Measurement of the time-dependent CP asymmetry in $B^0 \to \rho^+\rho^-$ decays and constraint on the CKM angle ϕ_2 at the Belle II experiment

Belle II実験における$B^0 \to \rho^+\rho^-$崩壊の時間依存CP非対称度の測定とCKM角ϕ_2の決定

Ryogo Okubo*

Graduate school of science, Nagoya Univ., Nagoya, Japan

March 2025

* rokubo@hepl.phys.nagoya-u.ac.jp

ABSTRACT

The Standard Model(SM) of particle physics can explain nearly all phenomena in particle physics. However, it leaves several significant questions that are not solved, such as the matter-antimatter asymmetry in the universe and the nature of dark matter. One of the most important challenges in physics is to experimentally observe phenomena from beyond the Standard Model (BSM) to address these issues. The Kobayashi-Maskawa theory describes the weak interaction involving six types of quarks across three generations, and its charge-parity (CP) violation effects using a Cabibbo-Kobayashi-Maskawa (CKM) matrix. The CKM matrix is a unitary matrix with three rows and three columns, and its elements form a triangle on the complex plane, known as the unitary triangle. BSM physics can appear through quantum effects in B meson decays, which contain bottom quarks from the third generation. These effects can be detected as shifts in the parameters of the unitary triangle. Precise measurements of the unitary triangle allow for searching effects from higher energy regions without relying on theoretical models.

This study focuses on measuring one of the unitary triangle angles, specifically the CKM angle ϕ_2, using the decays of neutral B mesons into a pair of rho mesons ($B^0 \to \rho^+\rho^-$). The angle ϕ_2 is defined as $\phi_2 = arg(-V_{td}V_{td}^*/V_{ud}V_{ub}^*)$, and it can be determined from the CP violation parameters resulting from the interference between tree decays of $b \to u\bar{u}d$ and $B^0\bar{B}^0$ mixing. However, CP-violating effects also arise from the interference of $b \to d$ loop decays and the differing CP eigenvalues associated with the polarization of the vector mesons rho. Therefore, in addition to measuring the CP asymmetry, it is necessary to determine the branching fraction and the longitudinal polarization fraction. These measurements are combined with the results from the decays $B^+ \to \rho^+\rho^0$ and Bß$\rho^0\rho^0$. This approach will allow us to isolate the CP-violating effects by $b \to u\bar{u}d$ decays and $B^0\bar{B}^0$ mixing, enabling us to extract the value of ϕ_2.

In this study, we measure the branching fraction, longitudinal polarization fraction, and two time-dependent CP violation parameters in $B^0 \to \rho^+\rho^-$ decays using a data sample of 365 fb^{-1} collected by the Belle II experiment at SuperKEKB, located at KEK. During the data analysis, we suppressed the background from the $e^+e^- \to q\bar{q}$ process and the photon backgrounds in the $\rho^+ \to \pi^+\pi^-$, $\pi^0 \to \gamma\gamma$ processes, which tend to be misidentified. This was accomplished using machine learning techniques. In addition, we validated and calibrated the analysis using simulated samples and control samples, and then determined the four observables.

As a result, we obtain $\mathcal{B}(B^0 \to \rho^+\rho^-) = (2.90^{+0.23+0.31}_{-0.22-0.30}) \times 10^{-5}$, $f_L = 0.921^{+0.024+0.017}_{-0.025-0.015}$, $S = -0.26 \pm 0.19 \pm 0.08$, and $A_{\rm CP} = 0.02 \pm 0.12^{+0.05}_{-0.06}$. It is consistent with the previous experiments. In addition, we improved the precision of CP-violating parameter scaled by the data size by 10% thanks to the improved event selection. Moreover, we extracted ϕ_2 by combining our results with the result of $B \to \rho^+\rho^-.\rho^+\rho^0$, and $\rho^0\rho 0$in the previous experiments. As a result, we obtain $\phi_2 = (92.6^{+4.5}_{-4.8})^\circ$. This is consistent with the standard model. This study improved the precision of the world average in ϕ_2 extraction using $B \to \rho\rho$ by 6%.

CONTENTS

1. INTRODUCTION

The elementary particles are six quarks, six leptons, four gauge bosons, and the Higgs boson. Understanding of the elementary particles is promising for a fundamental understanding of the universe. The Standard Model (SM) of particle physics was developed thanks to previous great experimental and theoretical work, and it explains all experimentally observed phenomena in particle physics except for neutrino mixing. However, there are several open questions, such as the matter-antimatter asymmetry in the universe, dark matter, gravity, grand unification, and so on. Thus, the SM is an approximate model for particle physics and can be extended to solve these questions. Many theoretical models have been proposed, but none has been confirmed experimentally. Experimental observations of the phenomena from physics beyond the Standard Model (BSM) make the SM more realistic, making it one of the most important challenges in physics today. This thesis discusses the search for the BSM effects through the precise measurement of the particle-antiparticle asymmetry, in other words, charge-parity (CP) asymmetry in B decays.

1.1. Kobayashi-Maskawa theory

Kobayashi-Maskawa theory describes the flavor structure of the quarks. It describes the flavor-changing weak interaction of the quarks using the Cabbibo-Kobayashi-Maskawa (CKM) matrix which includes the couplings. In the SM, the Lagrangian of the $W^{\pm}$ boson interactions for quarks are

$$-\mathcal{L} = \frac{g}{\sqrt{2}} \overline{u_{Li}} \gamma^{\mu} (V_{\mathrm{CKM}})_{ij} d_{Li} W^{+}_{\mu} + h.c. \tag{1}$$

The parameter g is the coupling constant of the weak interaction, γ^{μ} is gamma matrices, and W^{+}_{μ} corresponds to the charged weak boson fields, and the subscripts $i, j = 1, 2, 3$ are generation numbers. Equation 2 shows the CKM matrix where V_{ij} is the coupling of the quarks when up type quarks ($i = u, c, t$) changes to down type quarks ($j = d, s, b$). [1]

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2}$$

The CKM matrix is unitary. The matrix has $(n-1)^2$ degree of freedom when the number of the families of the quark is n. $n(n-2)/2$ parameters are the rotation angles which are real numbers, and $((n-1)(n-2))/2$ is the parameter which is a complex phase that causes the CP violation. In the SM, the number of the quark family is three, and the CKM matrix has three real parameters and an imaginary parameter. There are several ways to describe the CKM using 4 parameters and using Wolfenstein parameterization [2], the CKM matrix can be written as

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

Equation 4 shows one of the unitarity conditions of the CKM matrix, and it was tested in the previous experiments [3].

$$V_{ud}V^{*}_{ub} + V_{cd}V^{*}_{cb} + V_{td}V^{*}_{tb} \tag{4}$$

Figure 1 shows the constraint on the Unitarity triangle was performed by the CKMFitter group[3]. All observables agree with the unitarity condition, and the theory of CKM has been successfully confirmed experimentally so far. Table 1 shows the result for the three unitarity triangle angles, and the ϕ_2 has the largest uncertainty.

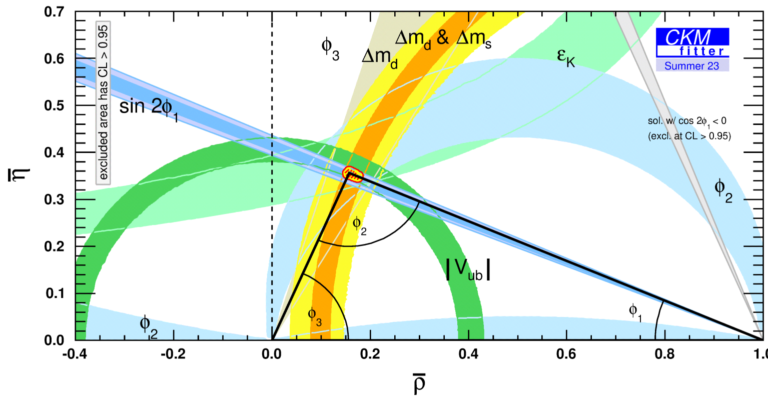


FIG. 1. Constraint on the unitarity triangle[3]

Parameter	Value
ϕ_1	$(22.53^{+0.33}_{-0.30})^\circ$
ϕ_2	$(86.2^{+3.9}_{-3.5})^\circ$
ϕ_3	$(65.9^{+3.3}_{-3.5})^\circ$

TABLE 1. World average for the three unitarity triangle angles [3]

1.2. Time-dependent CP violation in B^0 decays

Within the framework of the SM, CP violating effect occurs in B decays. Consider the decay when B^0 and $\bar{B}^0$ decay into the same states f, for example, $B^0, \bar{B}^0 \to \rho^+\rho^-$.

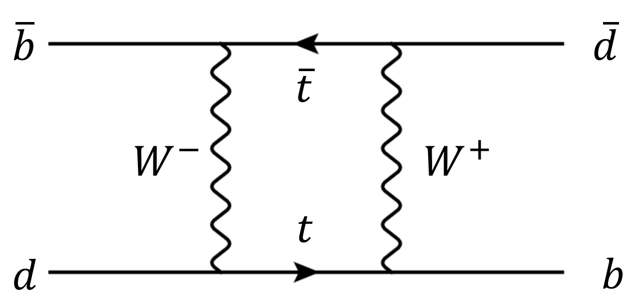


FIG. 2. Feinman diagram of the $B^0\bar{B}^0$ mixing. Only $t, \bar{t}$ that are dominant in the process are shown in the mixing diagram, but u and c quark also appear.

The effective Hermirtonian of the time evolution of neutral B mesons can be written as

$$\mathcal{H} = \mathcal{M} - \frac{i}{2}\Gamma, \tag{5}$$

where $\mathcal{M}$ is a mass matrix and Γ is a decay matrix, both are 2×2 matrix. The mass eigenstates of $\mathcal{H}$ are

$$|B_1\rangle = \frac{1}{|p|^2 + |q|^2}(p\,|B^0\rangle - q\,|\bar{B}^0\rangle) \tag{6}$$

$$|B_2\rangle = \frac{1}{|p|^2 + |q|^2}(p\,|B^0\rangle + q\,|\bar{B}^0\rangle). \tag{7}$$

The q/p is defined as

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} \tag{8}$$

The two eigenvalues are

$$\lambda_{1,2} = m_{1,2} - \frac{i}{2}\Gamma_{1,2}, \tag{9}$$

where $m_{1,2}$ is the mass and $\Gamma_{1,2}$ is the decay width. We consider the time-evolution of the B^0 meson states $\Psi(t)$ as

$$|\Psi(t)\rangle = b(t)\,|B^0\rangle + \bar{b}\,|\bar{B}^0\rangle. \tag{10}$$

The time-evolution of the states can be considered using the Hermirtonian $\mathcal{H}$ as

$$i\frac{d}{dt}\begin{pmatrix} b(t) \\ \bar{b}(t) \end{pmatrix} = \mathcal{H}\begin{pmatrix} b(t) \\ \bar{b}(t) \end{pmatrix} \tag{11}$$

solving the equation,

$$\begin{pmatrix} b(t) \\ \bar{b}(t) \end{pmatrix} = \begin{pmatrix} g_+(t) & -\frac{p}{q}g_-(t) \\ -\frac{p}{q}g_-(t) & g_+(t) \end{pmatrix}\begin{pmatrix} b(t=0) \\ \bar{b}(t=0) \end{pmatrix}. \tag{12}$$

