of Δt distribution function for $q\bar{q}$ background events calibrated for 361.6fb^{-1} real data.



(a) $B\bar{B}$ background (b) B^+B^- background

Figure 47: Fit result of Δt distributions to mixed(left) or charged (right) background events for 1 ab^{-1} MC dataset. The upper figures shows the comparison between event distribution and fitted PDF, and bottom figures shows the pull distrubiton of event distribution from PDF line shape.

event fraction f^l_{sig}, $f^l_{q\bar{q}}$ and $f^l_{B\bar{B}}$ are calculated for each event as

$$f^l_x = \frac{N^l_x \mathcal{P}^{\text{ext}}_x(M_{bc}, \Delta E, l)}{\sum_{x=\text{sig},q\bar{q},B\bar{B}} N^l_x \mathcal{P}^{\text{ext}}_x(M_{bc}, \Delta E, l)} \tag{106}$$

where N^l_x is the number of event of each event type in the given qr bin and calculated using the result of signal extraction fitting.

7.2.1 Validation using ToyMC

In order to check the validity of the resolution function for this $B^0 \to K^+K^-K^0_S$ analysis, lifetime fitting using the signal resolution function to a signal only Dalitz MC samples are performed with resolution function parameters calibrated for MC dataset as shown in Table. 15. Here only the B lifetime $\tau^{\text{fit}}_{B^0}$ is floated as a fitting parameter. Then B lifetime is fitted to be $\tau^{\text{fit}}_{B^0} = 1.490 \pm 0.033\,\text{ps}$ as shown in Fig. 48, which is consistent with MC input value $\tau_{B^0} = 1.519 \pm 0.007\,\text{ps}$.

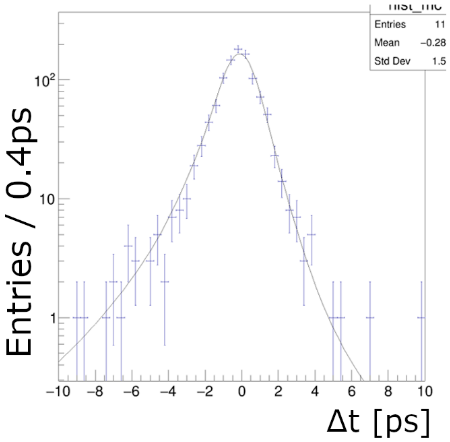


Figure 48: The result of lifetime fitting to signal only MC samples. The line and point represent fitted PDF and event distribution in MC dataset respectively.

Then we also checked lifetime fitting procedure including $q\bar{q}$ and $B\bar{B}$ background using Toy MC. 500 datasets corresponding to $361.6\,\mathrm{fb}^{-1}$ are generated based on the lifetime fitting PDF in Eq. 105, and for each dataset residual and pull distribution of the fitted lifetime τ_{B^0} is calculated as shown in Fig. 49. Those distribution is consistent with MC input values within statistical errors from fitting, and thus the resolution function and Δt modeling is describing distribution of Δt correctly.

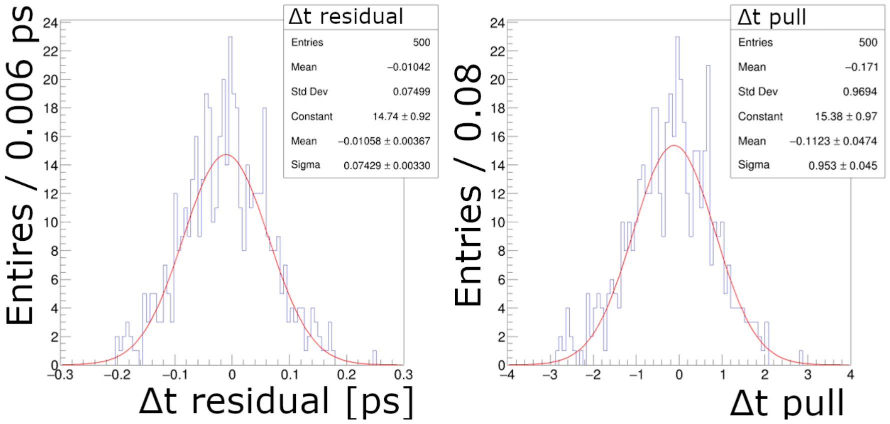


Figure 49: The result of Toy MC study of lifetime fitting including background events. The left figure shows the residual distribution of the lifetime fitting result, and the right figure shows the pull distribution. The red line represents Gaussian function fitted to each distribution.

7.2.2 Lifetime fitting to real data

Lifetime fitting to $361.6\,\mathrm{fb}^{-1}$ real data is performed, where signal resolution function or $q\bar{q}$ distribution is calibrated for data or data side-band respectively. The result of lifetime fitting is shown in Fig. 50 and Table. 18. The B^0 lifetime is obtained to be $1.48 \pm 0.09\,\mathrm{ps}$, which is consistent with PDG average value of B^0 lifetime.

	Fit Result	MC
$\tau_{B^0}[ps]$	1.489±0.093	1.519

Table 18: Lifetime fitting result using $361.6\,\mathrm{fb}^{-1}$ real data

7.3 PDF for CP fitting

The event-by-event PDF for extracting CP asymmetry parameters is defined as following:

$$\begin{aligned} P_{total} = f^l_{sig}(M_{bc}, \Delta E, qr) P_{sig}(m', \theta', \Delta t, qr) + \\ f^l_{q\bar{q}}(M_{bc}, \Delta E, qr) P_{q\bar{q}}(m', \theta', \Delta t, qr) + \\ f^l_{B\bar{B}}(M_{bc}, \Delta E, qr) P_{B\bar{B}}(m', \theta', \Delta t, qr). \end{aligned} \quad (107)$$

where l is the qr-bin, and f^l_{sig}, $f^l_{q\bar{q}}$ and $f^l_{B\bar{B}}$ is defined as the same way with Eq. 106.

7.3.1 Signal event

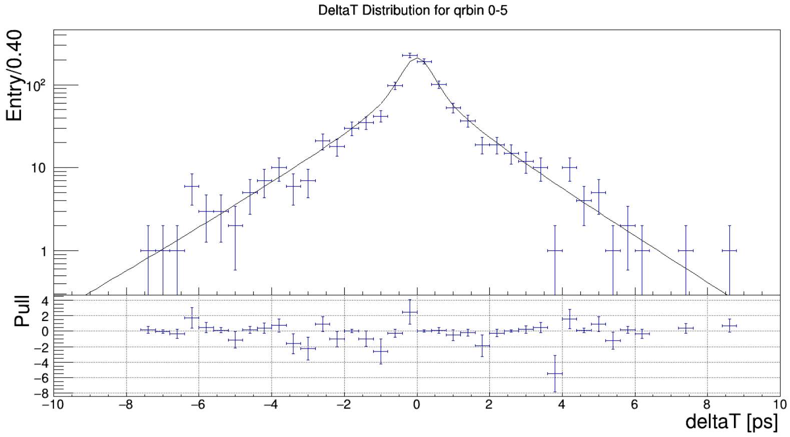
For signal events, PDF is written as

$$\begin{aligned} & P_{sig}(m', \theta', \Delta t, qr) \\ & = \epsilon(m', \theta') |detJ| \left((1 - w_l) \frac{d\Gamma}{d\Delta t}(\Delta t, q_{tag}) + w_l \frac{d\Gamma}{d\Delta t}(\Delta t, -q_{tag}) \right) \\ & \qquad \otimes R_{sig}(\Delta t) \\ & = \epsilon(m', \theta') |detJ| \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}} \left((1 - q_{tag}\Delta w_l) \left(|A|^2 + |\bar{A}|^2 \right) - \right. \\ & \qquad q_{tag}(1 - 2w_l) \left(|A|^2 - |\bar{A}|^2 \right) \cos \Delta m_d \Delta t + \\ & \qquad \left. 2q_{tag}(1 - 2w_l)\,\mathrm{Im}\left(\bar{A} A^* \right) \sin \Delta m_d \Delta t \right) \otimes R_{sig}(\Delta t) \end{aligned} \quad (108)$$

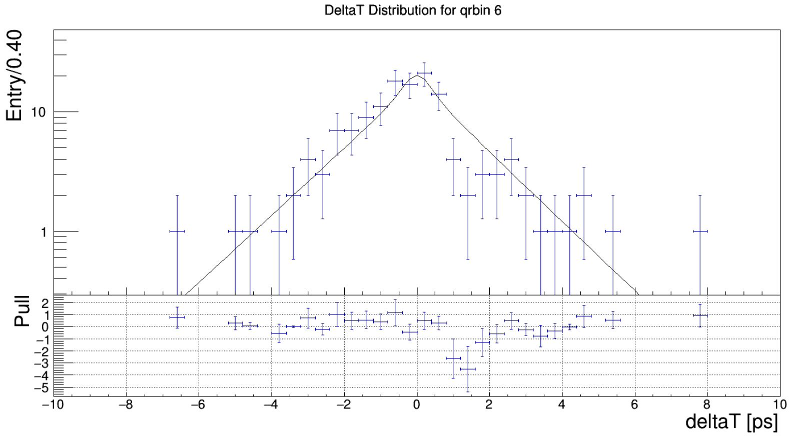
where decay amplitude $|A|^2$ is defined using isobar model as Eq. 62. The term $\epsilon(m', \theta')$ represents the reconstruction efficiency of signal events over square Dalitz plane defined as

$$\epsilon(m', \theta') = \frac{n^{\mathrm{recon}}}{n^{\mathrm{gen}}}. \quad (109)$$

where $\epsilon(m', \theta')$ is implemented as a histogram PDF over square Dalitz plane, and calibrated using 1M Dalitz signal MC samples. The MC samples are generated to be distributing flatly over square Dalitz plane. The square Dalitz plane is then divided into 50 bins for both m' and θ' direction, and the ratio of the number of reconstructed events n_{recon} against the number of generated events n_{gen} is calculated for each bin. The calculated $\epsilon(m', \theta')$ is shown in Fig. 51.



(a) Fitting to events in qr-bin 0-5



(b) Fitting to events in qr-bin 6

Figure 50: The lifetime fitting results of B^0 lifetime to $361.6\,\mathrm{fb}^{-1}$ real data for qr-bin [0-5] (a) and qr-bin [6] (b). The upper figures show comparison between event distribution and fitted Δt PDF, while bottom figures represent pull distribution of event distribution from fitted Δt model.

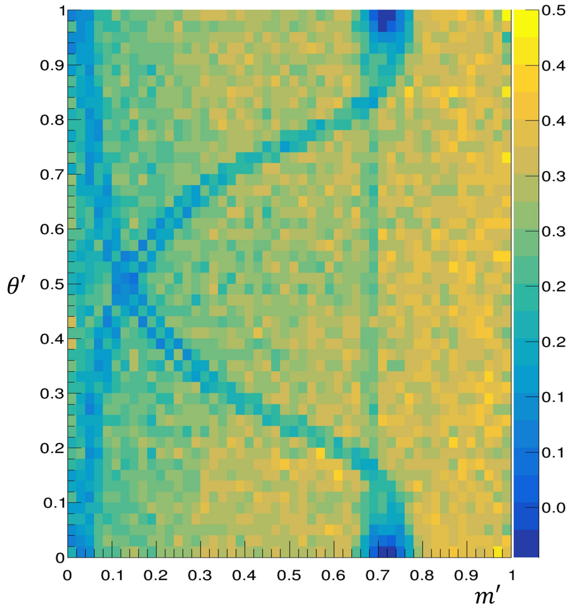


Figure 51: The result of $\epsilon(m', \theta')$ calibrated using 1M Dalitz signal sample. The whole square Dalitz plane is divided into 50 bins for both m' and θ' and implemented as a histogram PDF.

The signal PDF is normalized over square Dalitz plot and Δt, where normalization factor $\mathcal{N}$ is calculated as

$$\mathcal{N} = \int_{\text{sDP}} \int_{\Delta t} \frac{1}{2}(P_{sig}(m', \theta', \Delta t, qr = +1) + P_{sig}(m', \theta', \Delta t, qr = -1)) dm' d\theta' d\Delta t \quad (110)$$

where the normalization for $q_{\text{tag}} = +1$ or $q_{\text{tag}} = -1$ event are set to be equal in order to enhance the sensitivity of PDF against the direct CP asymmetry. In order to calculate the normalization factor efficiently, part of the PDF integration in Eq. 110, which is independent of CP fitting parameters, is calculated beforehand for each event numerically before CP fitting as

$$I_{ij} = \int_{\text{sDP}} \int_{\Delta t} \epsilon(m', \theta') |detJ| \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}} F_i(m', \theta') \bar{F}_j(m', \theta') \otimes R_{sig}(\Delta t) dm' d\theta' d\Delta t, \quad (111)$$

where i and j represents decay channels. Using the result of pre-calculation, the normalization factor $\mathcal{N}$ is then calculated by just linearly combining I_{ij} with CP asymmetry parameters a'_i, which is explained in Eq. 62.

7.3.2 $q\bar{q}$ background event

The PDF for continuum background events are implemented as a histogram PDF written as

$$P_{q\bar{q}} = H_{q\bar{q}}(m', \theta') P_{q\bar{q}}(\Delta t), \quad (112)$$

where $H_{q\bar{q}}(m', \theta')$ is the histogram PDF of the distribution of events over square Dalitz plane with 50 bins each. $H_{q\bar{q}}(m', \theta')$ is calibrated using event distribution in the Dalitz side-band region of the real data defined as $(5.24 < M_{bc} < 5.29\,\mathrm{GeV}/c^2) \wedge (|\Delta E| < 0.1\,\mathrm{GeV})$, which is defined apart from the side-band region used for signal extraction and Δt calibration. The histogram PDF obtained by the side-band data of $361.6\,\mathrm{fb}^{-1}$ real data is shown in Fig. 52.

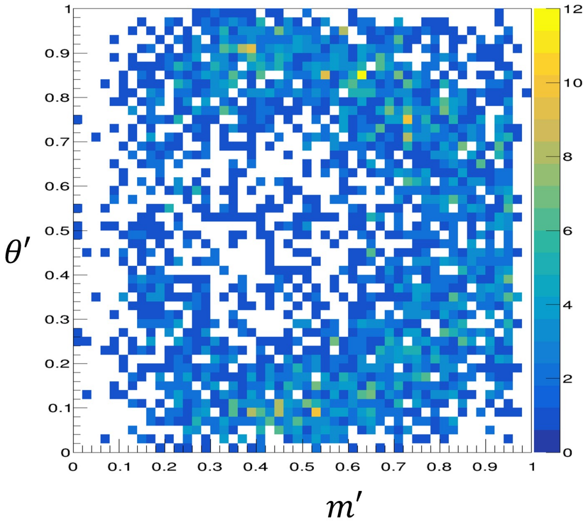


Figure 52: The $H_{q\bar{q}}(m', \theta')P_{q\bar{q}}$ distribution obtained from $361.6\,\mathrm{fb}^{-1}$ real data in the restricted side-band region defined as $(5.24 < M_{bc} < 5.29\,\mathrm{GeV}/c^2) \wedge (|\Delta E| < 0.1\,\mathrm{GeV})$.

The Dalitz side-band is defined as the narrow range closed to the signal region due to the difference in the event distribution on square Dalitz plane between signal region and normal side-band, which is defined as $M_{bc} < 5.26\,\mathrm{GeV}/c^2$. The m' distribution sliced by ΔE is shown in Fig. 53, and there exists large discrepancy of m' distribution between $|\Delta E| > 0.1\,\mathrm{GeV}$ region and $|\Delta E| < 0.1\,\mathrm{GeV}$ region. The m' distribution is also analyzed with real data as shown in Fig. 54, where m' distribution is compared between Dalitz side-band region and normal side-band region. We observed large discrepancy of m' distribution between those two regions especially at m' region where non-resonant component populates, and thus determined to model $q\bar{q}$ background events from only Dalitz side-band region.

7.3.3 $B\bar{B}$ background event

The PDF for $B\bar{B}$ background events are defined as a histogram PDF written as

$$P_{B\bar{B}} = f_{B^+B^-}H_{B^+B^-}(m', \theta')\frac{1 + q_{tag}A^l_{B^+B^-}(\theta')}{2}P_{B^+B^-}(\Delta t) + f_{B^0\bar{B}^0}H_{B^0\bar{B}^0}(m', \theta')\frac{1 + q_{tag}A^l_{B^0\bar{B}^0}(\theta')}{2}P_{B^0\bar{B}^0}(\Delta t), \quad (113)$$

where $H_{B^+B^-}$ or $H_{B^0\bar{B}^0}$ is the histogram PDF of the background event distribution originating from B^+B^- charged pairs or $B^0\bar{B}^0$ mixed pairs. The histogram PDF is divided into 20 bins in each m' and θ'. The events distributions from $1\mathrm{ab}^{-1}$ MC dataset are shown

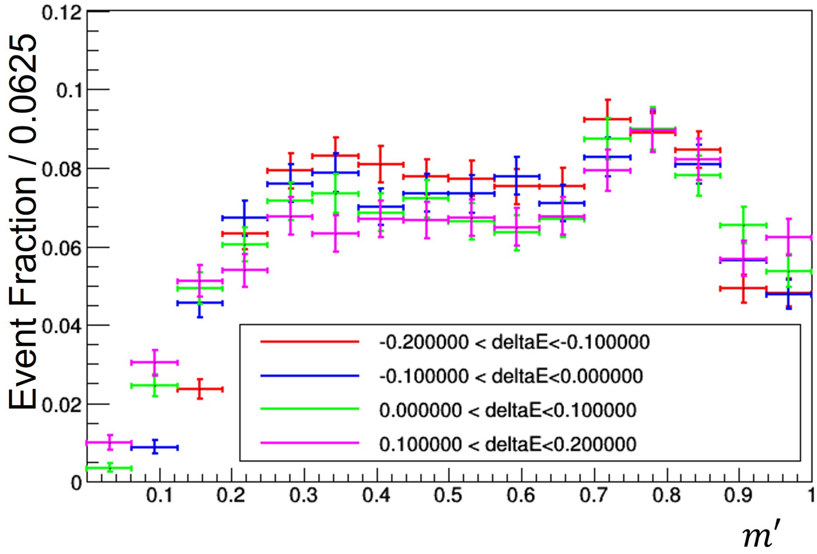


Figure 53: The scaled m' distribution for $q\bar{q}$ background events in MC dataset sliced by ΔE, where each color represents scaled m' distribution within each of ΔE region having edge at $\Delta E = [-0.2, 0.1, 0, 0.1, 0.2]\,\mathrm{GeV}$.