Using $\Delta m = m_2 - m_1$, $\Delta\Gamma = \Gamma_2 - \Gamma_1$, the $g_\pm(t)$'s are

$$g_\pm(t) = \frac{1}{2}e^{-\frac{\Gamma_1}{2}t}e^{-im_1 t}\left[1 \pm e^{\Delta\Gamma t/2}e^{-i\Delta m t}\right]. \tag{13}$$

Then, we consider the decay of neutral B meson which decays whose final state doesn't depend on B^0 and $\bar{B}^0$. The amplitude of $B^0 \to f$ and $\bar{B}^0 \to f$ are defined as $\mathcal{A}(B^0 \to f)$ and $\mathcal{A}(\bar{B}^0 \to f)$. The decay can be written as

$$|\langle f|\,\mathcal{H}\,|B^0\rangle|^2 = |g_+(t)\mathcal{A}(B^0 \to f) - \frac{q}{p}\mathcal{A}(\bar{B}^0 \to f)|^2 \tag{14}$$

$$|\langle f|\,\mathcal{H}\,|\bar{B}^0\rangle|^2 = |g_+(t)\mathcal{A}(\bar{B}^0 \to f) - \frac{q}{p}\mathcal{A}(B^0 \to f)|^2. \tag{15}$$

Then, we consider the time-dependent CP asymmetry $\frac{\Gamma(B^0\to f)-\Gamma(\bar{B}^0\to f)}{\Gamma(B^0\to f)+\Gamma(\bar{B}^0\to f)}$. It is,

$$\begin{aligned}
&\frac{\Gamma(B^0 \to f)(t) - \Gamma(\bar{B}^0 \to f)(t)}{\Gamma(B^0 \to f)(t) + \Gamma(\bar{B}^0 \to f)(t)} \\
&= cos(\Delta m t)[\Gamma(B^0 \to f) - \Gamma(\bar{B}^0 \to f)] + 2\sin(\Delta m t)\Gamma(B^0 \to f)\mathrm{Im}\left(\frac{q}{p}\frac{\mathcal{A}(\bar{B}^0 \to f)}{\mathcal{A}(B^0 \to f)}\right) \\
&\equiv A_{\mathrm{CP}}\cos(\Delta m t) + S_{\mathrm{CP}}\sin(\Delta m_t)
\end{aligned} \tag{16}$$

The parameter A_{CP} has non-zero value when $|\mathcal{A}(B^0 \to f)| \neq |\mathcal{A}(\bar{B}^0 \to f)|$, which comes from weak or strong phase differences between $B^0 \to f$ and $\bar{B}^0 \to f$. The parameter $S_{\rm CP} = {\rm Im}\left(\frac{q}{p}\frac{\mathcal{A}(B^0 \to f)}{\mathcal{A}(\bar{B}^0 \to f)}\right)$. Because of the complex phase in the CKM matrix the amplitudes of $B^0 \to f$ and $\bar{B}^0 \to f$ and $\frac{q}{p}$ have the complex phase, and it leads to the time-dependent CP violation in B^0 decays.

1.3. Contribution of the physics beyond the standard model in $B^0\bar{B}^0$ mixing

CKM picture of flavor mixing and CP violation was confirmed experimentally, but there is room for the physics Beyond the Standard Model (BSM) to contribute to the quark mixing within the uncertainty of the current measurements. Figure 2 shows the Feynman diagram of $B^0\bar{B}^0$ mixing. It contains a "loop" diagram in the transition from B^0 to $\bar{B}^0$. In the SM, mostly t quarks contribute to the mixing diagram. If we consider the BSM, it can also contribute to the mixing amplitude via quantum effect replacing t and W. In such cases, the mass scale of the BSM is not limited to the mass of b quark, and it can contribute even if it was TeV or higher mass. If there are contributions by the BSM in $B^0\bar{B}^0$ mixing, it can be the source of the CP-violation other than SM, and it can be detected by the precise measurement of the CP violation in B decays. It is called an "Indirect search" for the BSM. It is a promising way to search for BSM because we can consider various BSM models such as supersymmetric particles, Lepto quarks, and charged Higgses. Also, the accelerator energy doesn't limit the energy scale of the BSM models because of the quantum effect.

To be more specific, observables that can be measured without using $B^0\bar{B}^0$ mixing and any loop process can extract the CKM matrix using pure SM processes. For example, $|V_{ud}|$, $|V_{us}|$, $|V_{cb}|$, $|V_{ub}|$ can be extracted using tree decays. The CKM angle ϕ_3 is also one of the parameters that can be extracted without BSM contributions. It can be measured using the direct CP violation as a result of the interference between $B^- \to D^0K^-$ ($b \to cs\bar{u}$) and $B^- \to \bar{D}^0K^-$ ($b \to u\bar{c}s$). Although it accesses the complex phase of the CKM matrix, it is expected to be the same as the SM because this process doesn't contain any loop contributions.

On the other hand, observables that use the loop process are sensitive to BSM. For example, ϕ_2 is measured using CP violation which comes from the interference of $b \to u\bar{u}d$ tree decays with and without $B^0\bar{B}^0$ mixing. If there is a contribution from the BSM, the CP violating parameters in $b \to u\bar{u}d$ will be shifted from the SM. Similarly, ϕ_1 is measured via interference of $b \to c$ decays with and without $B^0\bar{B}^0$ mixing, and it is also sensitive to BSM. Another parameter is Δm_d, which is a $B^0\bar{B}^0$ mixing frequency. The frequency is also shifted by BSM.

In the CKM measurement, some of the parameters are sensitive to the contributions only from SM and the others can prove the effects from BSM. By comparing the Unitarity triangle shape which is extracted by parameters that are sensitive to BSM or not, we can obtain a strong constraint to the BSM.

The study to constrain the BSM effect in $B^0\bar{B}^0$ mixing was performed by the CKMFitter group [4]. In the paper, they parameterized the BSM contributions to the $B^0\bar{B}^0$ mixing as

$$M_d = (M_d)_{\mathrm{SM}}(1 + h_d e^{2i\sigma_d}) \tag{17}$$

where M_d is the $B^0\bar{B}^0$ mixing amplitude, $(M_d)_{\mathrm{SM}}$ is $B^0\bar{B}^0$ mixing only considering the SM, h_d is the relative BSM magnitude to SM, and σ_d is the relative phase of BSM to SM. Figure 3 shows the constraint on BSM in $B^0\bar{B}^0$ mixing. More than 20-30% magnitude of the BSM to SM contribution is excluded. In addition, the paper studied the future sensitivity, and it will be improved to be less than 10% BSM contributions by improving the sensitivity of CKM parameters by using the ongoing B-factory experiments (Belle II, LHCb) data. Measurements of CKM parameters by previous experiments was a huge success, but improving the precision of the CKM parameters still has an important role in searching for the contributions from BSM using the quantum effect.

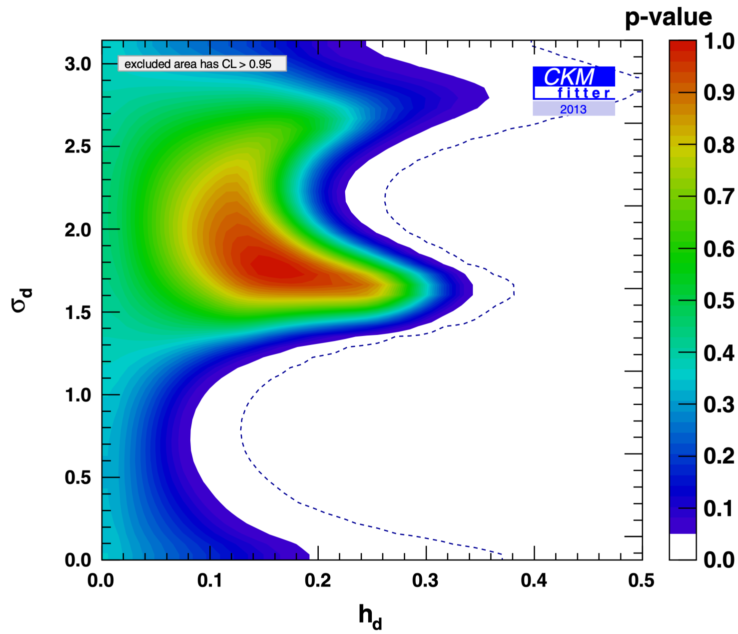


FIG. 3. Constraint on the BSM in $B^0\bar{B}^0$ mixing. h_d shows a relative magnitude of BSM compared to SM, and σ_d is the relative phase to SM of BSM. The dotted line shows 99.7% confidence interval regions.[4]

1.4. ϕ_2 measurement

There are many parameters in CKM matrix, but ϕ_2 is one of the most important observables in CKM because it is sensitive to the effects from BSM because of the $B^0\bar{B}^0$ mixing, and it has the largest uncertainty among the three angles in unitarity triangle. ϕ_2 can be written as

$$\phi_2 = \arg\left(-\frac{\mathrm{V_{tb}V_{td}^*}}{\mathrm{V_{ud}V_{ub}^*}}\right), \tag{18}$$

and it can be measured using the interference effect between $b \to u\bar{u}d$ decay and $B^0\bar{B}^0$ mixing. The interference effects appear as the CP-violating effect in the decay time difference. It is called a "time-dependent" CP violation, and it can be written as

$$\frac{\mathit{\Gamma}(\bar{B}^0(t) \to f) - \mathit{\Gamma}(B^0(t) \to f)}{\mathit{\Gamma}(\bar{B}^0(t) \to f) + \mathit{\Gamma}(B^0(t) \to f)} = S_{\mathrm{CP}} \sin(\mathit{\Delta} m_d t) + A_{\mathrm{CP}} \cos(\mathit{\Delta} m_d t) \tag{19}$$

where t is time, Δm_d is a mixing frequency of the $B^0\bar{B}^0$ mixing, and $\Gamma(\bar{B}^0(t) \to f)$ is a decay width of the $\bar{B}^0(t) \to f$ decay. A first, natural choice of the $b \to u\bar{u}d$ decay is $B^0 \to \pi^+\pi^-$. Only considering the tree $b \to u\bar{u}d$ decay and $B^0\bar{B}^0$ mixing, CP violating parameters of $B^0 \to \pi\pi$ are

$$\begin{aligned} S_{\mathrm{CP}}(B^0 \to \pi\pi) &= \sin(2\phi_2) \\ A_{\mathrm{CP}}(B^0 \to \pi\pi) &= 0 \end{aligned} \tag{20}$$

Time-dependent CPV measurement was performed at Belle Experiment [5], and Figure 4 shows the result.

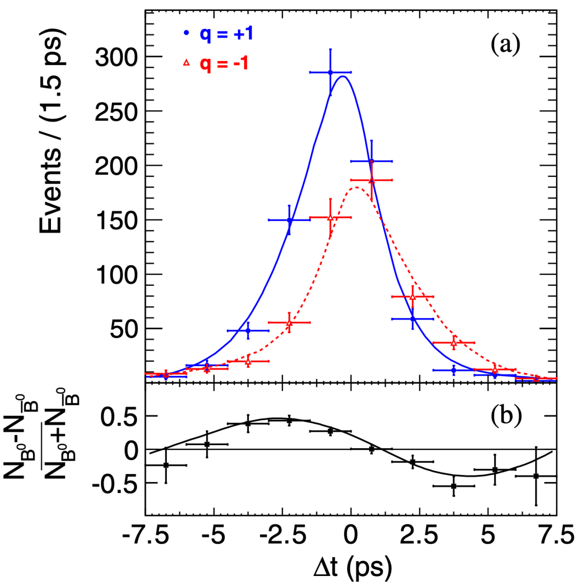


FIG. 4. Time dependent CPV measurement performed at Belle experiment [5]. (a) shows the Δt distribution for $B^0(q=1)$ and $\bar{B}^0(q=-1)$. (b) shows an asymmetry of (a).

Measured CPV parameters are

$$
\begin{aligned}
S_{\rm CP}(B^0 \to \pi\pi) &= -0.64 \pm 0.08(stat) \pm 0.03(syst) \\
A_{\rm CP}(B^0 \to \pi\pi) &= -0.33 \pm 0.06(stat) \pm 0.03(syst).
\end{aligned} \tag{21}
$$

The$A_{\rm CP}(B^0 \to \pi\pi)$ is significantly deviated from 0, it indicates that there are other contributions other than $b \to u\bar{u}d$ tree. Another contribution in $B^0 \to \pi^+\pi^-$ is the penguin decay, which is shown in Figure 5. The Penguin contribution in the $B^0 \to \pi^+\pi^-$ decay has a

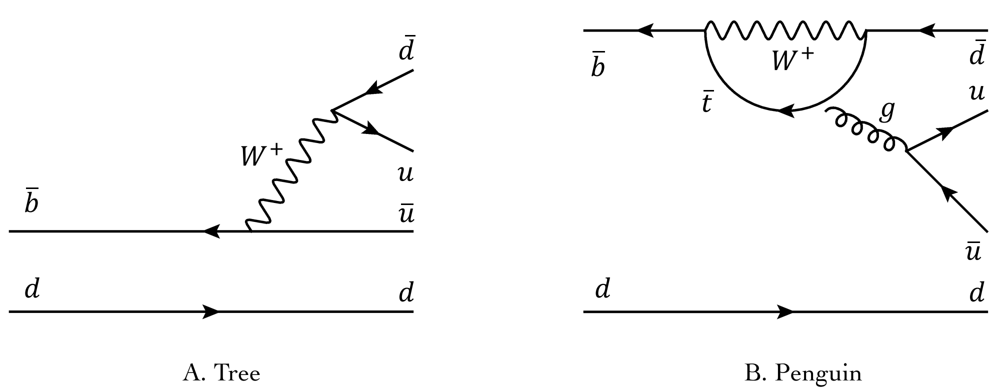


FIG. 5. Tree and Penguin contribution in $B^0 \to \pi^+\pi^-$ decay.

different phase to the tree decay, and it leads to non-zero $A_{\rm CP}$, and $S_{\rm CP}$ is also shifted from $\sin(2\phi_2)$ to $\sin(2\phi_2 + 2\Delta\phi_2)$. It is called "Penguin Pollution" in $B \to \pi\pi$ decay.
To extract ϕ_2, the Isospin analysis, which was suggested by [6], was performed. In the procedure, they used $B^0 \to \pi^+\pi^-$, $B^0 \to \pi^0\pi^0$, and $B^\pm \to \pi^\pm\pi^0$ to estimate the effects from the penguin. The states of pions are $|\pi^+\rangle = |I=1, I_3=1\rangle$, $|\pi^-\rangle = |1,-1\rangle$,

and $|\pi^0\rangle = |0,0\rangle$. Then the two pions final state. Firstly, $\pi\pi$ final states can have only $I = 0, 2$ states dute Bose statistics. State $|\pi^+\pi^-\rangle$ is $(\left|\pi_1^+\right\rangle \left|\pi_2^-\right\rangle + \left|\pi_1^-\right\rangle \left|\pi_2^+\right\rangle)/\sqrt{2}$. Writing $|1,1\rangle\,|1,-1\rangle = |1,-1\rangle\,|1,1\rangle = \sqrt{1/6}\,|2,0\rangle + \sqrt{1/3}\,|0,0\rangle$ using Clebsch Gordan coefficients, the $|\pi^+\pi^-\rangle$ states is

$$\left|\pi^+\pi^-\right\rangle = \sqrt{\frac{1}{3}}\,|2,0\rangle + \sqrt{\frac{2}{3}}\,|0,0\rangle\,. \tag{22}$$

Similarly, other final states are

$$\begin{aligned}\left|\pi^0\pi^0\right\rangle =& \sqrt{\frac{2}{3}}\,|2,0\rangle - \sqrt{\frac{1}{3}}\,|0,0\rangle\,,\\ \left|\pi^+\pi^0\right\rangle =& \sqrt{\frac{1}{3}}\,|2,0\rangle\end{aligned} \tag{23}$$

Then write the Hamiltonian and states of B^0 and B^+ as

$$\begin{aligned}\mathcal{H} =& \mathcal{A}_{3/2}\,|3/2,1/2\rangle + \mathcal{A}_{1/2}\,|1/2,1/2\rangle\,,\\ \left|B^0\right\rangle =& |1/2,-1/2\rangle\\ \left|B^+\right\rangle =& |1/2,+1/2\rangle\,.\end{aligned} \tag{24}$$

Using above, the amplitudes $\mathcal{A}_{ij} = \langle\pi^i\pi^j|\,\mathcal{H}\,|B^{i+j}\rangle$ of $B^0 \to \pi^+\pi^-$, $B^0 \to \pi^0\pi^0$, and $B^+ \to \pi^+\pi^0$ decays are

$$\begin{aligned}\mathcal{A}_{+-} =& \sqrt{\frac{1}{6}}\mathcal{A}_{3/2} + \sqrt{\frac{1}{3}}\mathcal{A}_{1/2}\\ \mathcal{A}_{+0} =& \sqrt{\frac{3}{4}}\mathcal{A}_{3/2}\\ \mathcal{A}_{00} =& \sqrt{\frac{1}{3}}\mathcal{A}_{3/2} - \sqrt{\frac{1}{6}}\mathcal{A}_{1/2}.\end{aligned} \tag{25}$$

Using Equation 25, the relationships of three amplitudes are

$$\begin{aligned}1/\sqrt{2}\mathcal{A}_{+-} + \mathcal{A}_{00} = \mathcal{A}_{+0}\\ 1/\sqrt{2}\bar{\mathcal{A}}_{+-} + \bar{\mathcal{A}}_{00} = \bar{\mathcal{A}}_{+0}.\end{aligned} \tag{26}$$

The loop diagram contributes to only $\Delta I = 1/2$ operator [6], thus the amplitude $\mathcal{A}_{3/2}$ comes from only the tree diagram, and the amplitude $\mathcal{A}_{1/2}$ comes from both tree and loop diagram. The $B^+ \to \pi^+\pi^0$ decays occur via only the tree diagram. It doesn't have any complex weak phases thus, we obtain $\mathcal{A}_{+0} = \bar{\mathcal{A}}_{-0}$ The relationship can be shown as triangles, as shown in figure 111. $\Delta\phi_2$ can be extracted by using the angle between the two isospin triangles.

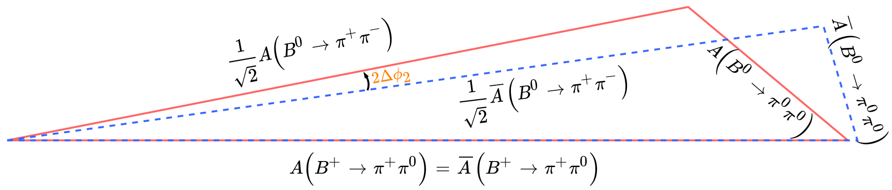


FIG. 6. Isospin triangle of $B \to \pi\pi$

The isospin analysis in $B \to \pi\pi$ decays was performed by Ref. [**?**]. The world average for the observables in $B \to \pi\pi$ decays are shown in Table 2. Due to the relatively large fraction of the loop amplitude in $B \to \pi\pi$ decays, $A_{\rm CP}$ of $B^0 \to \pi^+\pi^-$ has the non-zero value. In the Isospin analysis, the CP averaged side length of the triangle corresponds to the branching fraction of $B \to \pi\pi$ decays. A relatively large $\Delta\phi_2$ value is allowed because of the large $A_{\rm CP}$ in $B^0 \to \pi^+\pi^-$ decays and $\mathcal{B}$ in the $B^0 \to \pi^0\pi^0$ decays. Figure 7 shows the constraint on the ϕ_2 value using $B \to \pi\pi$ decays, and it has many solutions for ϕ_2 because of the large loop contribution leading a large $\Delta\phi_2$ value.

Observables	Values (World average)
$\mathcal{B}(\pi^+\pi^-)$	$(5.10 \pm 0.19) \times 10^{-6}$
$\mathcal{B}(\pi^+\pi^0)$	$(5.48 \pm 0.34) \times 10^{-6}$
$\mathcal{B}(\pi^0\pi^0)$	$(1.59 \pm 0.19) \times 10^{-6}$
$A_{\rm CP}(\pi^0\pi^0)$	0.34 ± 0.22
$S_{\rm CP}(\pi^+\pi^+)$	-0.672 ± 0.043
$A_{\rm CP}(\pi^+\pi^+)$	0.284 ± 0.039

TABLE 2. Observables for the Isospin analysis in the $B \to \pi\pi$ decays used in Ref. [7].

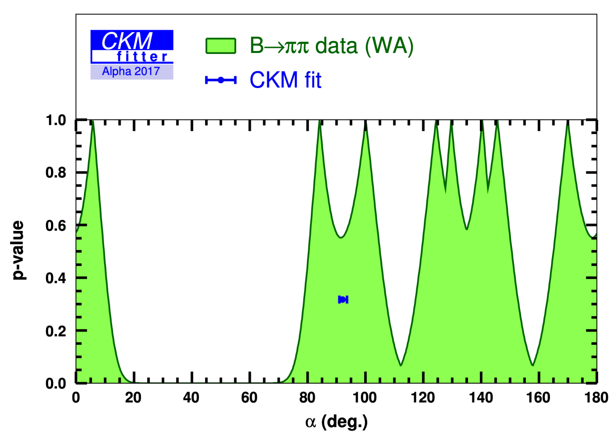


FIG. 7. Result of the Isospin analysis in $B \to \pi\pi$ decays performed by Ref. [7]

Another approach is to use the $B \to \rho\rho$ system. $B \to \rho\rho$ systems are experimentally more difficult than $B \to \pi\pi$ systems. $B \to \rho\rho$ decay has a polarization because ρ is a vector meson. The longitudinally polarized component is a CP-Even state, but the transversely polarized component is an admixture of the CP-Even and CP-Odd states. To extract the correct CPV parameters, the angular analysis for polarization extraction needs to be done simultaneously with the time-dependent CPV measurement. Overcoming the experimental difficulties, $B \to \rho^+\rho^-$ analysis was performed by **??**. The Isospin analysis in $B \to \rho\rho$ decays was also performed in Ref. [7]. The world average for the $B \to \rho\rho$ decays are shown in Table 3. The $A_{\rm CP}$ of $B^0 \to \rho^+\rho^-$ is much smaller than that of $B^0 \to \pi^+\pi^-$, and the branching fraction of the $B^0 \to \rho^0\rho^0$ compared to $B^0 \to \rho^+\rho^-$ is also much smaller than that of $B^0 \to \pi^0\pi^-$, that means the loop contribution is much smaller in $B \to \rho\rho$ decays compared to $B^0 \to \pi^0\pi^0$ decays. Experimentally, the $S_{\rm CP}$ of $B^0 \to \pi^0\pi^0$ has not been measured because π^0 decays into two photons dominantly, and it makes difficult to reconstruct the decay vertex position. In contrast, we can measure it in $B^0 \to \rho^0\rho^0$ decays because ρ^0 decays into two charged pions, which is another advantage in the Isospin analysis in $B \to \rho\rho$ decays. Figure 8 shows the result of the Isospin analysis performed in Ref. [7]. Thanks to the smaller loop contribution, the $\Delta\phi_2$ in $B \to \rho\rho$ system became much smaller, which is consistent with 0, and the Isospin analysis allowed only one solution. As a result, the ϕ_2 measurement in $B \to \rho\rho$ decay became most precise, that leads the ϕ_2 extraction.