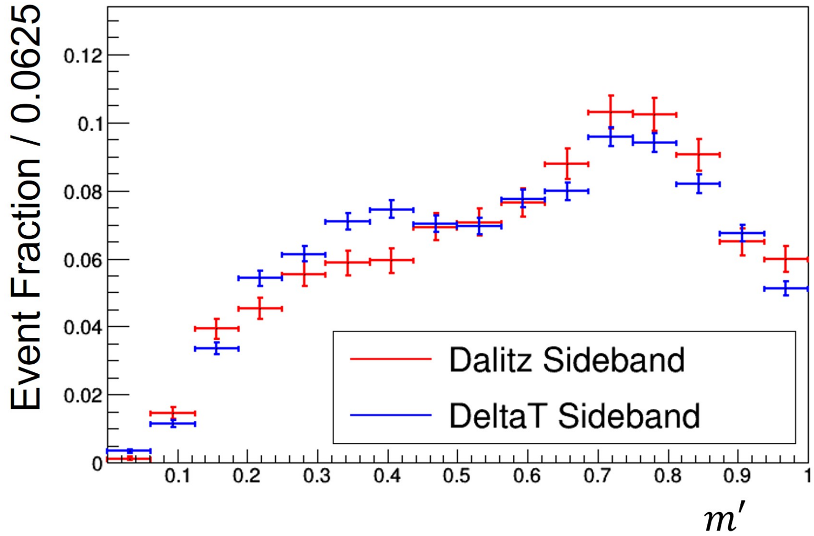


Figure 54: The scaled m' distribution for Dalitz side-band (blue) defined as $(5.24 < M_{bc} < 5.29\,\mathrm{GeV}/c^2) \wedge (|\Delta E| < 0.1\,\mathrm{GeV})$ or Δt side-band (red) defined as $M_{bc} < 5.26\,\mathrm{GeV}/c^2$ in real data.

in Fig. 55 and Fig. 56. The histogram PDF for the CP fitting is then implemented as

$$H_{BB}(m',\theta') = H^{MC}_{BB}(m',\theta') + H^{MC}_{BB}(m',1-\theta') \tag{114}$$

in order to compensate for low statistics and to have the distribution over square Dalitz plot symmetry over θ', because the asymmetry along θ' is taken into account by the parameter $A^l_{B^+B^-}$ and $A^l_{B\bar{B}}$ in Eq. 113. $f_{B^+B^-}$ or $f_{B^0\bar{B}^0}$ is the fraction of charged or mixed background component respectively, and is fixed to $f_{B^+B^-} = 0.5724$ or $f_{B^0\bar{B}^0} = 0.4276$, which is obtained from 1ab^{-1} MC sample. $A_{B^+B^-}$ or $A_{B^0\bar{B}^0}$ describes the asymmetry of event distribution over square Dalitz plane according to the qr bin, and calculated as

$$A^l_{BB}(\theta') = \frac{n_{qr>0} - n_{qr<0}}{n_{qr>0} + n_{qr<0}}. \tag{115}$$

The θ' distribution in each qr bin of B^+B^- or $B\bar{B}$ background component is shown in Fig. 57 and Fig. 58, and each distribution is fitted by a linear function as $A^l_{BB}(\theta') = x_i(2\theta'-1)$. Events for high qr bins tends to show large θ' distribution asymmetry due to BB background such as particle mis-identification of $\pi^\pm$ to $K^\pm$. The effect of the small statistics of BB background on the determination of PDF parameters will be taken into the systematic uncertainty. Those PDF parameters for $B\bar{B}$ background is fixed to the values obtained from 1ab^{-1} generic MC sample due to the low statistics in the 361.6fb^{-1} real data sample.

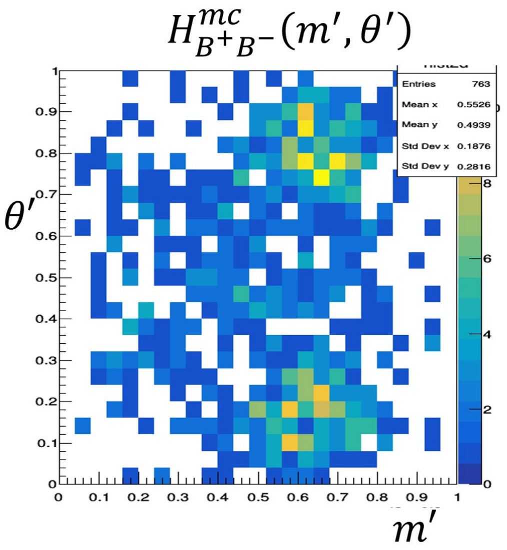


Figure 55: The $H_{B^+B^-}(m',\theta')$ distribution obtained from 1 ab^{-1} generic MC dataset.

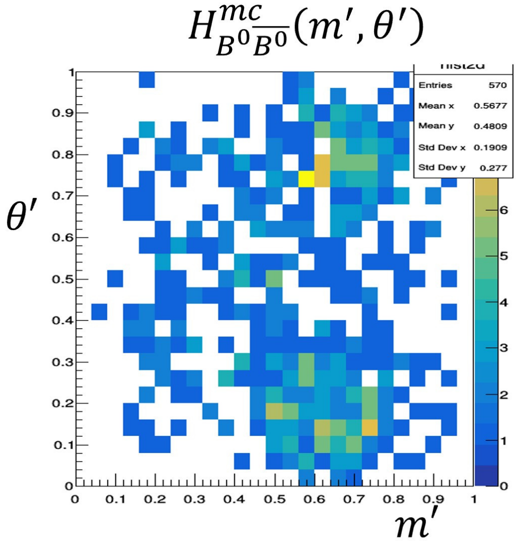


Figure 56: The $H_{B^0\bar{B}^0}(m', \theta')$ distribution obtained from $1\,\mathrm{ab}^{-1}$ generic MC dataset.

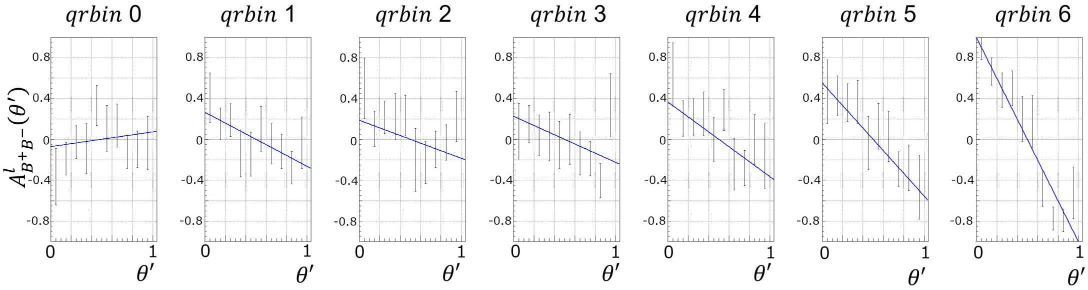


Figure 57: The $A_{B^+B^-}(m', \theta')$ distribution for each qr bin obtained from $1\,\mathrm{ab}^{-1}$ generic MC dataset. Each figure shows the raw asymmetry of the number of events with flavor tagging $qr > 0$ or $qr < 0$ in each θ' bin. Each figure is fitted by a 1 dimension polynomial and used for CP fitting PDF.

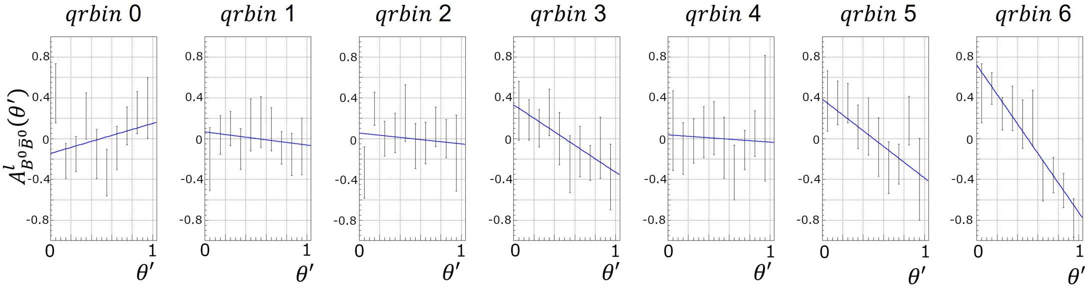


Figure 58: The $A_{B^0\bar{B}^0}(m', \theta')$ distribution for each qr bin obtained from $1\,\mathrm{ab}^{-1}$ generic MC dataset. Each figure shows the raw asymmetry of the number of events with flavor tagging $qr > 0$ or $qr < 0$ in each θ' bin. Each figure is fitted by a 1 dimension polynomial and used for CP fitting PDF.

7.4 CP fitting procedure

In the full PDF as written in Eq. 107, there exists $4 \times 7 = 28$ parameters describing decay amplitude and event distribution over square Dalitz plane and Δt, as described in Eq. 77. Among those parameters, Dalitz CP asymmetry parameter c_i and d_i for four decay channels, for $f_X K_S^0$,$K_S^0(K^+K^-)_{NR}$, $K^-(K^+K_S^0)_{NR}$ and $K^+(K^-K_S^0)_{NR}$ respectively, are merged into one parameter denoted as c_{others} and d_{others}. CP asymmetry parameters for $\chi_{c0}K_S^0$ is fixed to the world average of PDG value where $c_i = 0$ and $d_i = 21.5\,\mathrm{deg}$. The amplitude a_i and the phase difference b_i is relative parameters, and thus $a_{\phi K_S^0}$ and $b_{\phi K_S^0}$ is fixed as a reference for amplitudes and phases for other decay channels. The list of fixed or floated parameters are shown in Table. 19.

The CP fitting is then performed using unbinned maximum likelihood by minimizing log-likelihood $-2\ln\mathcal{L} = -2\sum_i \ln(P_{\mathrm{total}}(\vec{x}))$, where i runs over all event candidates in the signal region.

Due to the overlapping of multiple resonances over square Dalitz plane , there are expected to be multiple local minimums obtained by CP fitting. In order to find best local minimum, fitting to a single dataset is performed from random initial values of Dalitz fitting parameters repeatedly, and adopt local minimums to where fit result is converged from multiple random initial values of fitting parameters.

	a_i	b_i	c_i	d_i
ϕK_S^0	fix	fix	float	float
$f_0 K_S^0$	float	float	float	float
$\chi_{c0} K_S^0$	float	float	fix	fix
$f_X K_S^0$	float	float	float	float
$K_S^0(K^+K^-)_{NR}$	float	float		
$K^+(K_S^0K^-)_{NR}$	float	float		
$K^-(K_S^0K^+)_{NR}$	float	float		

Table 19: The list of parameters used for CP fitting.

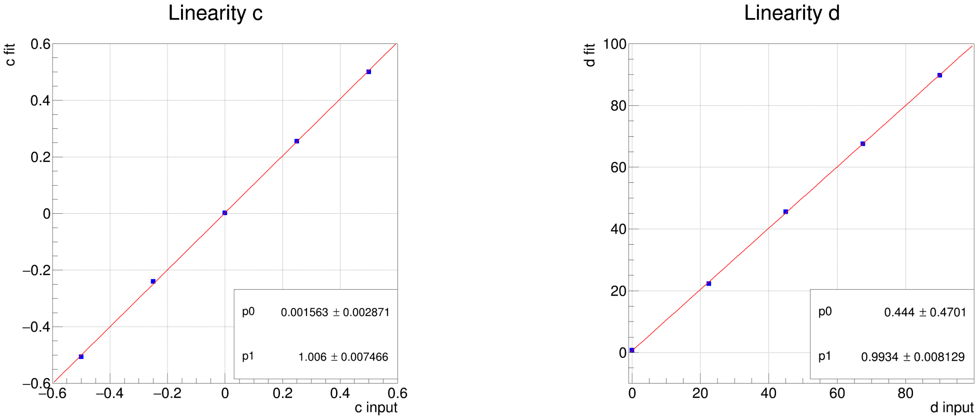
7.5 CP fitting validation using Toy MC

In order to validate the CP fitting procedure described above, validation of fitting procedure using Toy MC dataset are performed. In order to check the linearity between input value of a Dalitz fitting parameter $c_i^{\mathrm{input}}, d_i^{\mathrm{input}}$ and fit result $c_i^{\mathrm{fit}}, d_i^{\mathrm{fit}}$, toy MC dataset are generated based on the CP fitting PDF.

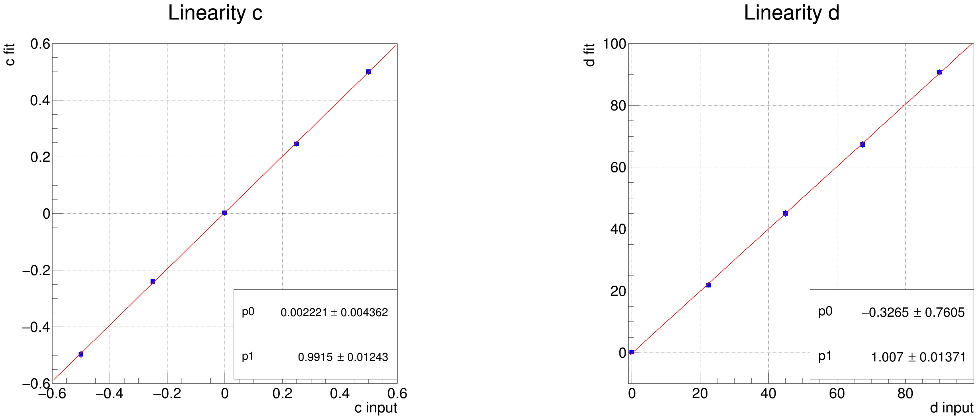
In the Toy MC dataset, the set of Dalitz CP violating parameters for $B^0 \to \phi K_S^0$ decay channels $(c_{\phi K_S^0}, d_{\phi K_S^0})$ are set to $(c_{\phi K_S^0} = 0, d_{\phi K_S^0} = [0, 22.5, 45.0, 67.5, 90])$ for $d_{\phi K_S^0}$ linearity test, and $(c_{\phi K_S^0} = [-0.5, -0.25, 0, 0.25, 0.50], d_{\phi K_S^0} = 32.2^\circ)$ for $c_{\phi K_S^0}$ linearity test. Here the value of $d_{\phi K_S^0} = 32.2^\circ$ is quoted from the result of previous Belle analysis [31]. For each set of $(c_{\phi K_S^0}, d_{\phi K_S^0})$, 1000 individual datasets corresponding to $361.6\mathrm{fb}^{-1}$ integrated luminosity are generated, and the CP fitting is performed to obtain the fit result distribution of $(c_{\phi K_S^0}, d_{\phi K_S^0})$.

The same linearity check is also conducted for both $(c_{f_0 K_S^0}, d_{f_0 K_S^0})$ and $(c_{\mathrm{others}}, d_{\mathrm{others}})$, where $d_{f_0 K_S^0}$ is fixed to 30.5° for $c_{f_0 K_S^0}$ linearity test and d_{others} is fixed to 24.4° for c_{others} linearity test quoted from the previous analysis.

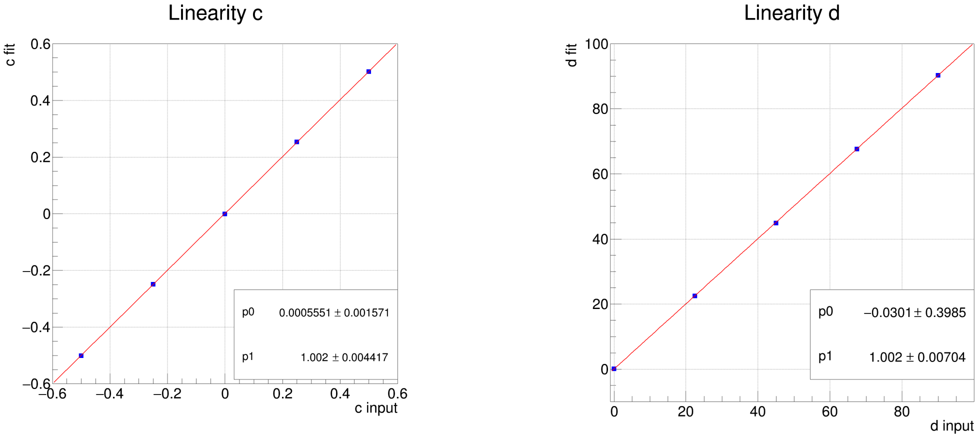
The result of the linearity check using toy MC is shown in Fig. 59, and no obvious deviation of the fit result of CP violating parameters from input value are observed. The possible fit bias due to the CP fitting procedure or low statistics in real data will be discussed in the systematic uncertainty section.