Observables	Values (World average)
$\mathcal{B}(\rho^+\rho^-)$	$(27.76 \pm 1.84) \times 10^{-6}$
f_L	0.990 ± 0.020
$\mathcal{B}(\rho^+\rho^0)$	$(24.9 \pm 1.9) \times 10^{-6}$
f_L	0.950 ± 0.016
$\mathcal{B}(\rho^0\rho^0)$	$(0.93 \pm 0.14) \times 10^{-6}$
f_L	0.71 ± 0.006
$S_{\rm CP}(\rho^+\rho^+)$	-0.15 ± 0.13
$A_{\rm CP}(\rho^+\rho^+)$	0.00 ± 0.09
$S_{\rm CP}(\rho^0\rho^0)$	0.3 ± 0.9
$A_{\rm CP}(\rho^0\rho^0)$	-0.2 ± 0.9

TABLE 3. Result of the Isospin analysis in $B \to \rho\rho$ decays performed by Ref. [7]

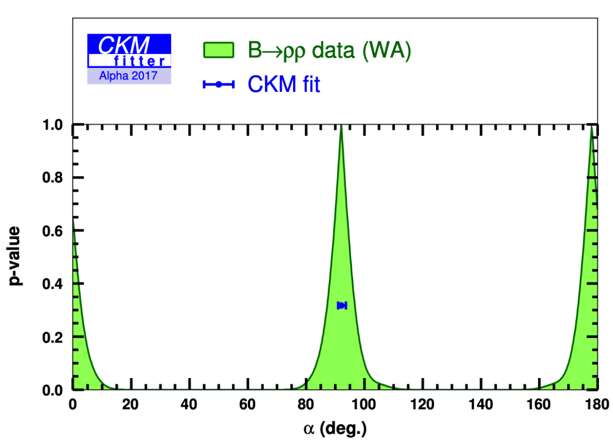


FIG. 8. Result of the Isospin analysis in $B \to \rho\rho$ decays performed by Ref. [7]

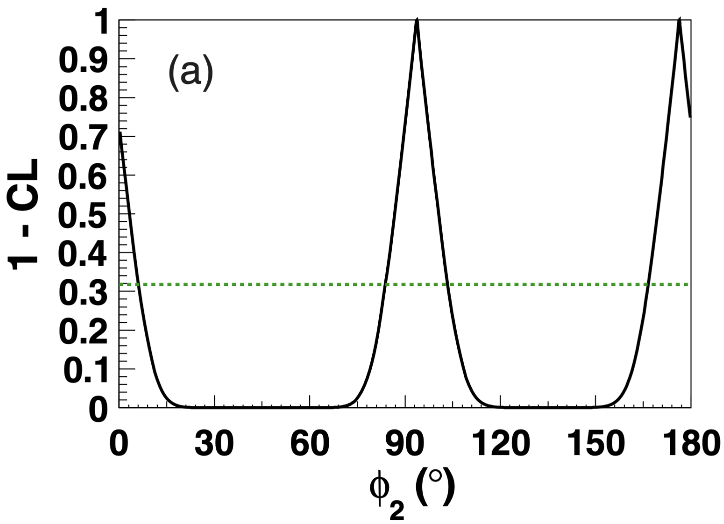


FIG. 9. ϕ_2 extraction using $B \to \rho\rho$ performed by [5]

1.5. Strategy of the study

In this analysis, we aim to measure ϕ_2 that has the largest uncertainty among three unitarity triangle angles. The $B^0 \to \rho\rho$ decays are promising to measureϕ_2, and the $S_{\rm CP}$ of $B^0 \to \rho^+\rho^-$ decays is one of the most important parameters to constrain ϕ_2 in $B \to \rho\rho$ Isospin analysis.

Table 4 shows the world average of the $B \to \rho\rho$ observables. We perform Isospin analysis using the world average, and Figure 10 shows the result of the ϕ_2 extraction. Also, we estimate the improvement by reducing the uncertainty of each observable by factor 2, and Figure 11 shows the expected improvement. The most promising observables are $S_{\rm CP}(\rho^+\rho^-)$ and $S_{\rm CP}(\rho^0\rho^0)$. Both are statistically limited and they can be improved by the huge statistics of the Belle II experiment.

Parameter	Value
$\mathcal{B}(\rho^+\rho^-)$	$(27.7 \pm 0.19) \times 10^{-6}$
$f_L(\rho^+\rho^-)$	$0.99^{+0.021}_{-0.019}$
$S_{cp}(\rho^+\rho^-)$	-0.14±0.13
$A_{cp}(\rho^+\rho^-)$	0.00±0.09
$\mathcal{B}(\rho^0\rho^0)$	$(0.96 \pm 0.15) \times 10^{-6}$
$f_L(\rho^0\rho^0)$	$0.71^{+0.08}_{-0.09}$
$S_{cp}(\rho^0\rho^0)$	0.3±0.7
$A_{cp}(\rho^0\rho^0)$	-0.2±0.9
$\mathcal{B}(\rho^+\rho^0)$	$(24.0 \pm 1.9) \times 10^{-6}$
$f_L(\rho^+\rho^0$	0.950 ± 0.016

TABLE 4. Inputs for the ϕ_2 extraction using world average[8]

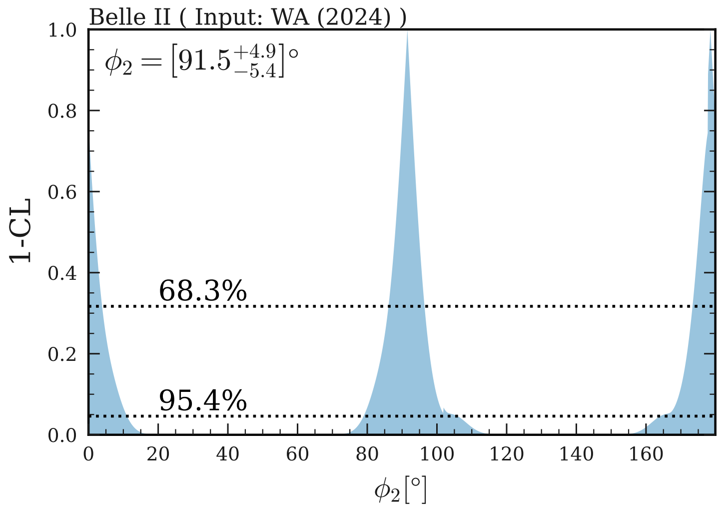


FIG. 10. ϕ_2 extraction using the world average

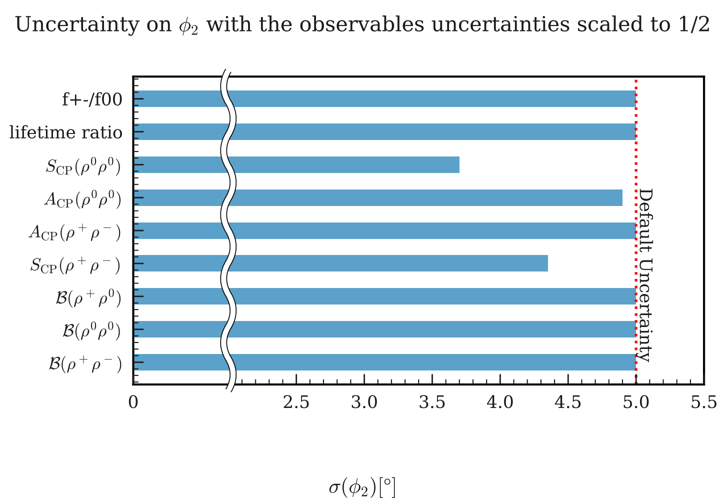


FIG. 11. Expected improvement in ϕ_2 when the uncertainty of each observable is improved by factor 2

In this study, we aim to measure $\mathcal{B}$, f_L, $S_{\rm CP}$, and $A_{\rm CP}$ of $B \to \rho^+\rho^-$ decay at Belle II experiment. The most important parameter is the $S_{\rm CP}$, and it is statistically dominated. Thus, by adopting the latest machine learning techniques which did not exist when the previous studies were performed, the statistical precision can be improved significantly.

Firstly, we improved the selection. The major backgrounds are $ee \to qq$ backgrounds and mis-reconstructed photons in $\rho \to \pi\pi^0(\to 2\gamma)$. To suppress such backgrounds, we developed dedicated Multivariate analyses(MVA)s. Then, we optimized all cut-based event selections for example, the energy, the angle, invariant masses, and MVA outputs for decay particles from the $B \to \rho\rho$ decay.

The next step is signal extraction. In this step, we extracted the longitudinal polarization and branching fraction using a 6-dimensional fitting to ΔE which is a difference between reconstructed energy and the beam energy, $m_{\pi\pi^0}$ which is an invariant mass of each ρ meson,

$\cos\theta_\rho$ which is a cosine helicity angle of the each ρ, and output of the MVA which suppress $ee \to qq$ backgrounds. Figure 12 shows the cosine helicity angle of the $B \to \rho\rho$ decay, it is defined as the angle between the direction of the B^0 and the direction of $\pi^\pm$ in $\rho^\pm$ rest frame. The longitudinal polarization makes a peak at $\cos\theta_{\rho^\pm} = \pm 1$, and the transverse polarization makes a peak at $\cos\theta_{\rho^\pm} = 0$.

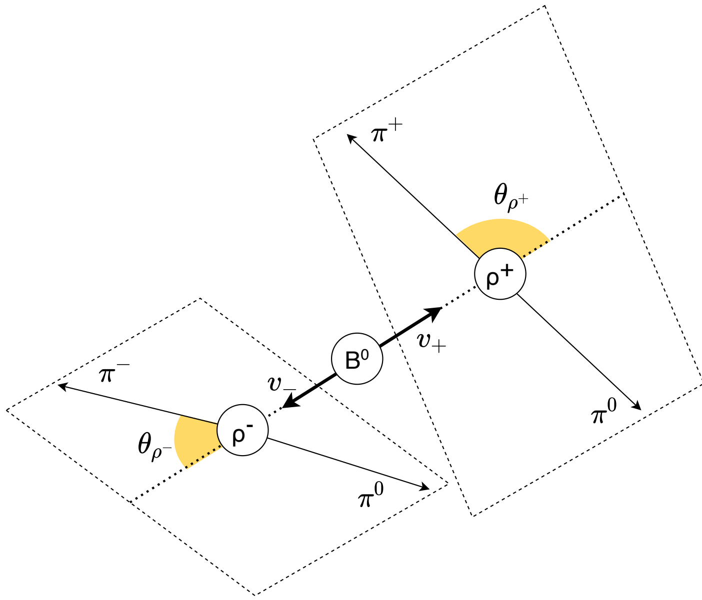


FIG. 12. Helicity angle of the $B \to \rho\rho$ decay. $\theta_{\rho^\pm}$ is the helicity angle of each ρ, and $v_\pm$ shows the direction of B^0 in the $\rho^\pm$ rest frame.

After extracting the branching fraction and longitudinal polarization, we measured the time-dependent CPV that is shown in Equation 19. In the Belle II experiment, the center-of-mass energy is set to 10.58 GeV, which is the same as the mass of $\Upsilon(4S)$ resonance. $\Upsilon(4S)$ decays in to $B^0\bar{B}^0$ pair. The produced $B^0\bar{B}^0$ state is entangled, and when one changes the CP by the $B^0\bar{B}^0$ mixing, the other also changes its states by the mixing. On the other hand, the $B^0\bar{B}^0$ state changes to the $\bar{B}^0 B^0$ state by the mixing, not to the $\bar{B}^0\bar{B}^0$ or $B^0 B^0$ states. In addition, the beam energy is asymmetric and the $\Upsilon(4S) \to BB$ is boosted to $\beta\gamma = 0.28$. The boosted BB pair enables time-dependent CPV measurement of B decay. Firstly, the B^0 is pair-produced, one decays into the signal decay, and another one decays into generic final states of B^0 decay. Using the particles from the decay which is not a signal, we can identify the flavor of B meson (B^0 or $\bar{B}^0$). In addition, by measuring the distance of the decay vertex which is projected onto the boost direction, we can evaluate the decay time difference Δt of the two B^0s. Measured Δt was affected by several reasons. The flight length of B^0 meson is about $\tau_{B^0}(\beta\gamma c) \sim \mathcal{O}(100\ \mu\mathrm{m})$, and measured decay vertex distance is affected by the detector resolution. In addition, the Δz is projected onto the boost axis, and this approximation also affects the measured Δt. To take these into account, we modeled the Δt resolution of the detector and approximation, then validated and calibrated it using data. Then, using the signal probability which was obtained from the fit for the branching fraction and longitudinal polarization and the model for the detector resolution, we performed a fitting to Δt and measured S_{CP} and A_{CP}.

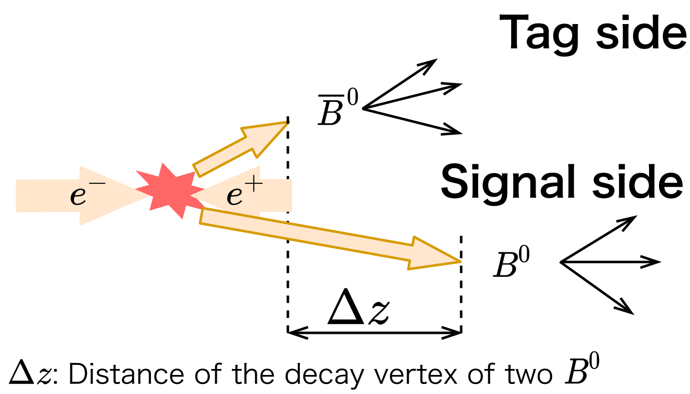


FIG. 13. Time-dependent CPV measurement at Belle II experiment

Finally, we performed the first ϕ_2 extraction at Belle II experiment with the $\mathcal{B}$, f_L, $S_{\rm CP}$, and $A_{\rm CP}$ of $B \to \rho\rho$. Also, we will discuss the constraint on the BSM in $B^0\bar{B}^0$ using the existing results and updated $B \to \rho\rho$ measurement performed by this study.

2. BELLE II EXPERIMENT

2.1. Overview of the Belle II experiment

The SuperKEKB/Belle II experiment is an electron-positron collider experiment at the center of mass energy 10.58 GeV. The main target of the experiment is for the precise measurement of B meson, especially the parameters of the CKM matrix. The center of mass energy corresponds to the resonance of $\Upsilon(4S)$, it decays to the B meson pair immediately after the production. It aims to accumulate 50 ab^{-1} of the integrated luminosity which corresponds to fifty times larger than that of the previous project, Belle Experiment, by achieving the world-highest peak luminosity of $6.5 \times 10^{35}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The Belle II detector is also upgraded so that provides enough performance in a high-rate environment due to high luminosity. By using a huge number of B mesons, we can measure B decays very precisely. The major physics goal is to measure the CKM parameters including ϕ_2, rare B decays search, precise CPV measurement in penguin decays, and so on. Another aspect of the Belle II experiment is the search for the lepton flavor violation decay of *tau*, precise measurement of τ, D meson by the high-statistics. By the precise measurement of such decays, it can seek a small deviation from the SM due to the physics beyond the standard model.

2.2. SuperKEKB

The SuperKEKB is an electron-positron collider located at Tsukuba, Japan. It consists of two circular accelerator rings, one is for a positron whose energy is 4 GeV, and the other is for an electron whose energy is 7 GeV. The asymmetric beam energy is for the decay time difference measurement of $B^0\bar{B}^0$ pairs by measuring the decay position difference in the boost axis. The unique point of the SuperKEKB is a way of the collision which is called a "nano-beam scheme". The idea is to squeeze the beam size to $\mathcal{O}$(nm) level at the collision to increase the luminosity. In general, when extremely focusing the beam, it spreads rapidly beyond the focusing point which is one of the causes of the luminosity limitation called as "hourglass effect". The idea is to squeeze the beam while avoiding the "hourglass effect", it collides beams with a relatively large crossing angle of 83 mrad. The SuperKEKB operation with physics data taking started in 2019, and it achieved the world-highest luminosity of $4.71 \times 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ on June 22, 2022. The integrated luminosity by May 2024 is 499.51fb^{-1}, which corresponds to about half of the Belle Experiment.

2.3. Belle II detector

The Belle II detector is a general-purpose detector that detects the particles produced by the e^+e^- collision, which is shown in Figure 14. It consists of the vertex detector, central drift chamber, particle identification system in the barrel and the forward end cap regions, the electromagnetic calorimeter, K_L, and μ detector, and 1.5 T solenoid magnet. This section shows the overview of each detector. The designed performance of the detector is shown in Table 5.

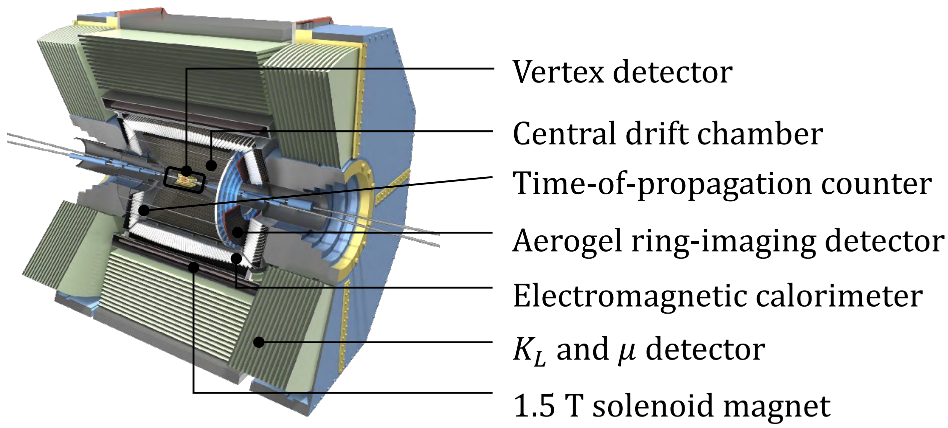


FIG. 14. The Belle II detector.

Detector	Performance
Vertex detector	$\sigma \simeq 20$ μm vertex resolution
CDC	$\sigma_{r,\phi} = 100$ μm $\sigma_z = 2$ mm $\sigma_{dE/dx} = 5\%$
TOP	99% K efficiency with $< 0.5\%$ π fake rate
ARICH	96% K efficiency with 1% π fake rate (4 GeV)
ECL	$\sigma E/E = 0.2\%/E \oplus 1.6\%/\sqrt[4]{E} \oplus 1.2\%$
KLM	$\Delta\phi, \Delta\theta = 20$ mrad for K_L(Barrel) $\Delta\phi, \Delta\theta = 10$ mrad for K_L(Endcap) $\sim 1\%$ hadron fake rate (μ)

TABLE 5. Designed performance of the Belle II detector [9]

2.3.1. Vertex detector

A Vertex detector is a detector which is located at the innermost part of the Belle II detector, which detects the charged particle from the interaction and reconstructs the decay vertex of B meson precisely for the decay-time difference measurement of $B\bar{B}$ pairs. It consists of two types of detectors, one is a Pixel Detector (PXD), and the other is a

Silicon Vertex Detector (VXD). PXD is located at the inner part of the vertex detector, and it consists of the two layers of the pixelated semiconductor sensors. Considering the environment of the Belle II experiment, the interaction point is very narrow and the interaction rate is very high. The signal detection rate increases by inversesquare of the distance from the vertex position, so the PXD which is located innermost part of the detector is required to handle the very high hit rate. To realize it, the pixelized sensor was adopted for the detector. The pixelized silicon sensor was successfully achieved in the LHC experiment, however, it was impossible to use the same sensor in the Belle II experiment. It is because the beam energy at Belle II is much smaller than that of LHC, and the effect of the multiple scattering becomes serious because the sensor is too thick for it. The DEPleted Field Effective Transistor (DEPFET) technique was on the semiconductor sensors for PXD, which achieve only 75 μm thickness.
The other part of the VXD is SVD. It has 4 layers of the detector, and it covers from 38 mm to 140 mm of radii, which is much larger than that of the PXD. To cover the large volume while keeping the resolution, a double-sided silicon detector was adopted. The strip is installed in two directions in each layer, one is a long strip with the same direction as the z axis, and the other is a short strip with a vertical direction to the beam axis. Combining the hit information of the two strips can measure the hit point of the charged particle. The strip pitch size in the beam direction is $50 - 70$ μm , and $75 - 240$ μm in the vertical direction to the beam axis depending on the radius.

2.3.2. Central Drift Chamber

The central drift chamber is a large gaseous tracking detector located at $160 - 1100$ mm of the radius. It has two important roles, one is the reconstruction of the charged track and the measurement of its momentum from the radius of the charged track in the magnetic field of 1.5 T, and the other is the particle identification from the energy loss in the gas volume. It consists of 14336 sense wires, and the gas is a mixture of 50% helium and 50% ethane.

2.3.3. Time-Of-Propagation Counter

The Time-Of-Propagation (TOP) counter is a particle identification detector located in the barrel region of the detector. It consists of quartz radiators and high time resolution photodetectors. Charged particles produce Cherenkov light, whose opening angle depends on the velocity of the particles. The momentum is measured by other detectors, so the opening angle depends on the type of particle. TOP measures the hit time and position of the Cherenkov photons, and evaluates the probability of each particle from the hit patterns, and identifies the particles. Micro-Channel-Plate Photo-Multiplier Tubes (MCP-PMT) is the photodetector for the TOP counter; it measures photon timing with a resolution of 30 ps; it gives excellent particle identification performance. One of the issues for the TOP counter is the lifetime of the photocathode of MCP-PMTs. The Quantum Efficiency (QE) will decrease by the accumulated output charge density of MCP-PMTs due to the outgassing from the MCPs. We have worked to improve the lifetime of the photocathode and installed in the TOP counter three tube types, the so-called Conventional type, Atomic Layer Deposition (ALD) type, and Life-extended ALD type. The expected lifetime is shown in Table 6. The Life-extended ALD type is expected to be able to be used until the end of the Belle II experiment, but the Conventional one needs to be replaced a few years after starting data taking, and the ALD type also needs to be replaced in the future.