(a) $c_{\phi K_S^0}$ (b) $d_{\phi K_S^0}[^\circ]$



(c) $c_{f_0 K_S^0}$ (d) $d_{f_0 K_S^0}[^\circ]$



(e) c_{other} (f) $d_{\text{other}}[^\circ]$

Figure 59: The result of linearity validation using toy MC datasets for ϕK_S^0, $f_0 K_S^0$ and other. Horizontal and vertical axes represents Dalitz CP violating parameters use for the generation of Toy MC dataset and fitted value of CP violation parameters respectively. Each point represents the mean and its error of the fit result distribution obtained by Gaussian fitting.

8 Systematic Uncertainties

In this section, the possible systematic uncertainty is discussed. In the discussion below, the systematic uncertainty of the measurement of Dalitz CP violating parameters c_i, d_i is discussed first, and then the uncertainty of c_i and d_i is converted to that of $\phi_1(i)$ and $A_{\rm CP}(i)$ followed by the relation as shown in Eq. 79 and Eq. 80, where $i = \phi K_S^0, f_0 K_S^0$, and other.

The expected sources of systematic uncertainties are listed in Table. 20. The total systematic uncertainty is calculated as the square root of the squared individual systematic uncertainties.

Parameter	$\delta\phi_1$			δA_{CP}		
Decay Mode	ϕK_S^0	$f_0 K_S^0$	others	ϕK_S^0	$f_0 K_S^0$	others
Analysis Model						
Fit bias	0.74	0.64	0.27	0.015	0.021	0.005
Observable correlation	0.15	0.24	0.13	0.007	0.006	0.006
Fixed Parameters	0.82	0.98	0.75	0.019	0.021	0.014
$\tau_d, \Delta m_d$	0.44	0.40	0.35	0.008	0.007	0.009
$B\bar{B}$ background	0.17	0.46	0.38	0.011	0.022	0.015
$q\bar{q}$ background	0.10	0.21	0.19	0.004	0.006	0.009
Multiple Candidate	0.15	0.12	0.12	0.004	0.003	0.003
Tag-side interference	0.54	0.27	0.42	0.004	0.005	0.005
Δt measurement						
Detector misalignment	0.19	0.21	0.17	0.009	0.008	0.006
Momentum scale	0.02	0.02	0.02	0.001	0.001	0.001
Beam spot	0.09	0.09	0.09	0.002	0.002	0.002
Kinematic Approximation	0.33	0.27	0.30	0.007	0.006	0.006
Dalitz Model						
Non-resonant modeling	0.79	1.22	1.16	0.013	0.019	0.021
Possible resonance	0.50	2.04	0.68	0.004	0.033	0.010
PDF binning	0.18	0.20	0.16	0.005	0.004	0.003
Reconstruction Efficiency	0.03	0.02	0.02	0.000	0.000	0.000
Total	1.69	2.79	1.76	0.035	0.055	0.036

Table 20: Summary Table of all the systematic uncertainty expected in this analysis.

8.1 Systematic for analysis models

8.1.1 Fit bias

The systematic uncertainty due to the CP fitting procedure is estimated using Toy MC study conducted in the Section 7.5. In order to take the possible fit bias due to the low statistics into account, the statistics of each Toy MC dataset was set to be equal with $361.6 fb^{-1}$ integrated luminosity. For each plot in Fig. 59, the mean of the fit result residual μ_i is calculated by Gaussian fitting. Then μ_i is fitted by a linear function as shown in Fig. 60. The systematic uncertainty due to the fit bias is then estimated as the largest deviation of the linear function from zero within the definition range of each fitting parameters, where the definition range of Dalitz CP violating parameter is $c_i = [-1, 1]$ and $d_i = [-45°, 135°]$ in this case.

8.1.2 Observable correlation

The systematic uncertainty due to the correlation between observable is estimated. In the parametrization of observable distribution in signal extraction fitting and CP fitting, there exists possible correlation between observable which is not included into PDF parametrization as shown in Appendix A. In order to estimated the possible effect of observable correlation such as $M_{bc} - \Delta E$ correlation in the signal event, fully generated signal MC dataset, where possible correlation between observable are simulated, are generated 500 times and signal extraction fitting and CP fitting procedures are performed again. The systematic is then estimated as the mean of the residual distribution for each CP violating parameter.

8.1.3 Fixed parameters

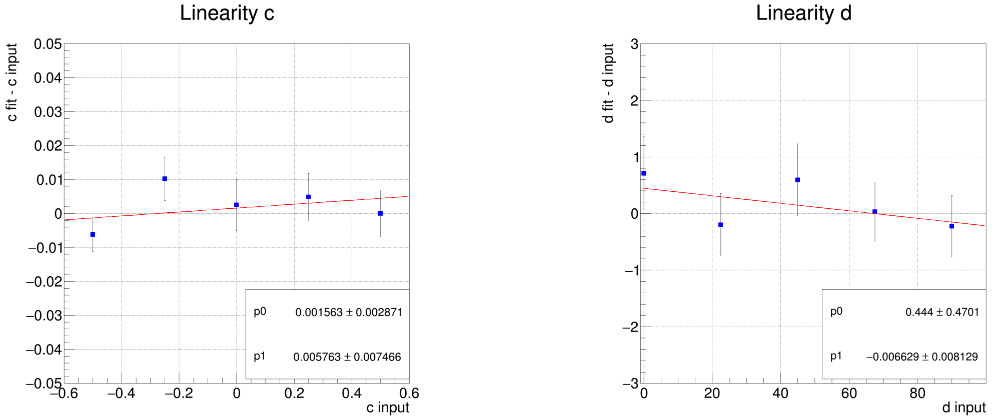
The following parameters are fixed by MC datasets or fitting to side-band data, in the process of signal extraction fitting and CP fitting.

- Signal extraction fit : $M_{bc}, \Delta E$ shape for $q\bar{q}$ and $B\bar{B}$ background events, $M_{bc}, \Delta E$ for signal event.
- Δt modeling : Δt shape for $q\bar{q}$ and $B\bar{B}$ background events, the resolution function for signal event.
- Wrong tag fraction : $w, \Delta w$

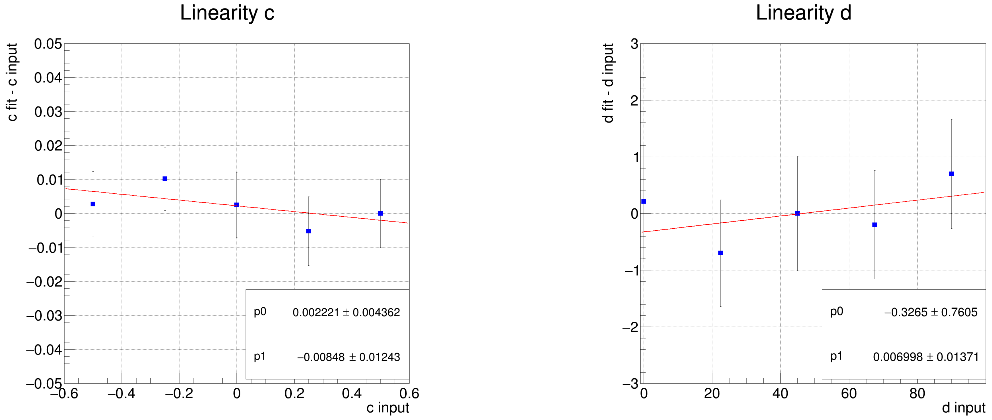
Those fixed parameters are divided into three category as signal extraction fit parameters, Δt modeling parameters, and wrong tag fraction parameters. In order to estimate uncertainty due to fixed parameters, the parameters for each category are fluctuated simultaneously based on the covariance matrix of the fitting result for real data. We repeated signal extraction fitting and CP fitting procedure with different fixed parameters against a same toy dataset, and took a standard deviation of the fitting result c_i, d_i as a systematic uncertainty due to fixed parameters. The result of systematic uncertainty for each category is shown in Table. 21.

8.1.4 $\tau_d, \Delta m_d$

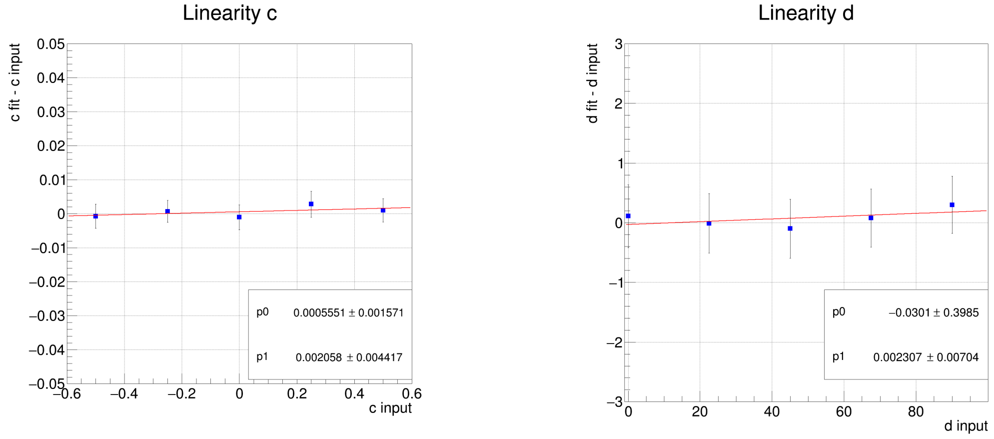
The systematic uncertainty due to the uncertainty of the lifetime and mixing frequency τ_d and Δm_d, which is fixed to the PDG average value in the CP fitting procedure, is taken into consideration. 500 datasets are generated with different τ_d and Δm_d, which is fluctuated from PDG average by the statistical error of PDG values[37], and CP fitting is performed for those datasets with the nominal τ_d and Δm_d values. The systematic uncertainty is estimated by the standard deviation of the distribution of fitted c_i and d_i.



(a) $c_{\phi K_S^0}$ (b) $d_{\phi K_S^0}[^\circ]$



(c) $c_{f_0 K_S^0}$ (d) $d_{f_0 K_S^0}[^\circ]$



(e) c_{other} (f) $d_{\text{other}}[^\circ]$

Figure 60: The result of the fit bias estimation using toy MC datasets for ϕK_S^0, $f_0 K_S^0$ and other. Horizontal and vertical axes represents CP violating parameters use for the generation of Toy MC dataset and the residual of CP violation parameters respectively. Each point represents the mean of the residual distribution and its error estimated by Gaussian fitting.

Parameter	δc_i			δd_i		
Decay Mode	ϕK_S^0	$f_0 K_S^0$	others	ϕK_S^0	$f_0 K_S^0$	others
Signal Extraction Fit	0.30	0.33	0.30	0.005	0.005	0.003
Δt modeling	0.49	0.65	0.40	0.006	0.006	0.004
Wrong tag fraction	0.59	0.66.	0.56	0.006	0.008	0.004
Total	0.82	0.98	0.75	0.009	0.011	0.007

Table 21: The systematic uncertainties for each category of fixed parameters and total systematic uncertainty in terms of Dalitz CP violating parameters c_i and d_i.

8.1.5 $B\bar{B}$ background

The systematic uncertainty due to the modeling of $B\bar{B}$ background are estimated. In the PDF modeling of $B\bar{B}$ background, the qr bin dependency of square Dalitz plane distribution is modeled using $1\,\mathrm{ab}^{-1}$ MC dataset, but mis-modeling of CP characteristics might lead to the bias in the CP fitting result. CP fitting to a $1\mathrm{ab}^{-1}$ dataset are performed 500 times with fluctuated CP asymmetries parameters for $B\bar{B}$ background modeling described in Eq. 115. The systematic uncertainty is then estimated by the standard deviation of the fitting result distribution.

8.1.6 $q\bar{q}$ background

As stated in Sec. 7.3.2, the $q\bar{q}$ background is modeled by only events in Dalitz side-band region, but $q\bar{q}$ background modeled by other region might results in the difference in the CP fitting result, and thus we included $q\bar{q}$ background modeling into systematic uncertainty. To estimate the systematic uncertainty, we performed CP fitting to the real data with different $q\bar{q}$ model constructed from region $(5.20 < M_{bc} < 5.29\,\mathrm{GeV}/c^2) \wedge (|\Delta E| < 0.1\,\mathrm{GeV})$, where only M_{bc} region is enlarged from the Dalitz side-band. The systematic uncertainty is then estimated as the difference in the CP fitting result.

8.1.7 Multiple Candidate

In order to estimated the effect of best candidate selection, CP fitting for the real data with or without multiple candidate selection are performed. The systematic uncertainty is estimated by the difference of CP fitting result.

8.1.8 Tag-side interference (TSI)

The interference between CKM-flavored decay and doubly-CKM-suppressed decay affect the measurement of CP violating parameters when flavor tagging is performed against coherent B meson pairs originating from $\Upsilon(4S)$[30]. The systematic uncertainty due to the TSI is estimated by assuming all the B_{tag} being tagged from hadronic decay modes, where the effect of TSI on the measurement of CP violating parameter is maximized[19]. We first calculated the systematic uncertainty of S_{CP} and A_{CP} in the quasi-two-body approach, and then converted the uncertainty of S_{CP} and A_{CP} back to the uncertainty in terms of ϕ_1 for each of ϕK_S^0,$f_0 K_S^0$ and other component respectively.

8.2 Systematics for Δt measurement

8.2.1 Detector misalignment

In order to estimate the effect of possible detector alignment, 1M signal MC samples with the nominal detector alignment and 4 possible misaligned scenario are generated respectively. CP fitting are performed to those 5 signal only datasets. The systematic uncertainty is calculated as the maximum deviation of the CP fitting result for mis-aligned scenarios from the nominal detector alignment.

8.2.2 Momentum scale and Beam spot

Momentum of the reconstructed tracks are calibrated for every bucket, and the variation of momentum is covered by ±0.1% envelope. However, it is possible that those momentum have shift from truth value, and thus systematic uncertainty due to the shift in track momentum is taken into account.

Apart from that, the beam spot parameters are also measured periodically, but there exist several sources for systematic uncertainty due to beam spot:

- Statistical uncertainty of beam spot measurement with $ee \to \mu\mu$ events
- Fluctuation of the beam spot within the interval of beam spot calibration
- Systematic uncertainty due to the beam spot calibration method.

As a systematic uncertainty for the momentum scale and beam spot, the evaluation of those systematics in $B^0 \to J/\psi K_S^0$ analysis [14] is quoted to this analysis.

8.2.3 Kinematic approximation

The effect of the conversion from Δt to Δl is taken into the consideration using MC study. CP fitting to a signal only MC dataset are performed with or without $\Delta t \to \Delta l$ approximation as described in Eq. 86, and the difference of CP fitting result are calculated as the systematic uncertainty.

8.3 Systematic uncertainty from Dalitz modeling

8.3.1 Non-resonant modeling

In the current analysis, the distribution of non-resonant $B^0 \to K^+K^-K_S^0$ component are empirically modeled with exponentially decaying model[31], but the possible difference between modeling of non-resonant component and real data lead to a bias in the CP fitting result. In order to estimate the effect of non-resonant component modeling onto

CP asymmetries measurement,, the different non-resonant component modeling are implemented into CP fitting. The parameters for the modeling are quoted from the previous BaBar analysis [29] where non-resonant component is modeled as :

$$A_{CP}^{NR}(s_{12}, s_{23}) = (a_{S_0} + a_{S_1}x + a_{S_2}x^2) + (a_{P_0} + a_{P_1}x + a_{P_2}x^2)P_1(\cos\theta_3), \tag{116}$$

where $x \equiv m_{12} - \Omega$ and Ω is an offset defined as

$$\Omega \equiv \frac{1}{2}(m_{B^0} + \frac{1}{3}(m_{K^0} + 2m_{K^\pm})) \tag{117}$$

and P_1 is the first Legendre polynomial. The s_{12} and s_{23} denotes the squared invariant mass of two Kaons. The fitting to the real data is performed with those 2 different non-resonant component, and the difference of CP fitting result are accounted for as the systematic uncertainty due to the non-resonant component modeling.

8.3.2 Possible resonance

ϕK_S^0, $f_0 K_S^0$, $f_X K_S^0$, and $\chi_{c_0} K_S^0$ are included as resonances for this analysis, but there exist possible resonances as is listed in Table. 22, which is expected to overlap against $f_0 K_S^0$ resonance and non-resonant component in the square Dalitz plane. In order to estimate the effect of the minor resonances which is not included in this analysis on the CP fitting result, $f_0(1710)K_S^0$ and $f_2'(1525)K_S^0$ are additionally considered into the CP fitting procedure. The systematic uncertainty is calculated as the difference of CP fitting result to the real-data with or without the additional minor resonances.

Resonance	Parameters (MeV)	Line Shape
$f_0(1710)$	$M = 1704 \pm 12$ $\Gamma = 123 \pm 18$	RBW
$f_2'(1525)$	$M = 1517.4 \pm 2.5$ $\Gamma = 86 \pm 5$	RBW

Table 22: Possible Resonances to be included in the estimation of systematic uncertainty. The parameters for those resonances are cited from Ref. [31]

8.3.3 PDF binning

$q\bar{q}$ and $B\bar{B}$ background distribution are implemented as PDF histograms over square Dalitz plane. The number of binning of those histogram PDF will affect CP fitting result, and thus need to be taken into systematic uncertainty. In order to estimate this uncertainty, the number of binning for $q\bar{q}$ and $B\bar{B}$ background are changed from 50 to 20 for $q\bar{q}$ histogram PDFs, and $B\bar{B}$ from 25 to10 respectively. Then CP fitting for a MC dataset are performed and the systematic uncertainty is calculated as the maximum difference of CP fitting result between different binning of histogram PDFs.