Type	Lifetime (C/cm^2)
Conventional	1.1 (Average)
ALD	10.4 (Average)
Life-extended ALD	13.6 (Minimum)

TABLE 6. Expected lifetime per MCP-PMT type measured on test samples.

2.3.4. Aerogel ring imaging Cherenkov counter

Aerogel Ring Imaging Cherenkov Counter (ARICH) is another detector for particle identification. It is located at the forward endcap of the Belle II detector, which corresponds to the downstream in the boost direction. The aim is the same as the TOP counter, but the location is different. It consists of a large aerogel that covers the endcap and photodetectors which are sensitive to the position. While TOP reconstructs the opening angle of the Cherenkov ring by the hit position and hit timing, the ARICH mainly measures it by hit patterns. One of the unique structures of ARICH is an aerogel radiator, which consists of two layers of aerogel with a different refractive index. By doing this, it can focus the Cherenkov ring at the photodetector plane. It also allows thicker aerogel layers and it helps to increase the number of the Cherenkov photons. The photodetector for ARICH is Hybrid Avaranche Photo Diode (HAPD). The avalanche photodiode sensor is put in the vacuum tube which has a photocathode. The photoelectron is generated at the photocathode, and it is accelerated by the electric field. It is detected at the avalanche photodiode sensor. The pixel size of the avalanche photodiode is $4.9 \times 4.9\text{mm}^2$ which is small enough to detect the position of the Cherenkov light.

2.3.5. Electromagnetic Calolimeter

The Electromagnetic Calorimeter (ECL) is located outside one layer of the TOP counter, and it covers not only the barrel region but also the end cap region. It consists of the thallium-doped cesium iodide crystal for the scintillator, and the PIN photodiode for the scintillation light detector. ECL plays many important roles in the Belle II experiment. It can monitor the luminosity, identify the electron, and give a trigger timing for the detector. In addition, the most important role is the photon detection. It can detect photons with a nice energy and angle resolution, and it allows to reconstruct π^0 efficiently. Belle II experiment has a clean environment compared with the hadron collider, and many decay modes that have π^0 in the final states can only be accessed by Belle II. In such Belle II unique decay including $B^0 \to \rho^+\rho^-$, the ECL plays one of the most important roles.

2.3.6. K_L and μ detector

K_L and μ detector located at the outermost layer of the Belle II detector, which detects K_L and μ. It has a structure in which iron plates and detectors are arranged alternately. It can identify K_L and μ because they are not absorbed by the crystal of ECL while the other except for the neutrino do mostly. The K_L makes a shower in the iron plate while the μ doesn't. By measuring the shape of the showers it can identify K_L and μ. The detection of μ plays an important role in the analysis because the charge of lepton from B^0 decays is useful input to identify the flavor of B^0 .

2.4. Flavor tagging

To measure the time-dependent CP violation of the B^0 decay, which is shown in Section 1.4, the flavor of B^0 needs to be identified. In the Belle II experiment, B^0 is pair produced. Using the remaining particles which are not used for the signal reconstruction, we can estimate the flavor. For example,$B^0 \to D^{*-}l^+\bar{\nu}$ decay can identify the flavor from the charge of the lepton, because only positively charged lepton can produced in $\bar{b} \to \bar{c}$ decay. Another example is $B^0 \to \bar{D}^0(\to K^+\pi^-\pi^0)\pi^0$. It can identify the flavor from the charge of $K^{\pm}$. The particle identification performance is important in flavor identification by the hadronic modes because the flavor can be identified from the charge of $K^{\pm}$ in many cases.

To identify the flavor of B^0 efficiently, a machine learning algorithm is adopted.[10] It is based on a Graph-neural-network structure, and it uses the information of the tracks as inputs. It uses the first 16 particles that do not belong to the signal side ordered by the momentum in the Lab frame. Then, the momentum, the distance of its Points Of Closest Approach (POCA) from the interaction point, the particle identification likelihood for 6 candidates, which are π, K, e, μ, proton, and deuteron of each particle were used for the inputs for the algorithm. The flavor tagger output using a Fast BDT algorithm which is developed before the GNN flavor tagger [11] is also used as an input for the GNN flavor tagger. The performance of the GNN flavor tagger is shown in Figure 15. $q = \pm 1$ corresponds to the flavor of the B^0, and r is the probability of it. The product of q and r is shown as flavor tagger output. The flavor tagger is calibrated using $B^0 \to D^{*-}\pi^+$ data, and the result is shown in Figure 16. The effective tagging efficiency estimated by the data is $(37.40 \pm 0.43 \pm 0.36)\%$, where the first uncertainty is statistical, and the second uncertainty is systematic.

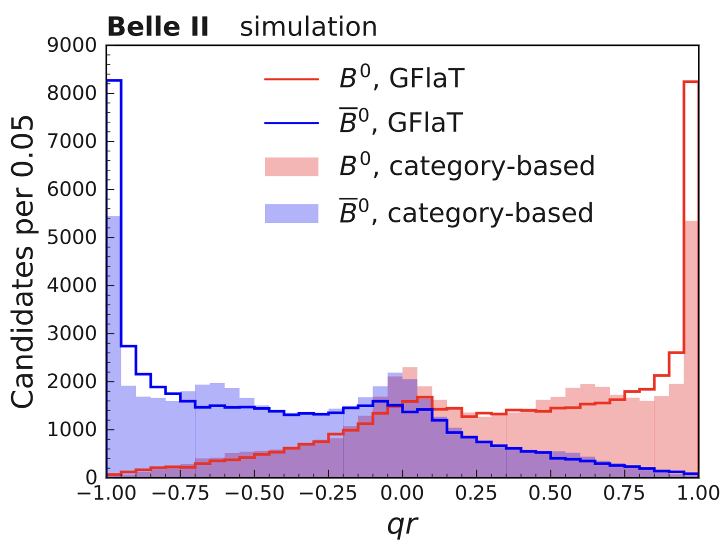


FIG. 15. The output for the flavor tagging. qr shows the output for the flavor tagger, and 1 corresponds to the B^0, and -1 corresponds to the $\bar{B}^0$. GFlaT shows the performance of the GNN-based flavor tagger, and category-based shows the performance of the flavor tagger developed using Fast BDT, which is a prior flavor tagger of GFlaT. [10]

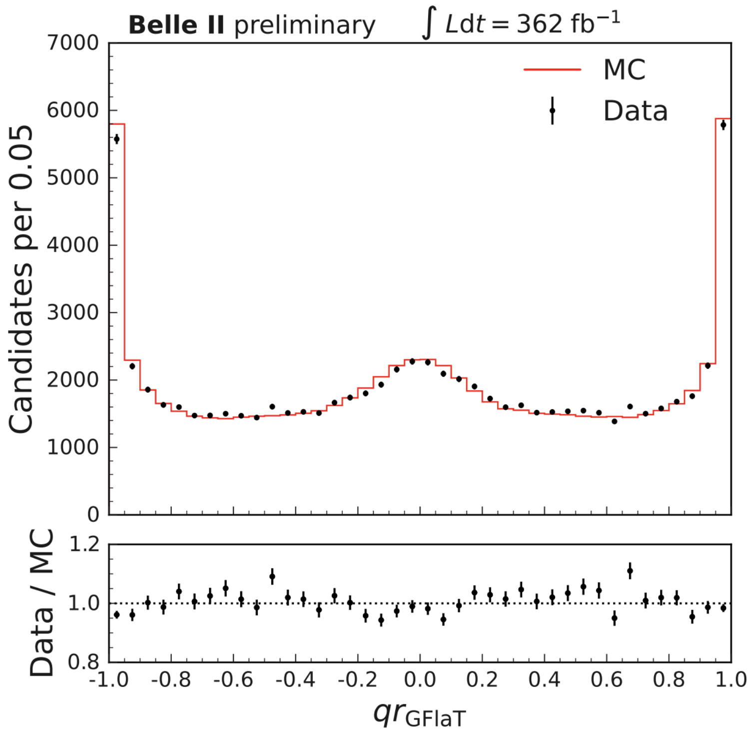


FIG. 16. Comparison between data and simulation after subtracting the background using $B^0 \to D^{*-}\pi^+$ decay.

3. OPTIMIZATION OF THE SIGNAL SELECTION

3.1. Samples

We used several samples for this study. We used simulated samples for $B\bar{B}$, $q\bar{q}$ ($q = udsc$), $B^0 to \rho^+\rho^-$ signals, and B decays that can affect the $B^0 to \rho^+\rho^-$ measurements. The simulated sample consists of two categories, one is run-dependent and the other is run-independent. The run-dependent sample takes into account the changes during the data taking, such as detector and background conditions, and the run-independent sample doesn't. On the other hand, run-independent samples are generated faster than run-dependent samples, so we used run-independent samples for the selection criteria optimization, and used run-dependent samples for the PDF modeling, validation, and fitter studies. The data was taken with two conditions, one is $\sqrt{s} = 10.58$ GeV which corresponds to the $\Upsilon(4S)$ resonance, and the other was taken 60 MeV below the $\Upsilon(4S)$ resonance for background studies.

Simulation	
Category	Sample size
$B\bar{B}, q\bar{q}$ (Run-independent)	1 ab^{-1}
$B\bar{B}, q\bar{q}$ (Run-dependent)	1.44 ab^{-1}
$B^0 \to \rho^+\rho^-$(Run-independent)	1.0×10^6 events
$B^0 \to \rho^+\rho^-$(Run-dependent)	5.0×10^6 events
B decays	$5.0 \times 10^5 - 5.0 \times 10^6$ events
Data	
Category	Sample size
$\Upsilon(4S)$	365 fb^{-1}
Off-resonance	42.7 fb^{-1}

TABLE 7. Data and simulated samples used for this study

3.2. Overview of the reconstruction and signal selection optimization

We optimized the event selection to maximize the statistical sensitivity of the observables $\mathcal{B}$, f_L, S_{CP}, and A_{CP}. To achieve this, we maximize the figure of merit (FoM) $S/\sqrt{S+B}$ in the signal-enhanced region. The signal-enhanced region is a tightened fit region that maximizes the FoM, S and B are signal and background yields, respectively. In the optimization of the selection process, we utilized both the $B^0 \to \rho^+\rho^-$ signal and its self-crossfeed events from the signal MC samples, and the other background events from the simulated samples.

The procedure for the optimization is below.

1. Train multivariate analysis (MVA)
2. Optimize cut-based event selection
3. Calculate FoM after applying the optimized selection and best candidate selection in the signal-enhanced region. The cut for signal enhancement using ΔE, M_{bc}, and CSMVA is also optimized to maximize the FoM.

4. Apply the optimized selections to the training data for MVAs, and back to 1.

In this process, MVAs are trained on samples passing through an optimized selection in the last iteration, and therefore, MVAs can be specialized in samples that are likely to remain after the other selection criteria have been applied. By training MVAs and optimizing selections in a series of processes, we aim to maximize the FoM efficiently. The details are described below.

3.3. Reconstruction and baseline selection

We reconstructed $B^0 \to \rho^+\rho^-$, $q\bar{q}$ ($e^-e^+ \to u\bar{u}, d\bar{d}, c\bar{c}$, and $s\bar{s}$) simulated samples with a loose event selection before the selection optimization. To efficiently extract samples of interest and make the analysis smooth, we first reconstruct $B^0 \to \rho^+\rho^-$ and skim more signal-like events with a loose selection. The decay chain of $B^0 \to \rho^+\rho^-$ is given by,

$$\begin{aligned} B^0 &\to \rho^+\rho^- \\ \rho^\pm &\to \pi^\pm\pi^0 \\ \pi^0 &\to \gamma\gamma. \end{aligned}$$

Its final state contains π^+, π^-, and four γs, and so two charged tracks and four ECL clusters can be detected in the Belle II detector system. We reconstruct $B^0 \to \rho^+\rho^-$ as follows:

1. reconstruction of two π^0s from ECL clusters
2. detection of two charged tracks of $\pi^\pm$ in CDC
3. reconstruction of two $\rho^\pm$s in the vertex fit with $\pi^\pm$ and π^0, and
4. reconstruction of B^0 from two ρs.

In this process, we apply cuts at each reconstruction stage. Table 8 shows the pre-selections. The E is the energy of the particles, dr and dz are the distance-of-closest-approach to the interaction point (IP) in the beam axis and the plane perpendicular to it, chiProb is a χ probability of the track. M_{bc} is a beam-constrained mass that corresponds to $M_{\mathrm{bc}} = \sqrt{E_B^2 - p_B^2}$ where E_B is half of the beam energy and p_B is a momentum of the reconstructed signal B^0, and $\Delta E = E - E_{\mathrm{beam}}/2$ is a difference of the beam energy E_{beam} and reconstructed energy$E$ in the center of mass frame.

3.4. Continuum suppression

Continuum suppression MVA (CSMVA) is used to suppress $e^+e^- \to q\bar{q}$ backgrounds ($q = u, d, s, c$). The MVA software applied in this analysis is TabNet [12], which is a kind of neural-network-based machine learning algorithm. The features used for the training are shown in Table 9. The training target is a binary classification for correctly reconstructed signals and all $ee \to q\bar{q}$ backgrounds. The size of the training dataset corresponds to 10M events for signal samples and 300 fb^{-1} for $q\bar{q}$ backgrounds. To avoid overfitting, the $q\bar{q}$ samples are maintained independently from those used in the cut-based selection optimization where differential evolution is applied. The samples are randomly under-sampled to equalize the

TABLE 8. Summary table of the pre-selection and fitting in reconstruction. The selections are applied in order from top to bottom.

target particle	selection and fitting
γ	$E > 40$ MeV
π^0	100 MeV $< M_{\pi^0} <$ 160 MeV
	mass-constraint kFit
	cosine of the helicity angle of the π^0 meson < 0.98
$\pi^{\pm}$	$\|dr\|$ <0.5 cm
	$\|dz\|$ <3 cm
	chiProb>0
$\rho^{\pm}$	600 MeV < M < 1100 MeV
B^0	M_{bc} >5.2 GeV
	$\|\Delta E\|$ <0.5 GeV
	flavor tagging
	IP-constraint treeFit
	vertex tagging with no constraint
	M_{bc} >5.24 GeV (again)
	$\|\Delta E\|$ <0.3 GeV (again)

number of signals and backgrounds in the training. The number of the training sample depends on the optimized selection, which is typically $\mathcal{O}(10^5)$. 95% of the data was allocated for the training, and the remaining 5% was allocated for the validation.

TABLE 9. Summary of variables applied for CSMVA.

Variable	Abbreviation
thrustOm	T_{tag}
thrustAxisCosTheta	cθT
cosTBz	cT s, b
cosTBTO	cT s, t
cosHelicityAngleMomentum	cHm
KSFWVariablesmm2	M^2_{miss}
KSFWVariableshoo0	$H_{\mathrm{oo}\,0}$
KSFWVariableshoo1	$H_{\mathrm{oo}\,1}$
KSFWVariableshoo2	$H_{\mathrm{oo}\,2}$
KSFWVariableshoo3	$H_{\mathrm{oo}\,3}$
KSFWVariableshoo4	$H_{\mathrm{oo}\,4}$
KSFWVariableshso00	$H_{\mathrm{so\ c}\,0}$
KSFWVariableshso02	$H_{\mathrm{so\ c}\,2}$
KSFWVariableshso04	$H_{\mathrm{so\ c}\,4}$
KSFWVariableshso10	$H_{\mathrm{so\ n}\,0}$
KSFWVariableshso12	$H_{\mathrm{so\ n}\,2}$
KSFWVariableshso14	$H_{\mathrm{so\ n}\,4}$
KSFWVariableshso20	$H_{\mathrm{so\,m0}}$
KSFWVariableshso22	$H_{\mathrm{so\,m2}}$
KSFWVariableshso24	$H_{\mathrm{so\,m4}}$
CMS_cosTheta	cθ^*

Figure 17 shows the output of the TabNet-based CSMVA model. Signals and backgrounds are classified well, and the performance of the training and validation dataset is consistent within the uncertainty, which means the model is not overfitted to the training samples.

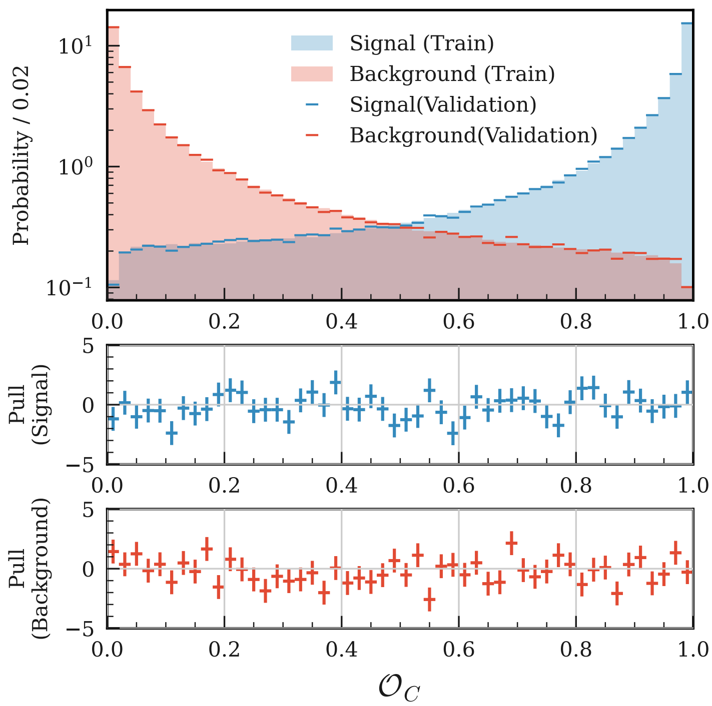


FIG. 17. Output of the TabNet-based CSMVA model using simulated samples.

3.5. Differential Evolution

To optimize the selection criteria, we use differential evolution (DE) [13], a genetic optimization algorithm. In this analysis, DE attempts to maximize the FoM while changing the event selection criteria and applying the best candidate selection in an iterative process. The P_{π^0} or E_γ strongly correlated to $\cos_{\rho^\pm}$, and too tight selection can affect the longitudinal polarization measurement. Thus, we set the limit of the selection for these variables in the selection optimization. All the selection variables and their upper or lower limits of the DE-search range are shown in Table 10.

TABLE 10. Selection that optimized by Differential Evolution

Category	Selection	Search Range
γ	$E_{forward}$	50 MeV - 300 MeV
	E_{barrel}	50 MeV - 300 MeV
	$E_{backward}$	50 MeV - 300 MeV
	PhotonMVA	0 - 1
π^0	P_{π^0}	0 MeV - 300 MeV
	M_{π^0}	100 MeV - 135 MeV (min.)
		135 MeV - 150 MeV (max.)
	daughterAngle	1 - π
	|daughterDiffOfPhi|	1 - π
$\pi^{\pm}$	PID(Binary)	0 - 1
Continuum suppression	CSMVA	0 - 1
B^0	M_{bc}	5.26 GeV - 5.28 GeV
fitting variables for signal-enhanced region	ΔE	-0.15 GeV - 0 GeV (min.)
		0 GeV - 0.15 GeV (max.)
	$M(\rho^{\pm})$	0.52 GeV - 0.775 GeV (min.),
		0.775 GeV - 1.2 GeV (max.)

3.6. Optimization

Figure 18 shows the results for each iteration of the optimization. We performed hyperparameter tuning for the MVA models prior to training them in each iteration. The FoM decreased in the third iteration because the hyperparameter tuning was not performed to check if the hyperparameter tuning was necessary to improve the FoM. Table 11 shows a summary of the optimized selections, and Figure 19 shows the distributions of the variables used in the selection.

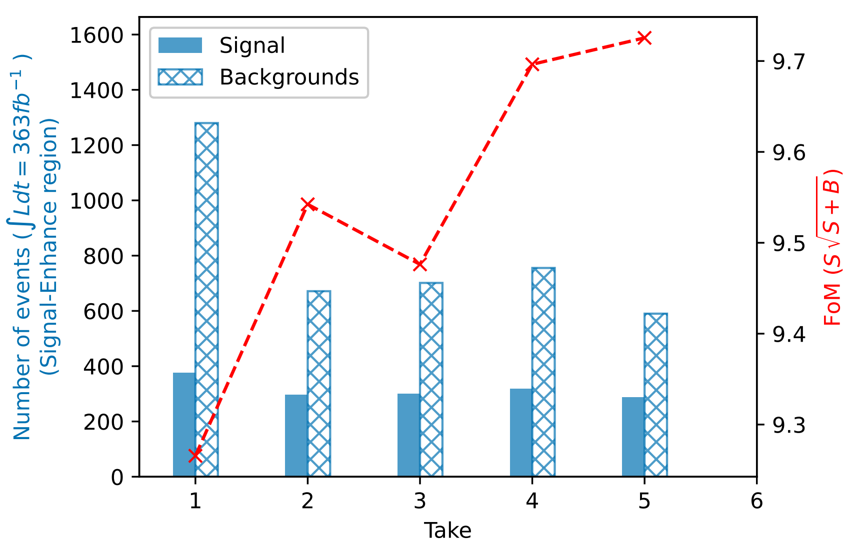


FIG. 18. Figure of merit in each iteration.

TABLE 11. Summary of the optimized selection.