8.3.4 Reconstruction Efficiency

The systematic uncertainty due to the different K selection efficiency between MC and real data needs to be included due to the efficiency term $\epsilon_{ij}(m', \theta')$, which is a component of signal PDF. For $K^{\pm}$ systematic uncertainty, the $K^{\pm}$ selection efficiency $\epsilon^{data}_{K^{\pm}}(p)$,$\epsilon^{MC}_{K^{\pm}}(p)$ for cut condition kaonID > 0.6 are calculated for both real data and MC, and the ratio of the selection efficiency $r_{K^{\pm}}(p) = \epsilon^{data}_{K^{\pm}}(p)/\epsilon^{MC}_{K^{\pm}}(p)$ are calculated. The selection efficiency and the selection efficiency ratio for $K^{\pm}$ is shown in Fig. 61.

For K^0_S systematic uncertainty, the K^0_S selection efficiency $\epsilon^{data}_{K^0_S}$,$\epsilon^{MC}_{K^0_S}$ using KsSelector with mvaV0 > 0.90 and mvaLambda > 0.11 cut criteria are calculated using the real data and $1\,\mathrm{ab}^{-1}$ MC samples, and the ratio of the K^0_S selection efficiency between data/MC is also calculated as $r_{K^0_S}(p)$. The selection efficiency and the selection efficiency ratio between data and MC for K^0_S is shown in Fig. 62.

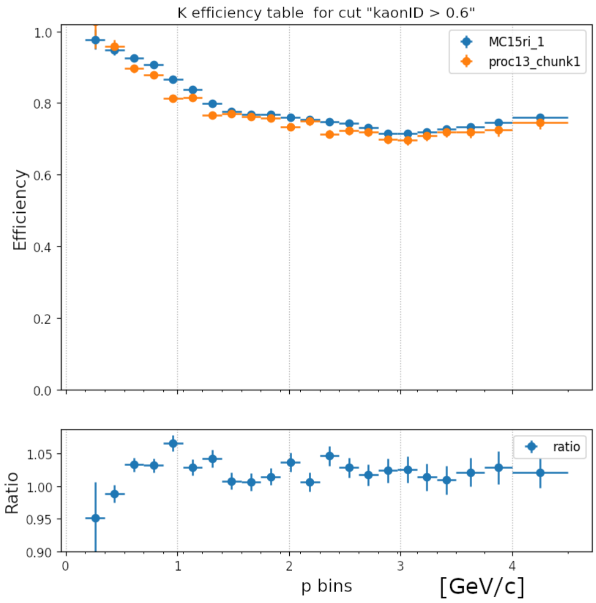


Figure 61: The upper figure shows the relation between $K^{\pm}$ momentum and $K^{\pm}$ reconstruction efficiency for MC (blue) and data (orange). The lower figure shows the ratio of the reconstruction efficiency data/MC for each momentum bin.

Those K selection efficiency ratios depends on its momentum, and thus each bin value of histogram PDF $\epsilon_{ij}(m', \theta')$ is fluctuated by Gaussian distribution with standard deviation σ_{ij} given as

$$\sigma_{ij} = \sqrt{(1 - r_{K^{\pm}}(p_{K^{+}}))^2 + (1 - r_{K^{\pm}}(p_{K^{-}}))^2 + (1 - r_{K^0_S}(p_{K^0_S}))^2} \tag{118}$$

where $p_{K^+}, p_{K^-}, p_{K^0_S}$ are the momentum of K with the given square dalitz parameters (m', θ'). CP fitting is performed to a same dataset with fluctuated efficiency term $\epsilon_{ij}(m', \theta')$ 500 times, and the standard deviation of the CP fitting result are calculated as the systematic uncertainty.

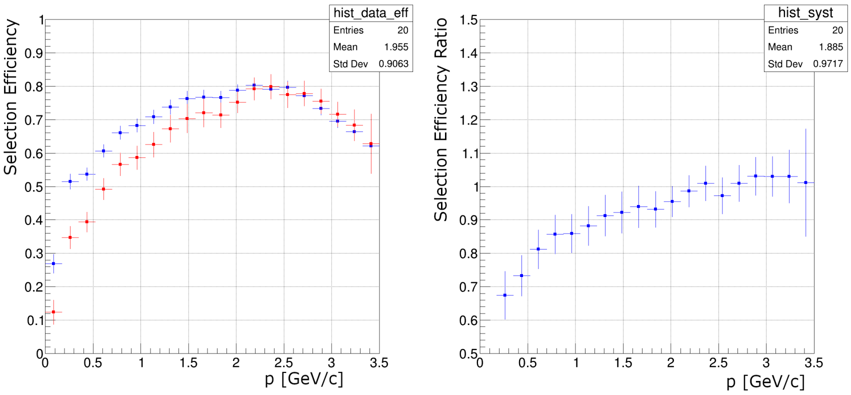


Figure 62: The left figure shows the relation between K_S^0 momentum and K_S^0 reconstruction efficiency for MC (blue) and data (red). The right figure shows the ratio of the reconstruction efficiency data/MC for each momentum bin.

9 Result and Discussion

In this section, we discuss the result of CP fitting result applied for real data and show constraint on new physics and future prospect of the measurement of CP asymmetries in $B^0 \to K^+K^-K_S^0$ three body decay.

9.1 Procedure for CP fitting

In order to find best solution, we performed CP fitting multiple times, with each of CP fitting starting from random initial values for the fitting parameters. The range of random values are floated within $a_i = [0.1, 10]$, $b_i = [-180°, 180°]$, $c_i = [-0.5, 0.5]$ and $d_i = [0°, 180°]$ respectively. Then a local minimum to which CP fittings converge multiple times are chosen to be a solution. As a result of 500 successful CP fitting from different initial parameters, we found a best local minimum in terms of log-likelihood, which is 9 unit better in terms of $-2\ln\mathcal{L}$ than the second best local minimum. We also found 2nd, 3rd and 4th best solution. We compared $-2\ln\mathcal{L}$ of each local minimum to determine best fitting result among multiple local minimums and to observe discrepancy between local minimums.

9.2 Result of CP asymmetries fitting

The result of the best CP fitting result is shown in Table. 23. Results of the 2nd, 3rd and 4th best local minimums found in this analysis is also summarized in Appendix C. In terms of the amplitude term a_i in the CP fitting, the four solutions are consistent for $B^0 \to \phi K_S^0$, $B^0 \to \chi_{c0} K_S^0$ and $B^0 \to (K^\pm K_S^0)_{NR} K^\mp$ amplitude, but the large fluctuation in the a_i for $B^0 \to f_0 K_S^0$, $B^0 \to f_X K_S^0$ and $B^0 \to (K^+K^-)_{NR} K_S^0$ is observed due to the large overlapping of the event distribution for those three channels in square Dalitz Plane. In particularly, the amplitude $a_{(K^+K^-)_{NR}K_S^0}$ varies from 2.51 to 7.29 between solutions. For the direct CP violating term c_i, $c_{\phi K_S^0}$ is close to zero for all solutions, but $c_{f_0 K_S^0}$ and c_{others} showed different fitting result for each local minimums due to the interference. For the mixing-induced CP violating term d_i, the result of CP fitting result is consistent between four solutions due to the large statistic error, and all showed preference of ϕ_1 against $90° - \phi_1$.

9.2.1 The difference of log-likelihood between minimums

In order to analyze the discrepancy and significance of local minimums found in this analysis, a study with ToyMC is conducted. We generated 1000 datasets corresponding to current statistics or 10 times larger statistics respectively based on the CP fitting PDF with parameters fixed to that of the best local minimum. Then we performed CP fitting with initial values of fitting parameters set to the values for 1st, 2nd, 3rd and 4th local minimum respectively, and compared the difference of $-2\ln\mathcal{L}$. Fig. 63a shows the difference of $-2\ln\mathcal{L}$ between local minimums for the current statistics, with Gaussian fitted for each distribution as shown in Table. 24. From the fitted parameters for the likelihood difference between 1st and 2nd local minimum, 1st minimum is estimated to be distinguished from 2nd minimum with over 2σ significance, and thus we selected 1st

Parameter	Fit result
$a_{f_0 K_S^0}$	3.438 ± 0.506
$a_{\phi K_S^0}$	1.000
$a_{f_X K_S^0}$	0.425 ± 0.0612
$a_{\chi_{c0} K_S^0}$	0.245 ± 0.022
$a_{(K^+K^-)_{NR}}$	7.294 ± 0.544
$a_{(K_S^0 K^+)_{NR}}$	4.200 ± 0.224
$a_{(K_S^0 K^-)_{NR}}$	3.630 ± 0.217
$b_{f_0 K_S^0}[^\circ]$	92.05 ± 9.73
$b_{\phi K_S^0}[^\circ]$	-33.48
$b_{f_X K_S^0}[^\circ]$	66.91 ± 9.08
$b_{\chi_{c0} K_S^0}[^\circ]$	120.53 ± 9.57
$b_{(K^+K^-)_{NR}}[^\circ]$	110.51 ± 5.18
$b_{(K_S^0 K^+)_{NR}}[^\circ]$	-123.37 ± 5.52
$b_{(K_S^0 K^-)_{NR}}[^\circ]$	12.25 ± 5.94
$c_{f_0 K_S^0}$	0.177 ± 0.146
$c_{\phi K_S^0}$	-0.034 ± 0.092
$c_{\chi_{c0} K_S^0}$	0
c_{others}	0.078 ±0.046
$d_{f_0 K_S^0}[^\circ]$	29.49 ± 13.71
$d_{\phi K_S^0}[^\circ]$	28.89 ± 10.12
$d_{\chi_{c0} K_S^0}[^\circ]$	21.9
$d_{\text{others}}[^\circ]$	30.67 ± 9.04

Table 23: Summary of the CP fitting result for the best minimum obtained using 361.6fb^{-1} data, where the errors is statistical only.

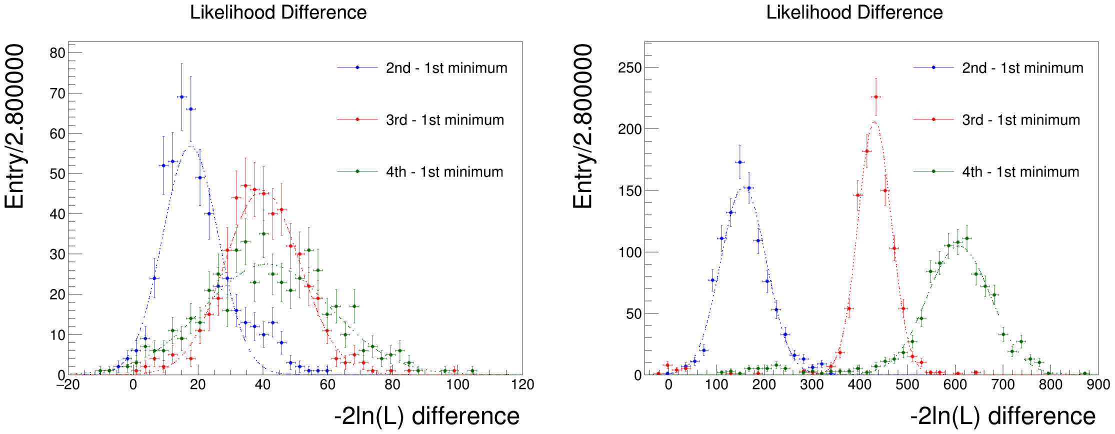
local minimum as the result of this analysis.Fig. 63b also shows the expected likelihood difference with 10 times larger statistics, where we expect almost no overlap of likelihood between 1st and 2nd minimum.

9.2.2 Fit Fraction

In addition to the discussion of $-2\ln\mathcal{L}$ difference in Sec. 9.2.1, we defined fit fraction to compare the fit result with other analysis and branching ratio. Here the fit fraction of each decay channel FF_i is defined as

$$FF_i = \frac{\int_{sDP}(|a_i' F_i|^2 + |\bar{a_i'}\bar{F}_i|^2)dm'd\theta'}{\int_{sDP}(|A|^2 + |\bar{A}|^2)dm'd\theta'}, \tag{119}$$

where the A and F_i is defined in Eq. 62, and i denotes each decay channel. The sum of those fit fraction $\sum_i FF_i$ is not 1 due to the interference between decay channels. We



(a)

(b)

Figure 63: The likelihood difference between local minimums, evaluated using ToyMC with the current statistics (left) and 10 times larger statistics than the current statistics (right). The blue, red or green point represents the distribution of difference in $-2\ln\mathcal{L}$ between the 1st minimum and 2nd, 3rd or 4th minimum respectively. The dotted line represents the Gaussian fitted to each distribution.

$-2\ln\mathcal{L}$ difference	μ	σ
$2\ln\mathcal{L}_1 - 2\ln\mathcal{L}_2$	18.3 ± 0.4	8.9 ± 0.3
$2\ln\mathcal{L}_1 - 2\ln\mathcal{L}_3$	40.0 ± 0.4	11.7± 0.4
$2\ln\mathcal{L}_1 - 2\ln\mathcal{L}_4$	42.0 ± 0.4	19.7± 0.5

Table 24: The summary of the mean μ and width σ of the Gaussian fitted to the likelihood difference between local minimums. The $\mathcal{L}_i$ denotes log-likelihood for i-th local minimum.

additionally define the interference term of fit fraction FF_{ij} as

$$FF_{ij} = 2\,\mathrm{Re}\,\frac{\int_{sDP}(a'_i a'^*_j F_i F^*_j + \bar{a'_i}\bar{a'^*_j}\bar{F_i}\bar{F^*_j})dm'd\theta'}{\int_{sDP}(|A|^2 + |\bar{A}|^2)dm'd\theta'}. \tag{120}$$

Then FF_i and FF_{ij} sum to 1 by definition. The fit fractions calculated from the best fitting result is shown in Table. 25. The sum of the diagonal element of Table. 25 is calculated to be 213 ± 20 %, and the fit fraction of other solutions are also summarized in Appendix C. For the fit fractions $FF_{f_0 K^0_S}$ and $FF_{\phi K^0_S}$, where we expect larger fit fraction for $FF_{f_0 K^0_S}$ against $FF_{\phi K^0_S}$ due to the large branching fraction of $B^0 \to f_0(\to K^+K^-)K^0_S$ against $B^0 \to \phi(\to K^+K^-)K^0_S$, 1st, 3rd and 4th minimum is consistent with the branching fraction. For the fit fraction of $B^0 \to \chi_{c_0} K^0_S$, we expect small fluctuation in the amplitude $a_{\chi_{c0} K^0_S}$ due to the fixed CP violating parameter in the CP fitting and its m' peak position apart from other resonance decays. The Belle [31] and BaBar [29] analysis measured $FF_{\chi_{c0} K^0_S}$ to be 3.4 and 3.8 respectively, and thus 1st and 2nd minimus is consistent with previous results. Thus the best local minimum, which have the smallest $-2\ln\mathcal{L}$ among local minimums, is also most consistent with external measurements.

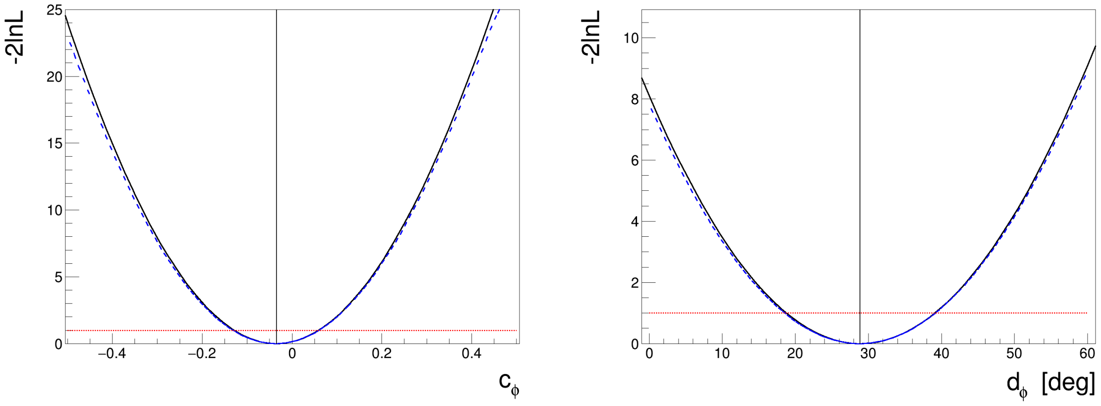
9.2.3 Interpretation of the best fitting result

The following discussion focus on the best local minimum. To validate the fitting result and error estimation, the log-likelihood scan is performed for the best CP fitting result as shown in Fig. 64, where the one fitting parameter is fixed to a value and other parameters are re-fitted. The result of the log-likelihood scan is consistent with the errors obtained by the CP fitting process. The comparison of the distribution of $m(K^-K^0_S)$, $m(K^+K^0_S)$, and m' between the data and fitted PDF is shown in Fig. 65, and observed no significant deviation of PDF from data, while there exist slight underestimate in the event yield for $B^0 \to f_0 K^0_S$ channel.

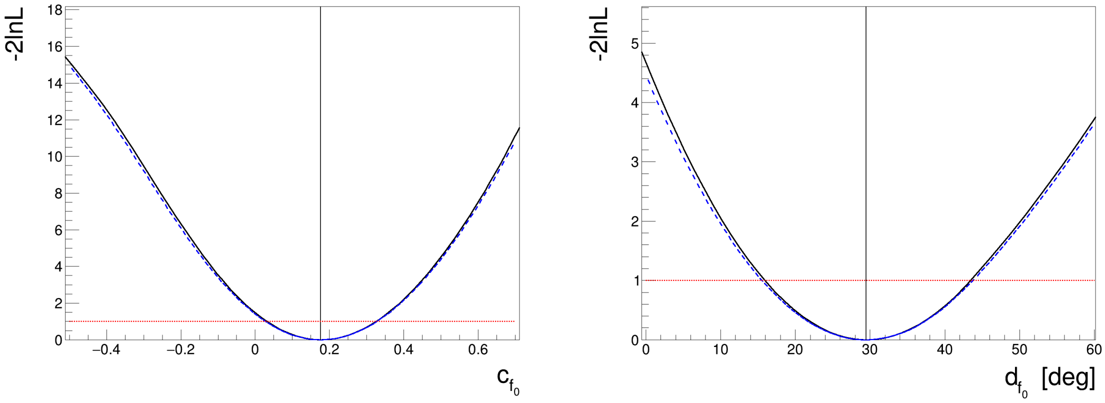
The CP fitting result is converted to the quasi-two-body approach CP asymmeties, S_{CP} and A_{CP}. The result is shown in Table. 26, where the statistic and systematic error is calculated from the errors for dalitz CP parameters c_i and d_i for each decay channels.

Fraction [%]	$f^0 K^0_S$	ϕK^0_S	$f_X K^0_S$	$\chi_{c_0} K^0_S$	$(K^+K^-)_{NR} K^0_S$	$(K^+K^0_S)_{NR} K^-$	$(K^-K^0_S)_{NR} K^+$
$f_0 K^0_S$	15.26 ±5.87	0.00 ±0.00	-0.01 ±0.84	0.06 ±0.05	-48.89 ±19.75	15.33 ±1.57	-4.09 ±2.17
ϕK^0_S		14.89 ±0.81	-0.00 ±0.00	-0.00 ±0.00	-0.00 ±0.00	-0.00 ±0.01	0.00 ±0.01
$f_X K^0_S$			2.67 ±0.79	0.03 ±0.01	4.02 ±0.25	2.85 ±0.68	-2.62 ±0.52
$\chi_{c_0} K^0_S$				3.14 ±0.54	-0.15 ±0.09	1.27 ±0.18	-1.18 ±0.14
$(K^+K^-)_{NR} K^0_S$					112.33 ±18.54	-44.02 ±6.71	-9.27 ±0.69
$(K^+K^0_S)_{NR} K^-$						37.12 ±4.64	-26.47 ±3.68
$(K^-K^0_S)_{NR} K^+$							27.72 ±3.88

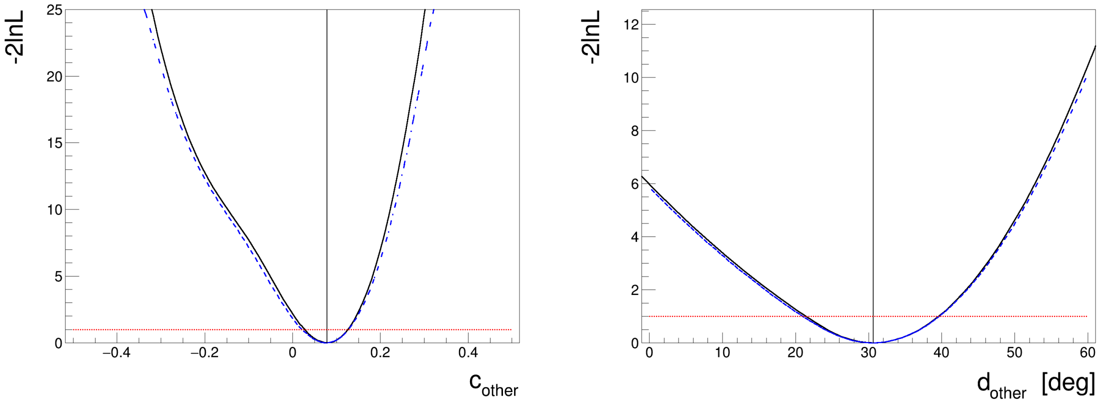
Table 25: Values of the fit fraction FF_{ij} calculated from CP fitting result in the unit of %.



(a) $c_{\phi K^0_S}$ (b) $d_{\phi K^0_S}$

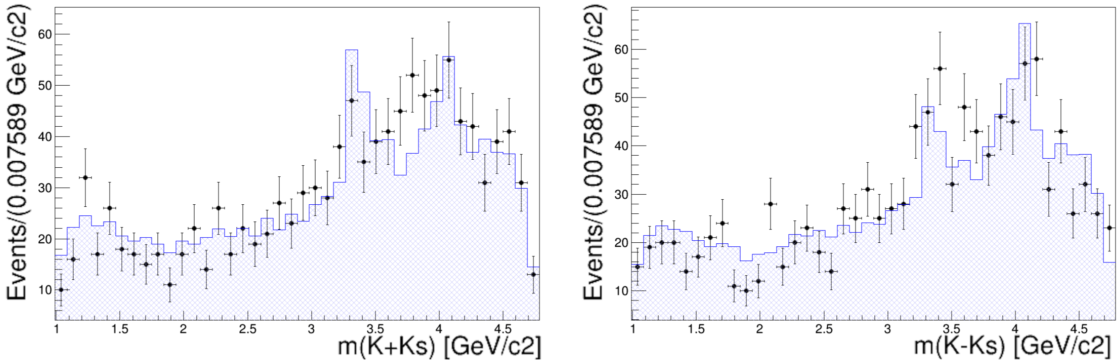


(c) $c_{f_0 K^0_S}$ (d) $d_{f_0 K^0_S}$

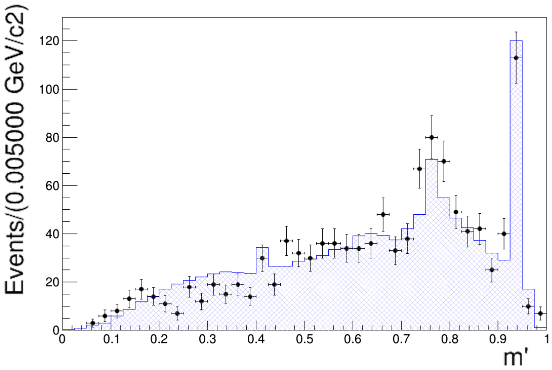


(e) c_{others} (f) d_{others}

Figure 64: The result of log-likelihood scan for the best local minimum found in this analysis. The black line represents log-likelihood with only statistical error, and the blue dashed line represents log-likelihood with both statistical and systematic errors. The vertical line shows the value of fitting parameters obtained by CP fitting, and the red vertical line shows 1σ threshold. The log-likelihood is re-calculated by fixing one parameter and re-fitting other parameters.



(a) $m(K^+K_S^0)$ (b) $m(K^-K_S^0)$



(c) m'

Figure 65: Distribution of $m(K^-K_S^0)$, $m(K^+K_S^0)$, and m' for data and PDF. The black points represent the data distribution with its statistic error, and the blue histogram represents the fitted PDF projected to each variables.

	A_{CP}	S_{CP}
ϕK_S^0	$0.07 \pm 0.18 \pm 0.04$	$0.84^{+0.13+0.03}_{-0.24-0.03}$
$f_0 K_S^0$	$-0.34 \pm 0.26 \pm 0.06$	$-0.81^{+0.31+0.05}_{-0.13-0.04}$
others	$-0.16 \pm 0.09 \pm 0.04$	$-0.87^{+0.19+0.03}_{-0.10-0.03}$

Table 26: Summary of the the quasi-two-body CP violating parameters delivered from Dalitz CP fitting result, where the errors is statistical and systematic respectively.

9.2.4 Raw asymmetry distribution

The Δt distribution is compared between data and PDF to visualize the CP asymmetries. Events are divided into two regions: ϕ mass region defined as $1.01 < m(K^+K^-) < 1.03\,\mathrm{GeV}$ and not-ϕ mass region defined as the rest of region excluding ϕ mass region. We also compared distribution for only high qr-bin events defined as qr-bin $= 4, 5$ or 6, which mean $|qr| > 0.6$. Here a raw asymmetry for each Δt interval is defined as $(N_+ - N_-)/(N_+ + N_-)$, where $N_\pm$ is the number of event candidates with $q = \pm 1$. The comparison of Δt distribution between all the event in the signal region and the full CP fitting PDF is shown in Fig. 66 for ϕ mass region and not-ϕ mass region respectively, and the raw asymmetries are shown in Fig. 67. The distribution of signal event and signal component of PDF is also compared by adopting sPlot technique [33] as shown in Fig. 68 and Fig. 69. As shown in Figures, CP-odd events are dominant in ϕ mass region, and CP-even events are dominant in not-ϕ mass region.

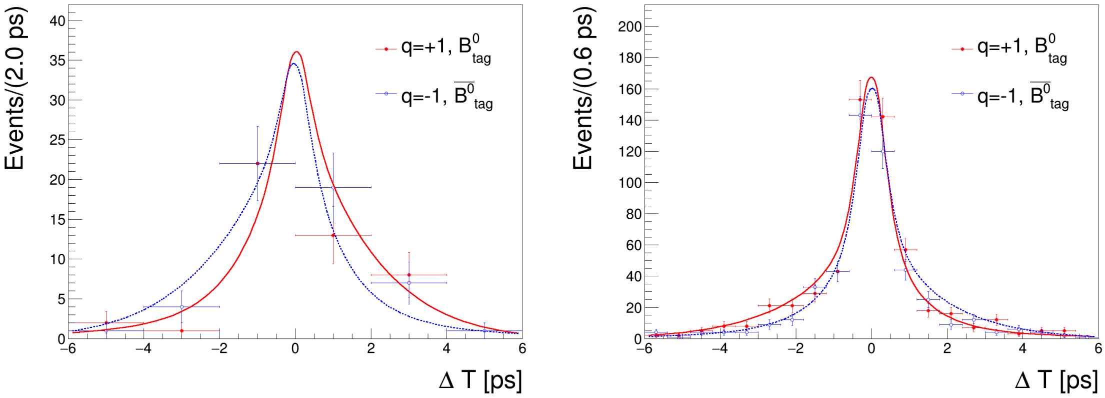
9.3 Summary of CP asymmetries and significance

We obtained CP asymmetries to be

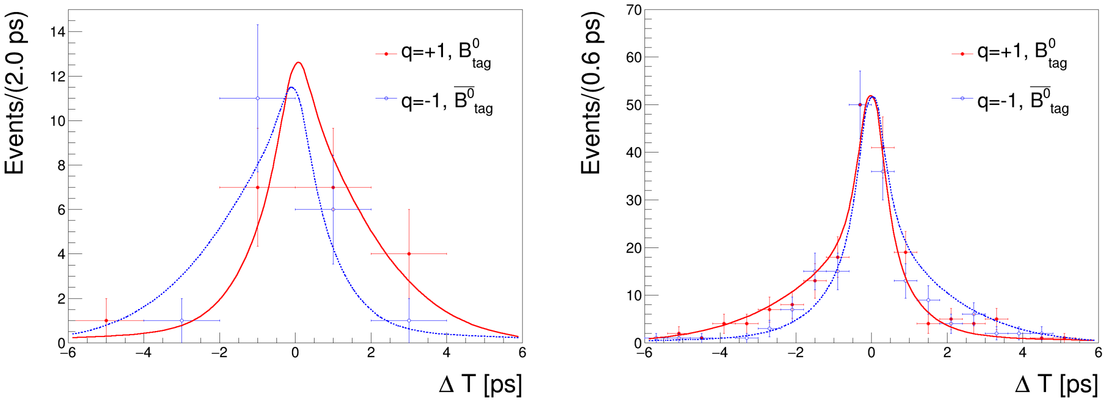
$$\begin{aligned}
\phi_1(\phi K_S^0) &= 28.9 \pm 10.1 \pm 1.7^\circ \\
A_{CP}(\phi K_S^0) &= 0.07 \pm 0.18 \pm 0.04 \\
\phi_1(f_0 K_S^0) &= 29.5 \pm 13.7 \pm 2.8^\circ \\
A_{CP}(f_0 K_S^0) &= -0.34 \pm 0.26 \pm 0.06 \\
\phi_1(\text{others}) &= 30.7 \pm 9.0 \pm 1.8^\circ \\
A_{CP}(\text{others}) &= -0.16 \pm 0.09 \pm 0.04
\end{aligned} \tag{121}$$

where the errors is statistical and systematic respectively.

The log-likelihood scan shown in Fig. 64 is used in order to estimate the significance of CP violation measurement. The likelihood is convoluted with Gaussian with its width set to the systematic uncertainty in order to take both statistical and systematic uncertainty into consideration, then the significance is estimated from log-likelihood distribution. For ϕ_1 measurement with $B^0 \to \phi K_S^0$ channel, we observed 2.8σ deviation from no CP violation hypothesis $\phi_1 = 0$, and the result is consistent with SM prediction within 0.6σ. For the direct CP asymmetry A_{CP} with $B^0 \to \phi K_S^0$ channel, we observed consistent result with SM prediction $A_{CP} = 0$ within 0.4σ.

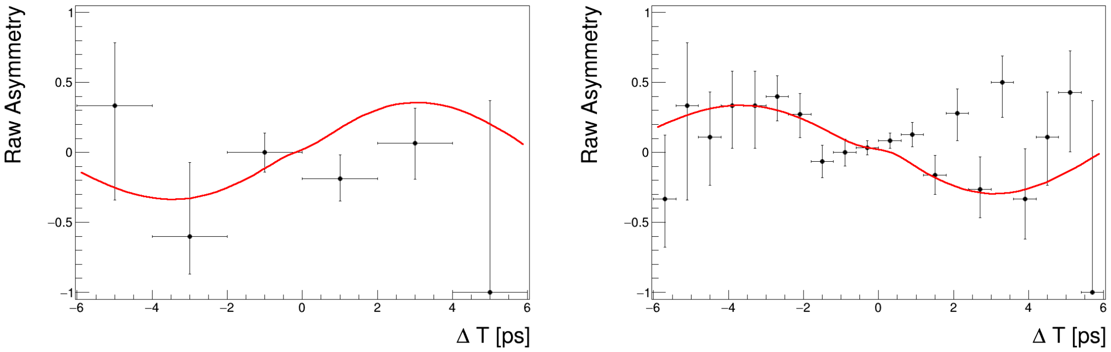


(a) ϕ mass region, including all bin (b) not-ϕ mass region, including all bin



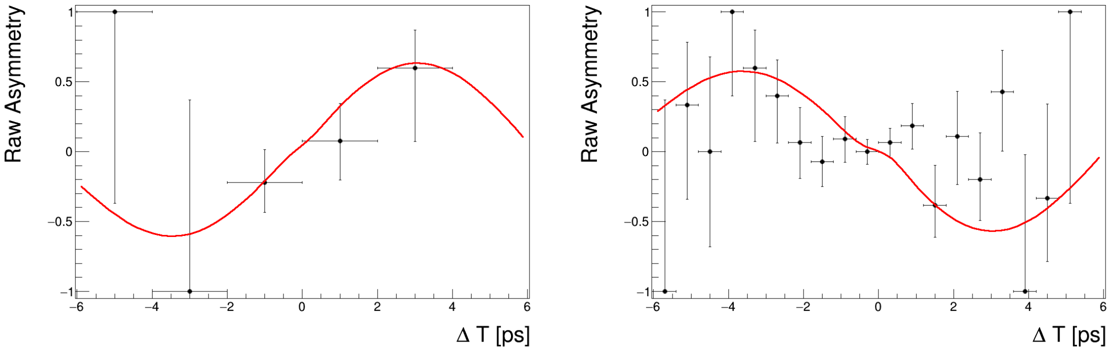
(c) ϕ mass region, including only good bin (d) not-ϕ mass region, including only good bin

Figure 66: Comparison of Δt distribution between data and PDF for all event. The ϕ mass region is defined as $1.01 < m(K^+K^-) < 1.03\,\mathrm{GeV}/c^2$, and the good bin is defined as qr-bin $= 4, 5$ or 6. The red (blue) points and lines represent $q_{\mathrm{tag}} = +1(-1)$ event candidates and PDF respectively.



(a) ϕ mass region, including all bin

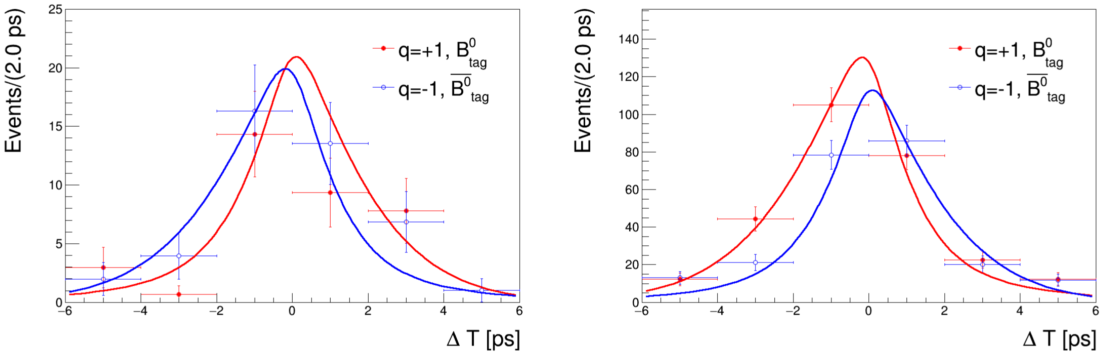
(b) not-ϕ mass region, including all bin



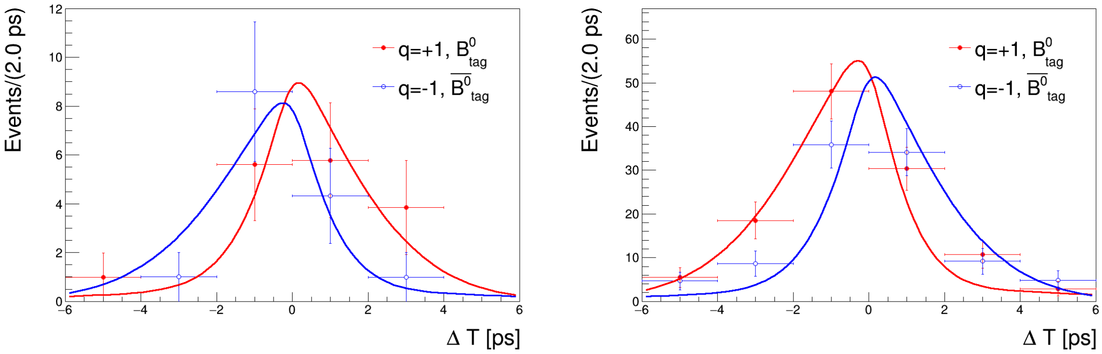
(c) ϕ mass region, including only good bin

(d) not-ϕ mass region, including only good bin

Figure 67: Comparison of raw asymmetries in each Δt region between data and PDF for all event. The ϕ mass region is defined as $1.01 < m(K^+K^-) < 1.03\,\mathrm{GeV}/c^2$, and the good bin is defined as qr-bin $= 4, 5$ or 6. The red line and black points represent raw asymmetry calculated from PDF and raw asymmetry in the event candidates respectively.

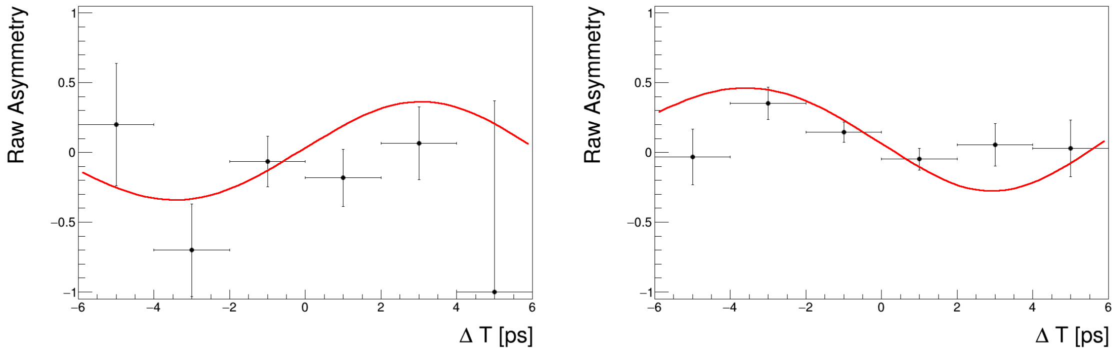


(a) ϕ mass region, including all bin (b) not-ϕ mass region, including all bin



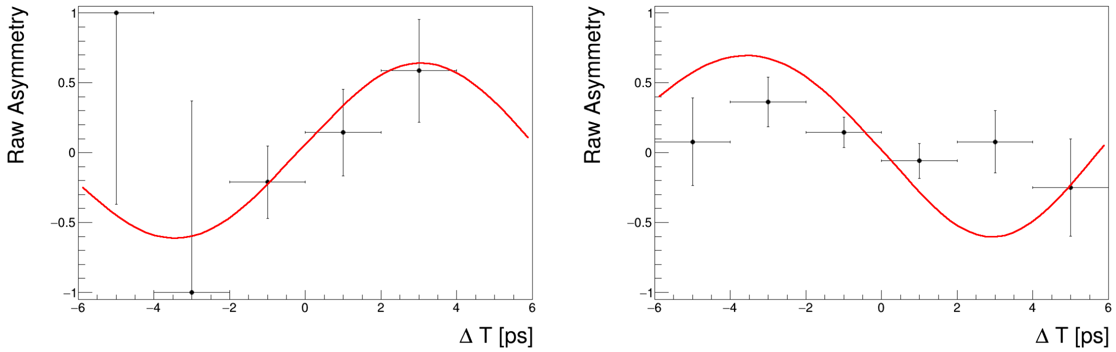
(c) ϕ mass region, including only good bin (d) not-ϕ mass region, including only good bin

Figure 68: Comparison of Δt distribution between sWeighted data and signal component of PDF. The ϕ mass region is defined as $1.01 < m(K^+K^-) < 1.03\,\mathrm{GeV}/c^2$, and the good bin is defined as qr-bin $= 4, 5$ or 6. The red (blue) points and lines represent $q_{\mathrm{tag}} = +1(-1)$ sWeighted event candidates and signal component of PDF respectively.



(a) ϕ mass region, including all bin

(b) not-ϕ mass region, including all bin



(c) ϕ mass region, including only good bin

(d) not-ϕ mass region, including only good bin

Figure 69: Comparison of Δt raw asymmetries between sWeighted data and signal component of PDF. The ϕ mass region is defined as $1.01 < m(K^+K^-) < 1.03\,\mathrm{GeV}/c^2$, and the good bin is defined as qr-bin $= 4, 5$ or 6. The red line and black points represent raw asymmetry calculated from signal component of PDF and raw asymmetry in the sWeighted event candidates respectively.

9.4 Comparison with previous experiments

Belle and BaBar experiment previously measured CP asymmetries in $B^0 \to K^+K^-K_S^0$ three body decay with time-dependent Dalitz plot analysis [31] [29].

Belle experiment measured CP asymmetries with 657×10^6 $B\bar{B}$ pairs. The most preferable solution from the Belle experiment is

$$\begin{aligned} \phi_1(\phi K_S^0) &= 32.0^{+8.8}_{-8.3} \pm 2.0^\circ \\ A_{CP}(\phi K_S^0) &= 0.04 \pm 0.20 \pm 0.10 \\ \phi_1(f_0 K_S^0) &= 30.5^{+8.6}_{-8.3} \pm 3.8^\circ \\ A_{CP}(f_0 K_S^0) &= -0.32^{+0.27}_{-0.29} \pm 0.13 \\ \phi_1(\text{others}) &= 24.4^{+6.2}_{-5.8} \pm 3.3^\circ \\ A_{CP}(\text{others}) &= -0.14 \pm 0.11 \pm 0.09, \end{aligned} \tag{122}$$

where the error is statistical and systematic respectively. When comparing the result from Belle experiment and this analysis, the measured CP asymmetries in this analysis is consistent with the previous result within its uncertainties due to the large statistical errors.

In terms of the statistical uncertainty, the statistic error in this analysis is consistent with that of Belle experiment analysis when take the effect of small statistics in this analysis consideration. On the other hand, the systematic uncertainty is estimated to improve from the previous analysis. The major contribution for this improvement is expected to be the good Δt resolution in Belle II experiment. The vertex reconstruction and Δt measurement was one of the major systematic uncertainties in the Belle analysis, but in this analysis the systematic uncertainties from those components are estimated to be small compared to that of Belle experiment, leading to better systematic uncertainties.

9.5 Future prospect

9.5.1 Improvement of uncertainties

The uncertainties for CP asymmetry measurement are currently dominated by statistic uncertainty in this analysis. With the increase in the integrated luminosity, part of systematic uncertainties are expected to become smaller along with the statistical uncertainty. Notable improvement in the systematic uncertainties expected in the future is as follows:

- Fit bias : The uncertainty due to fit bias include the possible effect of data statistics itself being small, and thus expected to be reduced with the increased statistics.
- $B\bar{B}$ background : The uncertainty from $B\bar{B}$ background modeling arise from the low statistics of background events from $B\bar{B}$ events. It will also be possible to calibrate $B\bar{B}$ background from data side-band with $B\bar{B}$ background enhanced region such as $(M_{bc} > 5.26\text{GeV}/c^2) \wedge (\Delta E < 0\text{GeV})$ as shown in Fig. 40.
- $q\bar{q}$ background : The uncertainty from $q\bar{q}$ background modeling arise from possible difference of m' and θ' distribution for $q\bar{q}$ background events between signal region

and side-band region. In this analysis we adopt narrow Dalitz side-band conservatively as stated in Sec. 7.3.2, but with more statistics we can investigate m' and θ' distribution for $q\bar{q}$ events in more detail to enlarge Dalitz side-band and might be possible to model $q\bar{q}$ background precisely.

- Fixed Parameters : The uncertainty due to fixed parameters include the effect of uncertainties from flavor tagging and resolution function. The uncertainties for those parameters are currently dominated by the statistical uncertainty due to the small statistics used for calibration, and thus expected to improve in the future with large statistics, although the systematic errors are irreducible.
- Non-resonant modeling : The current non-resonant component modeling and its shape parameter follows the empirical modeling from the previous analysis[31]. With increased statistics, it should become possible to include the line-shape parameters of non-resonant modeling itself into the fitting procedure, which will greatly improve the uncertainty deriving from the discrepancy between non-resonant modeling and data distribution.
- Possible resonance : In the current analysis, the minor resonances are excluded from CP fitting PDF to avoid possible affects from the number of event for minor resonances being too small. With increased statistics, it will become possible to correctly handle those resonances and include in CP fitting PDF, leading to reduction in the systematic uncertainty.

Along with those improvement, systematic uncertainties for fit bias, $B\bar{B}$ background, $q\bar{q}$ background, and part of fixed parameters are expected to be reduced proportional to $1/\sqrt{\mathcal{L}}$, where $\mathcal{L}$ is the integrated luminosity. Under this assumption, the relation between the integrated luminosity and uncertainties for the measurement of ϕ_1 through $B^0 \to \phi K_S^0$ mode are shown in Fig. 70 along with the uncertainty in Belle analysis and this analysis. The total uncertainty for ϕ_1 measurement is expected to be improved from Belle experiment result when we collect the same statistics as Belle experiment. In the Belle experiment, the most dominant systematic uncertainty arose from the vertex reconstruction, while we expect smaller systematic uncertainty due to the misalignment of detector and Δt measurement owing to the larger and more precise vertex detector in Belle II experiment. We also expect higher K_S^0 selection efficiency in this analysis. For the K_S^0 selection method in Belle experiment, K_S^0 selection efficiency was 86.9% with K_S^0 purity 94.0% [32], and thus we expect higher K_S^0 selection efficiency and better background rejection in this analysis compared to the previous analysis.

As shown in Fig. 70, the statistical uncertainty will be comparable against systematic uncertainty at about $20\,\mathrm{ab}^{-1}$ if we adopt the same analysis method due to the large systematic uncertainty originating from Dalitz modeling. Above this integrated luminosity, it will be necessary to improve Dalitz model with more precise non-resonant handling and additional minor resonances to achieve further precise measurement of ϕ_1.

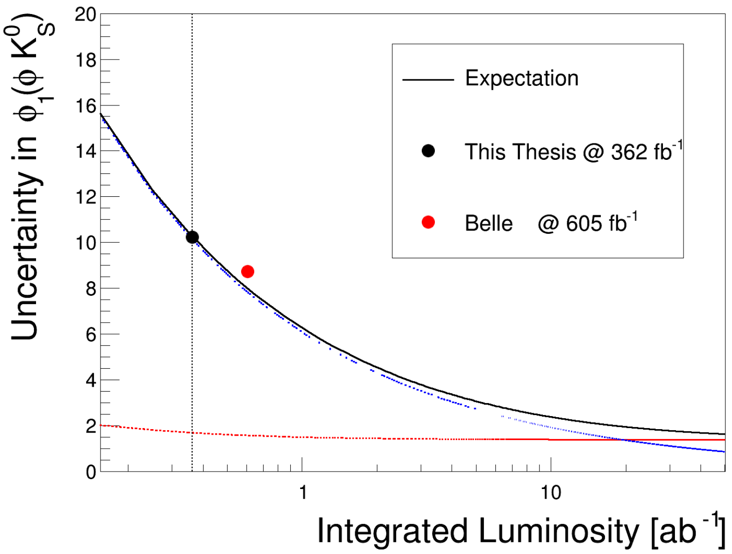


Figure 70: The future prospect of statistical and systematic uncertainties for the measurement of ϕ_1 through $B^0 \to \phi K^0_S$. The black, blue and red line represent total, statistical and systematic uncertainty respectively. The black or red point represents uncertainty in this analysis or Belle analysis respectively.

9.5.2 Improvement of analysis method

In this analysis we adopt analysis strategy with low signal yield and high purity close to the Belle analysis, but the other approach as BaBar adopted is possible for future analysis.

They included $f'_2(1525)$ and $f_0(1710)$ into CP fitting, and also they are adopting different non-resonant component modeling as mentioned in Sec. 8.3.1 in exchange for the loose cut criteria for event selection and much larger background events. They found 1579 signal events against 6881 continuum events and 116 $B\bar{B}$ background. The purity of signal event in BaBar is 0.18 and this is much worse than Belle and this analysis, while it is possible to make use of the large statistics to construct proper model for non-resonant component and to include minor resonances which cannot be used for this analysis due to too small signal yield. The possible disadvantage of this approach is the large affect of background events onto non-resonant component modeling, but there is possibility to improve systematic uncertainty by construct detailed background and non-resonant modeling with large statistics in the future.

The another possible improvement is to include $K^0_S \to \pi^0\pi^0$ decay mode in addition to the current $K^0_S \to \pi^+\pi^-$ decay channel. BaBar found 160 signal events through $K^0_S \to \pi^0\pi^0$ decay in compensate for 2751 background event. In order to improve sensitivity by including $K^0_S \to \pi^0\pi^0$ we need more precise modeling of background event distribution and development of K^0_S reconstruction method through $K^0_S \to \pi^0\pi^0$, nevertheless this have possibility to reduce statistic uncertainty after collecting full data set in the future.

The flavor tagging is also expected to improve in the near future. Apart from the category-based flavor tagger, a flavor tagging method with deep-learning neural network is currently developed [4]. In this flavor tagging method, the B flavor is directly determined using Graph-Neural-Network (GNN) technique without pre-categorisation of B meson decay products. The effective tagging efficiency for GNN flavor tagger is expected to be improved from current category-based flavor tagger [34], and thus we can expect improvement in the statistical uncertainties for CP asymmetry measurement in the near future.

In this analysis, we adopted isobar model to describe interference between decay channels. Theoretically, the $K^+K^-K_S^0$ final state should be treated with the unitarity kept, while isobar model is an approximate model which does not satisfy the unitarity of the three-body final state[7]. In Ref.[24], they adopted unitary coupled-channels model to include the re-scattering effect of multiple resonances and showed a discrepancy between unitary coupled-channels model and isobar model in three body decays such as $D^0 \to \pi^+\pi^-\pi^0$. The re-scattering effect is expected to be small for B meson decay than for D meson decay in general, however, it is not fully known yet how large the re-scattering effect in B meson three-body decay is. This leads to the possibility that the systematic uncertainty in this analysis is underestimated depending on the discrepancy between isobar model and unitary coupled-channels model, and we await detailed theoretical prediction to discuss possible improvement of the treatment of interference in the future analysis.

9.6 Constraint on the new physics parameter

As discussed in Sec. 2.4, the measurement of ϕ_1 through $B^0 \to \phi K_S^0$ decay can give a constraint on the NP parameter. Following Eq. 47, we can plot the relation between θ_{NP}, which is the weak phase contributing from NP, and $S_{CP}^{\phi K_S^0}$ for different $|A_{NP}/A_{SM}|$ ratio as shown in Fig. 71 [15], where we assume δ_{12}=0. The red line represents the central value of measured $S_{CP}^{\phi K_S^0}$, and the gray region represent the excluded region with 95.45% confidence level. Each line in Fig. 71 represents the relation between $S_{CP}^{\phi K_S^0}$ and θ_{NP} with assumption of the amplitude ratio $|A^{NP}/A^{SM}|$ to be 0.2 (dashed), 0.4 (dashed-dotted), 0.6 (dotted) and 0.8 (dashed-double-dotted) respectively. We also estimated confidence regions over $(\theta_{NP}, |A^{NP}/A^{SM}|)$ plane as shown in Fig. 72, where red or black line shows 68.27% or 95.45% confidence level contour respectively. Here we adopted the frequentist approach to estimate the confidence interval on $(\theta_{NP}, |A^{NP}/A^{SM}|)$ plane. As statistics increase in the future, the yellow region with high confidence level is expected to spread over the plane, leading to the exclusion of large $|A^{NP}/A^{SM}|$ value for large $\sin\theta_{NP}$ region.

Fig. 72 shows the model-independent constraint on the effect of NP on $B^0 \to \phi K_S^0$, while it is also possible to give model-dependent constraint by assuming certain NP model. One example is the mass insertion as discussed in Sec. 2.4, and the constraint on the mass insertion parameters is discussed in Appendix D.

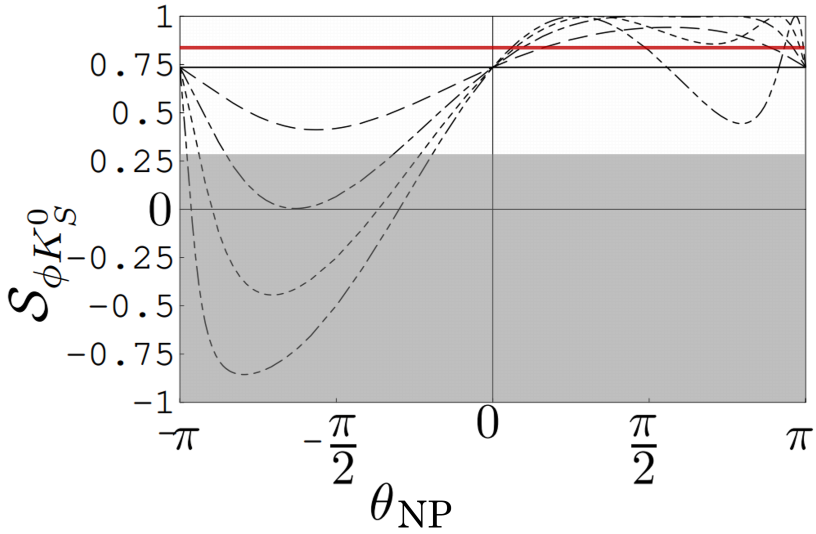
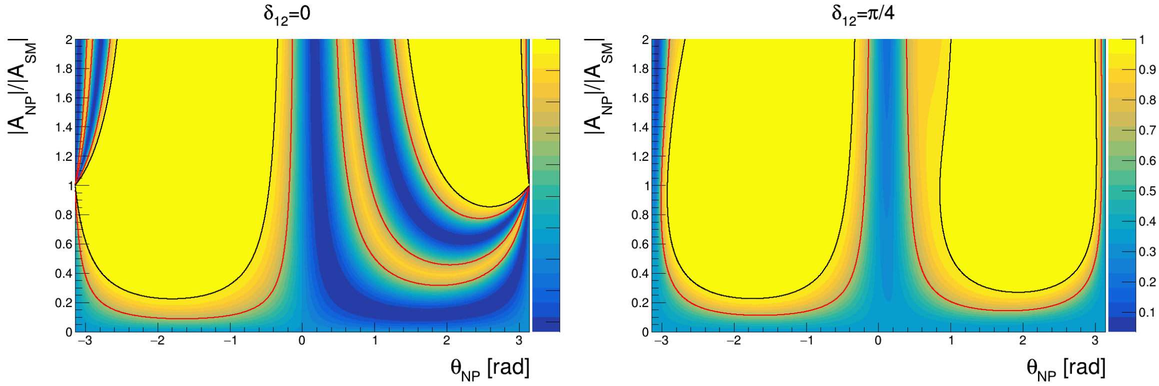
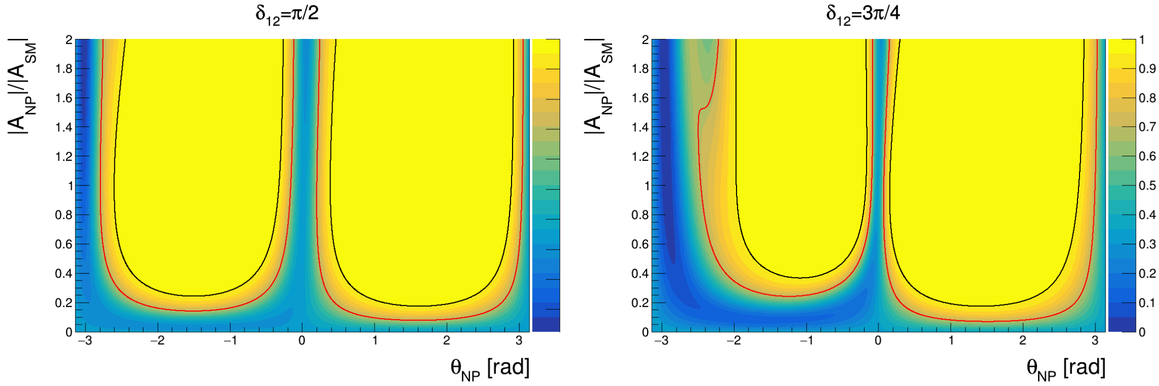


Figure 71: The relation between the NP weak phase θ_{NP} and $S_{CP}^{\phi K_S^0}$ assuming $\delta_{12} = 0$ [15]. Each line represents the value of $S_{CP}^{\phi K_S^0}$ as a function of θ_{NP} for different ratio of amplitude $|A^{NP}/A^{SM}| = 0.2$ (dashed), 0.4 (dashed-dotted), 0.6 (dotted) and 0.8 (dashed-double-dotted) respectively. The red line represents the central value of measured $S_{CP}^{\phi K_S^0}$. The gray region represent the excluded region with 95.45% confidence level.



(a) $\delta_{12} = 0$ (b) $\delta_{12} = \pi/4$



(c) $\delta_{12} = \pi/2$ (d) $\delta_{12} = 3\pi/4$

Figure 72: The confidence interval over the $(\theta_{NP}, |A^{NP}/A^{SM}|)$ plane assuming different δ_{12} value. The red or black line shows 68.27% or 95.45% confidence level contour respectively. The yellow (blue) region represents region with high (low) confidence level.

10 Conclusion

$B^0 \to \phi K_S^0$ decay process is dominated by penguin transition, and is sensitive to the effect from new physics. In order to handle the interference between $B^0 \to \phi K_S^0$ and other channels decaying into the same final state $B^0 \to K^+K^-K_S^0$ corretly, the Dalitz plot technique is adopted. We reported measurement of time-dependent CP asymmetries in the $B^0 \to K^+K^-K_S^0$ three body decay using 387×10^6 $B\bar{B}$ pairs collected at Belle II experiment. We obtained

$$\begin{aligned} \phi_1(\phi K_S^0) &= 28.9 \pm 10.1(\text{stat}) \pm 1.7(\text{syst})^\circ \\ A_{CP}(\phi K_S^0) &= 0.07 \pm 0.18(\text{stat}) \pm 0.04(\text{syst}) \end{aligned} \tag{123}$$

which is consistent with the previous measurements from Belle experiment analysis. The existence of CP violation is confirmed at 2.4σ significance, and the result is consistent with SM within 0.6σ. We also discussed expected uncertainties for CP asymmetries measurement, and showed future prospect of the measurement through $B^0 \to K^+K^-K_S^0$ three body decay and possible improvement.

Acknowledgement

I would like to thank Prof. Yutaka Ushiroda for giving the opportunity to participate in the Belle II experiment and supporting my analysis. I also thank Dr. K. Sumisawa, Dr. K. Hara, Dr. T. Higuchi and Dr. A. Ishikawa for their helpful discussion and comment for my whole analysis.

I thank members Belle II time-dependent CP violation group for cooperation in the CP violation measurement and advices.

I thank all the members of the Belle II collaboration and the SuperKEKB accelerator group for maintaining the Belle II experiment and offering data set for this analysis.

I thank Dr. H. Tanigawa, Mr. Y. Uematsu and Mr. B. Hanwook and the members of our laboratory for continuous discussion and cooperation.

References

[1] R. Aaij, B. Adeva, and LHCb. Dalitz plot analysis of $B_s^0 \to \overline{d}^0 K^- \pi^+$ decays. *Phys. Rev. D*, 90:072003, Oct 2014.

[2] T Abe, I Adachi, K Adamczyk, S Ahn, H Aihara, K Akai, M Aloi, L Andricek, K Aoki, Y Arai, et al. Belle II technical design report. *arXiv preprint arXiv:1011.0352*, 2010.

[3] M. Ablikim, J.Z. Bai, et al. Resonances in and. *Physics Letters B*, 607(3-4):243–253, Feb 2005.

[4] F. Abudinén et al. B-flavor tagging at Belle II. *The European Physical Journal C*, 82(4), Apr 2022.

[5] H. Albrecht, T. Hamacher, et al. Measurement of the polarization in the decay $B \to J/\psi K^*$. *Physics Letters B*, 340(3):217–220, 1994.

[6] W. Altmannshofer et al. The Belle II physics book. *PTEP*, 2019(12):123C01, 2019.

[7] Ignacio Bediaga and Carla Göbel. Direct CP violation in beauty and charm hadron decays. *Progress in Particle and Nuclear Physics*, 114:103808, 2020.

[8] M. Beneke. Corrections to $\sin(2\beta)$ from CP asymmetries in $B^0 \to (\pi^0, \rho^0, \eta, \eta', \omega, \phi)$ K_S decays. *Physics Letters B*, 620(3):143–150, 2005.

[9] A. J. Bevan et al. The physics of the B factories. *Eur. Phys. J.*, C74:3026, 2014.

[10] John Markus Blatt and Victor Frederick Weisskopf. *Theoretical nuclear physics.* Springer, New York, 1952.

[11] Ashton B. Carter and A. I. Sanda. CP nonconservation in cascade decays of B mesons. *Phys. Rev. Lett.*, 45:952–954, Sep 1980.

[12] J Charles, A Hocker, Heiko Lacker, Sandrine Laplace, FR Le Diberder, Julie Malcles, Jose Ocariz, Muriel Pivk, and Lydia Roos. CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories. *The European Physical Journal C-Particles and Fields*, 41(1):1–131, 2005.

[13] 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.*, 13:138–140, 1964.

[14] Thomas Czank, Thibaud Humair, Yosuke Yusa, and Radek Zlebcik. Measurement of the CP Violation Parameter $\sin 2\phi_1$ in $B^0 \to J/\psi K^0$ Decay using 189.26 fb^{-1} data set. May 2022.

[15] A. G. Akeroyd et al. Physics at super B factory, 2004.

[16] S.M. Flatté. Coupled-channel analysis of the $\pi\eta$ and $K\bar{K}$ systems near $K\bar{K}$ threshold. *Physics Letters B*, 63(2):224–227, 1976.

[17] Particle Data Group et al. Review of Particle Physics. *Progress of Theoretical and Experimental Physics*, 2022(8), 08 2022. 083C01.

[18] L.J. Hall, V.A. Kostelecky, and S. Raby. New flavor violations in supergravity models. *Nuclear Physics B*, 267(2):415–432, 1986.

[19] Takeo Higuchi. Inference of the true CP-violating parameters from observed CP-violating parameters affected by the tag-side interference effect. Dec 2022.

[20] Thibaud Humair, Justin Skorupa, Jakub Kandra, Radek Zlebcik, Caspar Schmitt, and Vladimir Chekelian. Measurement of the B^0 lifetime and mixing frequency using hadronic decays with the Moriond 2022 dataset. Oct 2021.

[21] R Itoh, M Nakao, and S Yamada. Belle II data acquisition system. *HEP Vol.33,No.3 (2014), p. 196*, 2014.

[22] Yoshihito Iwasaki, ByungGu Cheon, Eunil Won, and Gary Varner. Level 1 trigger system for the Belle II experiment. In *2010 17th IEEE-NPSS Real Time Conference*, pages 1–9. IEEE, 2010.

[23] Zhou Jianshe, Zhou Xingyu, and Shen. Chengping. Offline measurements of integrated luminosities of Moriond 2023 datasets. *BELLE2-NOTE-PH-2022-056*, 2022.

[24] H. Kamano, S. X. Nakamura, T.-S. H. Lee, and T. Sato. Unitary coupled-channels model for three-mesons decays of heavy mesons. *Phys. Rev. D*, 84:114019, Dec 2011.

[25] Thomas Keck. Fastbdt: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification, 2016.

[26] S. Khalil and E. Kou. On supersymmetric contributions to the CP asymmetry of the $B \to \phi K_S$ process. *Phys. Rev. D*, 67:055009, Mar 2003.

[27] Makoto Kobayashi and Toshihide Maskawa. CP violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.*, 49:652, 1973.

[28] J.-F. Krohn et al. Global decay chain vertex fitting at Belle II. *Nucl. Instrum. Meth.*, A976:164269, 2020.

[29] J. P. Lees, V. Poireau, and others. Study of CP violation in dalitz-plot analyses of $B^0 \to K^+K^-K^0_S$, $B^+ \to K^+K^-K^+$, and $B^+ \to K^0_S K^0_S K^+$. *Phys. Rev. D*, 85:112010, Jun 2012.

[30] Owen Long, Max Baak, Robert N. Cahn, and David Kirkby. Impact of tag-side interference on time-dependent CP asymmetry measurements using coherent B^0 $\bar{B}^0$ pairs. *Physical Review D*, 68(3), Aug 2003.

[31] Y Nakahama et al. Measurement of CP violating asymmetries in $B^0 \to K^+K^-K^0_S$ decays with a time-dependent dalitz approach. *Phys. Rev. D*, 82:073011, Oct 2010.

[32] Hiroshi Nakano. Search for new physics by a time-dependent CP violation analysis of the decay B→Ks eta gamma using the Belle detector.

[33] Muriel Pivk and Francois R. Le Diberder. sPlot: A statistical tool to unfold data distributions. *Nucl. Instrum. Meth.*, A555:356–369, 2005.

[34] Yo Sato, Thibaud Humair, and Petros Stavroulakis. Graph-neural-network based flavor tagger. Mar 2023.

[35] T Unno, Y Iwasaki, and H Nakazawa. Belle II trigger system. *HEP news Vol.34,No.2 (2015), p. 105*, 2015.

[36] Lincoln Wolfenstein. Parametrization of the Kobayashi-Maskawa matrix. *Phys. Rev. Lett.*, 51:1945–1947, Nov 1983.

[37] R. L. Workman et al. Review of Particle Physics. *PTEP*, 2022:083C01, 2022.

[38] W-M Yao et al. Review of particle physics. *Journal of Physics G: Nuclear and Particle Physics*, 33(1):1, Jul 2006.

[39] Laura Zani. Track reconstruction efficiency measurement using $e^+e^- \to \tau^+\tau^-$ events at Belle II. *PoS*, ICHEP2020:797, 2021.

[40] Q. D. Zhou, S. Yamada, P. Robbe, D. Charlet, R. Itoh, M. Nakao, S. Y. Suzuki, T. Kunigo, E. Jules, E. Plaige, M. Taurigna, H. Purwar, O. Hartbrich, M. Bessner, K. Nishimura, G. Varner, Y.-T. Lai, T. Higuchi, R. Sugiura, D. Biswas, and P. Kapusta. PCI-express based high-speed readout for the Belle II DAQ upgrade. *IEEE Transactions on Nuclear Science*, 68(8):1818–1825, 2021.

A Observable correlation

The correlation between observable $(M_{bc}, \Delta E, qr, m', \theta', \Delta t, \sigma\Delta t)$ is analyzed to consider the effect of observable correlations onto CP asymmetries measurement. The correlation between M_{bc} and ΔE in the signal MC samples can be seen in Fig. 73. The distribution of M_{bc} and ΔE for signal MC $q\bar{q}$ background events are also shown in Fig. 74, Fig. 75. The correlation between square Dalitz parameter m', θ' and Δt is shown in Fig. 76 and Fig. 77.

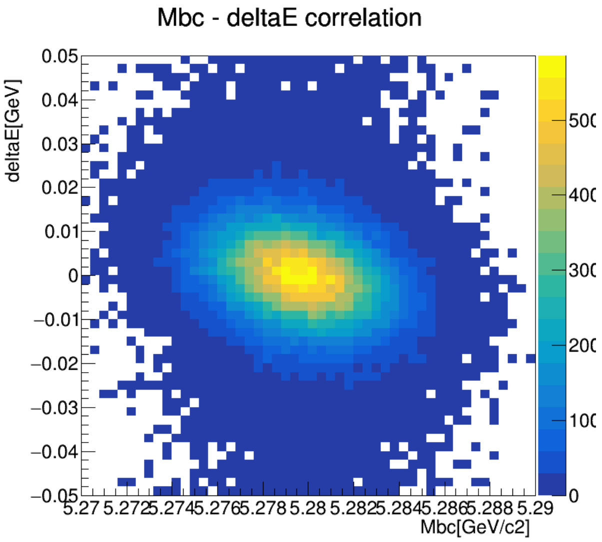
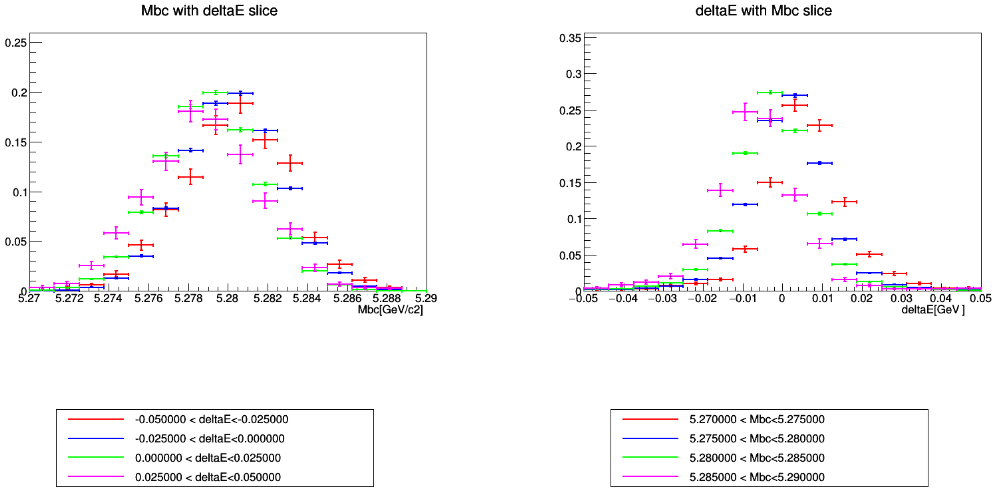


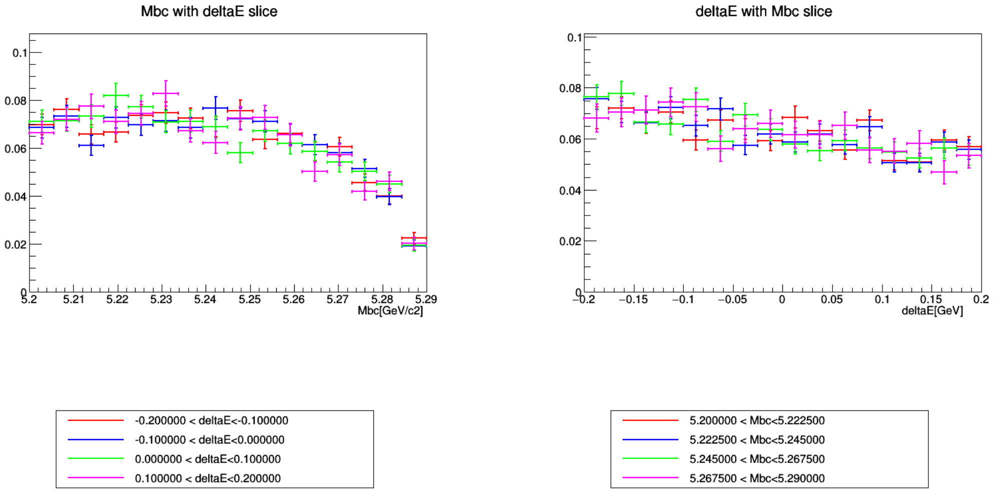
Figure 73: The 2-dimension scattering plot of $M_{bc} - \Delta E$ distribution in MC signal events.



(a) M_{bc} over ΔE slice

(b) ΔE over M_{bc} slice

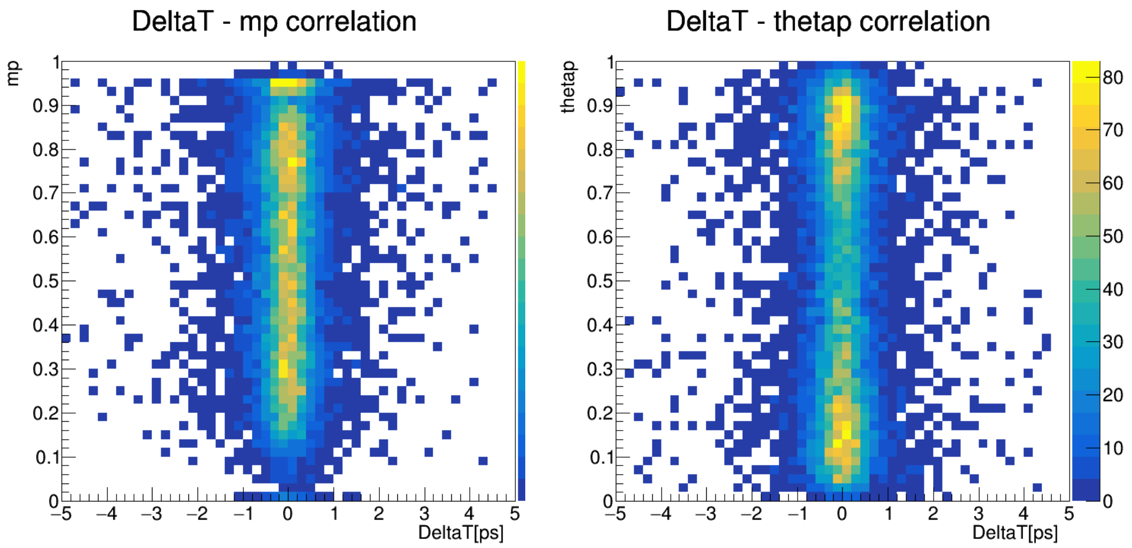
Figure 74: The distribution of M_{bc} with ΔE slices(left) and the distribution of ΔE with M_{bc} slices(right) for signal MC events.



(a) M_{bc} over ΔE slice

(b) ΔE over M_{bc} slice

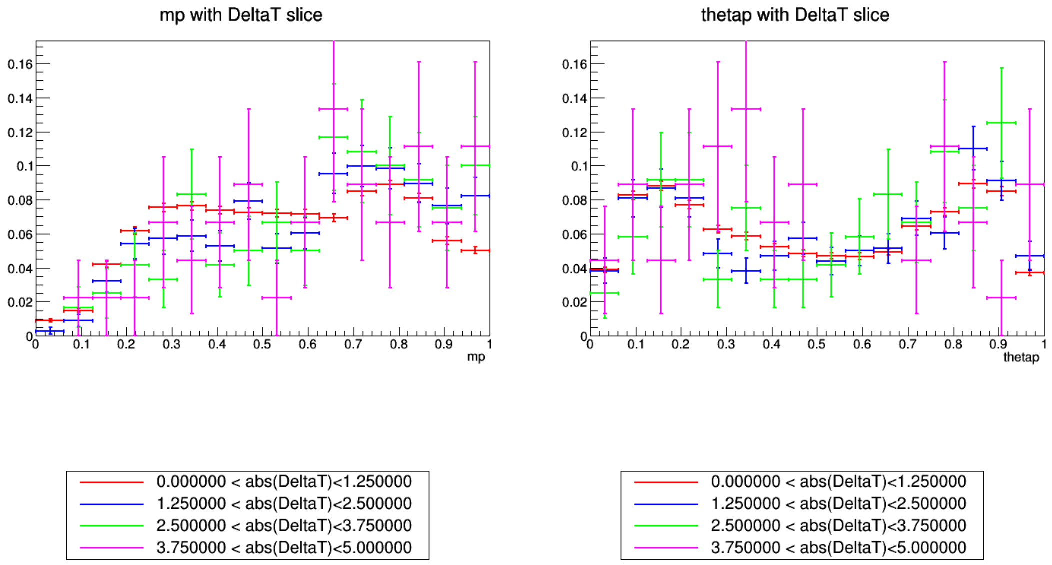
Figure 75: The distribution of M_{bc} with ΔE slices(left) and the distribution of ΔE with M_{bc} slices(right) for background events.



(a) Correlation $m' - \Delta t$

(b) Correlation $\theta' - \Delta t$

Figure 76: The 2-dimension scattering plot of $m' - \Delta t$ (left) and $\theta' - \Delta t$ (right) distribution in MC background events.



(a) m' over ΔT slice

(b) θ' over ΔT slice

Figure 77: The distribution of m' with Δt slices(left) and the distribution of θ' with Δt slices(right) for background events.

B KSFW Moment

KSFW moment [9] uses the energy and momentum of particles in the e^+e^- CMS frame to discriminate continuum events from signal candidates. Each of the track momentum is divided into three category : a charged particle, a neutral particle, and a missing particle. In addition to the normal tracks from particles, the missing momentum of an event is also treated as an additional particle. Then each KSFW moment is calculated as

$$H^{so}_{xl} = \sum_i \sum_{jx} |p_{jx}| P_l(\cos\theta_{i,jx}) \qquad (l = 0, 2, 4, x = 0, 1, 2) \tag{124}$$

$$H^{oo}_{l} = \sum_j \sum_{k} |p_{j}||p_k| P_l(\cos\theta_{j,k}) \qquad (l = 0, 2, 4) \tag{125}$$

$$H^{oo}_{l} = \sum_j \sum_{k} Q_j Q_k |p_{j}||p_k| P_l(\cos\theta_{j,k}) \qquad (l = 1, 3). \tag{126}$$

Here x denotes track categories as $x = 0, 1$, and 2 for charged, neutral, and missing track respectively. The index i runs over $B_{\rm CP}$ daughters, and jx over the $B_{\rm tag}$ daughters for each track category. The index j and k runs over $B_{\rm tag}$ daughter tracks. Q_i or Q_{jx} denotes the charge of each particle, and p_{jx} denotes the momentum of each particle. $P_l(\theta_{x,y})$ denotes the l-th Legendre polynomial of the cosine of the angle between particle x and particle y.

C Other local minimums

We also found 3 solutions other than the best local minimum discussed in the text. The fitting result for all the solutions are summarized in Table. 27. The summary of fit fractions for 2nd, 3rd and 4th best solution is shown in Table. 28, Table. 29 and Table. 30 respectively.

Parameter	1st	2nd	3rd	4th
a_{f_0}	3.44 ±0.51	2.68 ±0.56	4.17 ±0.44	3.90 ±0.41
a_{ϕ}	1.00 ±0.00	1.00 ±0.00	1.00 ±0.00	1.00 ±0.00
a_{f_X}	0.42 ±0.06	0.31 ±0.06	0.45 ±0.06	0.30 ±0.06
$a_{\chi_{c0}}$	0.24 ±0.02	0.25 ±0.02	0.21 ±0.03	0.24 ±0.02
$a_{(K^+K^-)_{NR}}$	7.29 ±0.54	6.91 ±0.59	4.73 ±0.65	2.51 ±0.58
$a_{(K^0_SK^+)_{NR}}$	4.20 ±0.22	3.95 ±0.22	3.79 ±0.26	3.78 ±0.26
$a_{(K^0_SK^-)_{NR}}$	3.63 ±0.22	3.86 ±0.22	2.78 ±0.28	2.84 ±0.27
b_{f_0}	-87.95 ±9.73	93.16 ±11.52	-117.61 ±8.54	38.35 ±9.65
b_{ϕ}	-33.48 ±0.00	-33.48 ±0.00	-33.48 ±0.00	-33.48 ±0.00
b_{f_X}	-113.09 ±9.08	55.71 ±11.59	-118.36 ±9.20	51.40 ±12.65
$b_{\chi_{c0}}$	-59.47 ±9.57	132.25 ±9.59	-45.14 ±13.51	170.18 ±10.03
$b_{(K^+K^-)_{NR}}$	-69.49 ±5.18	117.50 ±5.69	-84.06 ±5.59	121.65 ±11.58
$b_{(K^0_SK^+)_{NR}}$	56.63 ±5.52	14.93 ±6.44	36.56 ±6.31	66.34 ±7.05
$b_{(K^0_SK^-)_{NR}}$	-167.75 ±5.94	-119.70 ±5.98	-114.58 ±6.76	-144.32 ±7.60
c_{f_0}	0.18 ±0.15	0.08 ±0.19	0.05 ±0.14	-0.05 ±0.10
c_{ϕ}	-0.03 ±0.09	-0.06 ±0.08	-0.00 ±0.09	-0.03 ±0.09
$c_{\chi_{c0}}$	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00
c_{other}	0.08 ±0.05	-0.05 ±0.05	0.15 ±0.06	0.03 ±0.06
d_{f_0}	29.49 ±13.71	21.50 ±13.97	34.05 ±15.13	30.19 ±13.03
d_{ϕ}	28.89 ±10.12	22.30 ±10.69	25.50 ±13.29	22.71 ±10.50
$d_{\chi_{c0}}$	21.50 ±0.00	21.50 ±0.00	21.50 ±0.00	21.50 ±0.00
d_{other}	30.67 ±9.04	17.20 ±9.58	21.63 ±10.71	15.49 ±10.42
$-2\ln\mathcal{L}$	0	9.82	32.21	41.32

Table 27: The summary of all the solution found in this analysis, where the 1st solution is discussed detailed in the main text.

Fraction [%]	$f^0K_S^0$	ϕK_S^0	$f_XK_S^0$	$\chi_{c0}K_S^0$	$(K^+K^-)_{NR}K_S^0$	$(K^+K_S^0)_{NR}K^-$	$(K^-K_S^0)_{NR}K^+$
$f^0K_S^0$	8.56 ±3.92	0.00 ±0.00	-0.02 ±0.07	0.15 ±0.04	-28.38 ±13.31	-6.94 ±0.76	13.43 ±3.23
ϕK_S^0		14.98 ±1.40	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	-0.03 ±0.01	-0.08 ±0.01
$f_XK_S^0$			1.42 ±0.56	0.03 ±0.01	5.10 ±0.86	-2.18 ±0.56	1.61 ±0.55
$\chi_{c0}K_S^0$				3.29 ±0.61	-0.26 ±0.15	-1.24 ±0.18	1.29 ±0.26
$(K^+K^-)_{NR}K_S^0$					100.83 ±19.16	-14.48 ±1.75	-35.22 ±6.07
$(K^+K_S^0)_{NR}K^-$						32.83 ±4.64	-26.03 ±3.87
$(K^-K_S^0)_{NR}K^+$							31.35 ±4.49

Table 28: Values of the fit fraction FF_{ij} calculated from 2nd best CP fitting result in the unit of %.

Fraction [%]	$f^0K_S^0$	ϕK_S^0	$f_XK_S^0$	$\chi_{c0}K_S^0$	$(K^+K^-)_{NR}K_S^0$	$(K^+K_S^0)_{NR}K^-$	$(K^-K_S^0)_{NR}K^+$
$f^0K_S^0$	19.17 ±5.14	0.00 ±0.00	3.99 ±0.92	0.27 ±0.07	-22.85 ±11.03	19.24 ±4.23	-13.67 ±3.89
ϕK_S^0		13.85 ±2.02	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	-0.05 ±0.01	-0.06 ±0.01
$f_XK_S^0$			2.84 ±0.91	0.03 ±0.01	1.38 ±0.22	2.82 ±0.77	-1.61 ±0.52
$\chi_{c0}K_S^0$				2.15 ±0.69	-0.41 ±0.19	1.03 ±0.25	-0.71 ±0.22
$(K^+K^-)_{NR}K_S^0$					44.72 ±14.61	-21.07 ±6.00	26.14 ±7.97
$(K^+K_S^0)_{NR}K^-$						28.61 ±6.08	-21.22 ±4.96
$(K^-K_S^0)_{NR}K^+$							15.39 ±4.05

Table 29: Values of the fit fraction FF_{ij} calculated from 3rd best CP fitting result in the unit of %.

Fraction [%]	$f^0K_S^0$	ϕK_S^0	$f_XK_S^0$	$\chi_{c0}K_S^0$	$(K^+K^-)_{NR}K_S^0$	$(K^+K_S^0)_{NR}K^-$	$(K^-K_S^0)_{NR}K^+$
$f^0K_S^0$	17.45 ±5.14	0.00 ±0.00	3.51 ±1.30	0.24 ±0.06	10.33 ±3.88	-13.01 ±4.30	13.53 ±4.06
ϕK_S^0		14.43 ±3.11	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	-0.08 ±0.02	-0.04 ±0.01
$f_XK_S^0$			1.29 ±0.59	0.03 ±0.01	2.04 ±0.96	-1.21 ±0.55	1.37 ±0.54
$\chi_{c0}K_S^0$				2.93 ±0.80	-0.31 ±0.12	-1.20 ±0.32	0.66 ±0.18
$(K^+K^-)_{NR}K_S^0$					12.83 ±6.58	12.69 ±3.64	-1.18 ±0.81
$(K^+K_S^0)_{NR}K^-$						28.99 ±7.47	-21.64 ±5.98
$(K^-K_S^0)_{NR}K^+$							16.37 ±4.74

Table 30: Values of the fit fraction FF_{ij} calculated from 4th best CP fitting result in the unit of %.

D Model dependent constraint on the new physics parameter

In Sec. 9.6 we discussed model-independent constraint on the new physics parameter, but it is also possible to give a model-dependent constraint assuming certain new physics model. One of the new physics contributing to $B^0 \to \phi K_S^0$ is SUSY mass insertion as discussed in Sec. 2.4. The ratio of the SUSY amplitude against SM amplitude assuming $m_{\tilde{g}} \simeq 500\,\mathrm{GeV}$ is as follows [26]

$$\frac{A^{SUSY}(\phi K_S^0)}{A^{SM}(\phi K_S^0)} = (0.14+0.02i)\left((\delta_{LL}^d)_{23} + (\delta_{RR}^d)_{23}\right) + (65+11i)\left((\delta_{LR}^d)_{23} + (\delta_{RL}^d)_{23}\right). \quad (127)$$

Here $(\delta_{AB}^d)_{23}$ $((A,B)=(L,R))$ denote mass insertion terms.

We consider SUSY models with mass insertion where one of the mass insertion term is dominant. When we assume $(\delta_{LL}^d)_{23}$ to be dominant, the maximum SUSY amplitude is obtained to be $|A^{SUSY}/A^{SM}| \simeq 0.14$ [26]. Under assumption of $|A^{SUSY}/A^{SM}| = 0.14$, then we can give a constraint on the $\arg(\delta_{LL}^d)$ along with δ_{12}. The confidence regions over $(\arg(\delta_{LL}^d), \delta_{12})$ plane is shown in Fig. 78.

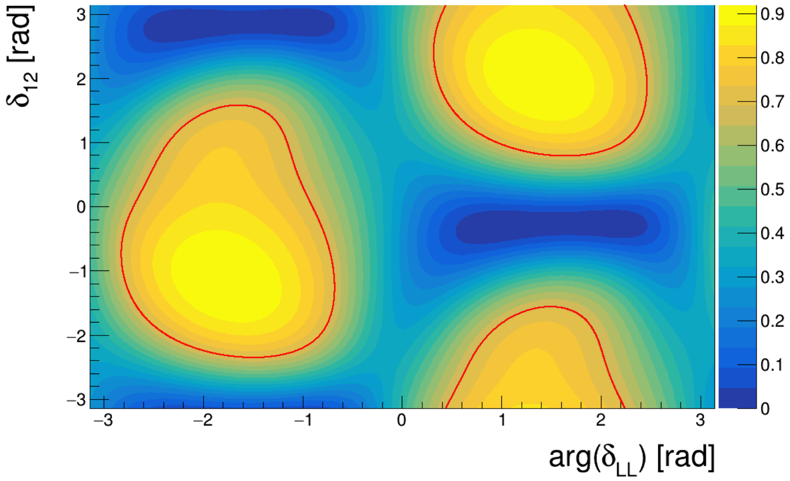


Figure 78: The confidence interval over $(\arg(\delta_{LL}^d), \delta_{12})$ plane. The red line shows 68.27% confidence level contour. The yellow (blue) region represents region with high (low) confidence level.