Target	selection
γ	$E_{forward} > 90$ MeV $E_{barrel} > 60$ MeV $E_{backward} > 90$ MeV clusterNHits > 1.5 |clusterTiming| <200 ns PhotonMVA > 0.1
π^0	120 MeV$< M_{\pi^0} <$ 150 MeV $P_{\pi^0} > 210$ MeV daughterAngle<2.5 |daughterDiffOfPhi| < 1.3
$\pi^{\pm}$	PID(Binary)>0.1
Continuum suppression	CSMVA>0.96
B^0	$M_{bc} > 5.275$ GeV
Signal-Enhance	-0.08 $< \Delta E$[GeV] <0.04 0.6 $< m_{\pi^{\mp}\pi^0}$[GeV/c^2] <0.9

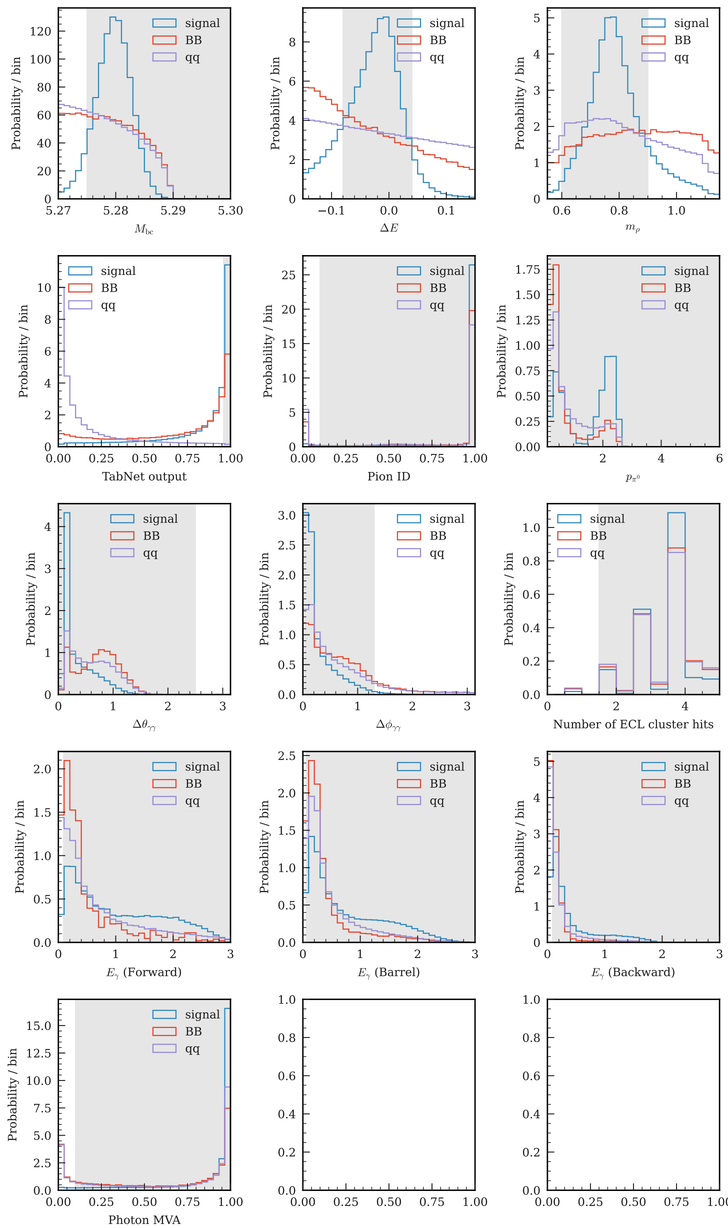


FIG. 19. Distributions of the variables used in the selection. The gray region corresponds to the selected region in the optimal selection.

3.7. Best candidate selection

We rank signal candidates in each event based on the sum of reduced chi-squared values given by,

$$\text{sum of reduced}\,\chi^2 = \frac{\chi^2_{\mathrm{B}^0}}{\mathrm{ndf}_{\mathrm{B}^0}} + \sum_{\pi^0_1,\,\pi^0_2} \frac{\chi^2_{\pi^0}}{\mathrm{ndf}_{\pi^0}}.$$

$\chi^2_{B^0}$ and $\chi^2_{\pi^0}$ are calculated in the vertex fit and π^0 mass-constraint fit, respectively. Figure 20 shows the number of candidates in longitudinal- and transverse-polarized signals after applying the optimized selection described in Section 3. On average, there are 1.189 candidates per event for longitudinal-polarized signal samples and 1.050 candidates per event for transverse-polarized signal samples. The efficiency of this best candidate selection is 94% for longitudinal-polarized signal samples and 98% for transverse-polarized signal samples.

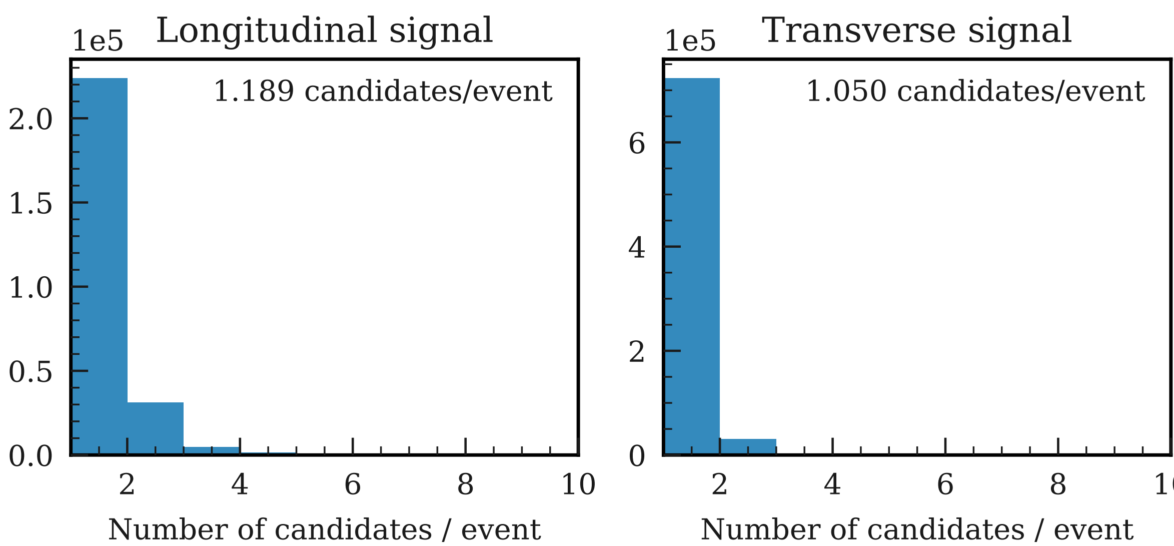


FIG. 20. (MC) Number of candidates in longitudinal- (left) and transverse-polarized (right) signal samples after applying the optimized selection

3.8. Summary of the selection optimization

Figure 21 and Table 12 shows the number of events at each selection step in run-independent simulated samples.

Selection	Signal	Self-crossfeed	qq	BB
Pre-selection	1770.1 (16.3%)	8679.4	10027239.8	187870.6
PID	1701.0 (15.7%)	8011.0	6692667.3	138702.3
γ	1477.2 (13.6%)	2690.9	2246153.2	40492.6
π^0	1307.5 (12.1%)	1786.1	1516142.1	24266.5
M_{bc}	1196.2 (11.0%)	866.9	321055.3	5608.3
CSMVA	504.4 (4.6%)	275.4	2486.5	1433.8

TABLE 12. Number of events at each cut in the run-independent simulated sample

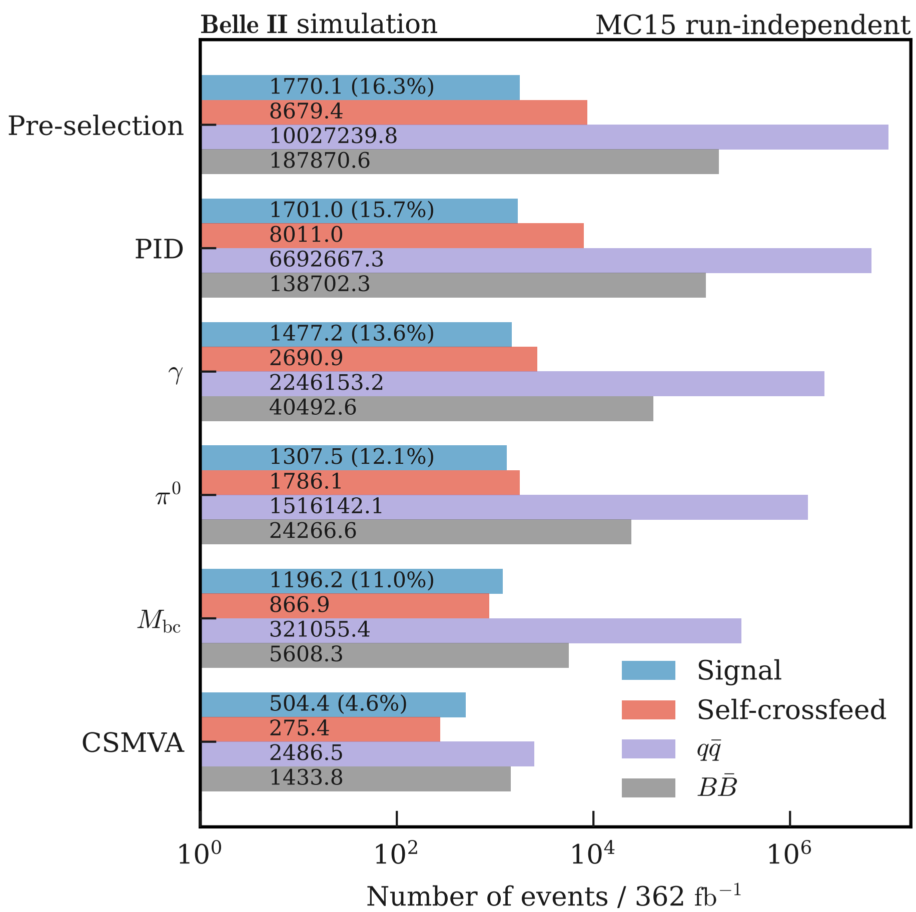


FIG. 21. Number of events at each cut in the run-independent simulated sample

Table 13 shows the number of events remaining after the optimized selection in the run-dependent simulated samples and run-dependent signal MC scaled to 363 fb^{-1}.

TABLE 13. Number of events after the selection with run-dependent simulated samples (363 fb^{-1})

Components	Number of events
Longitudinal signal	406.8
Self-crossfeed (Long. signal)	141.1
Transverse signal	7.4
Self-crossfeed (Trans. signal)	0.6
qq	2265.5
BB	906.2

4. SIGNAL EXTRACTION FITTING

4.1. Fitting strategy

In this analysis, we aim to measure $\mathcal{B}$, f_L, S_{CP}, and A_{CP}. The fitting is split into two steps: signal-extraction fitting and CP violation fitting. CP violation fitting needs flavor tagging information, which can be correlated with variables used in the signal-extraction fitting and potentially introduce bias. The two-step fitting process helps to avoid this bias. In the signal-extraction fitting, only $\mathcal{B}$ and f_L are measured without flavor tagging information. The fit parameters are ΔE, $m_{\pi^\pm\pi^0}$, $\cos\theta_\rho^\pm$, and CSMVA outputs, without flavor tagging output. After that, an event-by-event signal fraction is calculated from the results of the signal-extraction fitting using ΔE, $m_{\pi^\pm\pi^0}$, $\cos\theta_\rho^\pm$, and qr bins. The flavor tagger output qr is binned with the bin edge 0, 0.10, 0.25, 0.45, 0.60, 0.725, 0.875, and 1.0. The CP violation fitting is performed using this fraction. The fit parameters in the CP fitting are only S_{CP} and A_{CP}. In this section, we discuss signal-extraction fitting.

4.2. Signal extraction fitting

To measure the yield of each component, we fit ΔE, $m_{\pi^\pm\pi^0}$, $\cos\theta_{\rho^\pm}$, and CSMVA output. CSMVA output is transformed using probability integral transformation that makes the signal distributions flatten [14] to make PDF modeling simpler. We call the transformed CSMVA output as $\mathcal{T}_C$. Figure 22 shows the variables' distribution in the fit region.

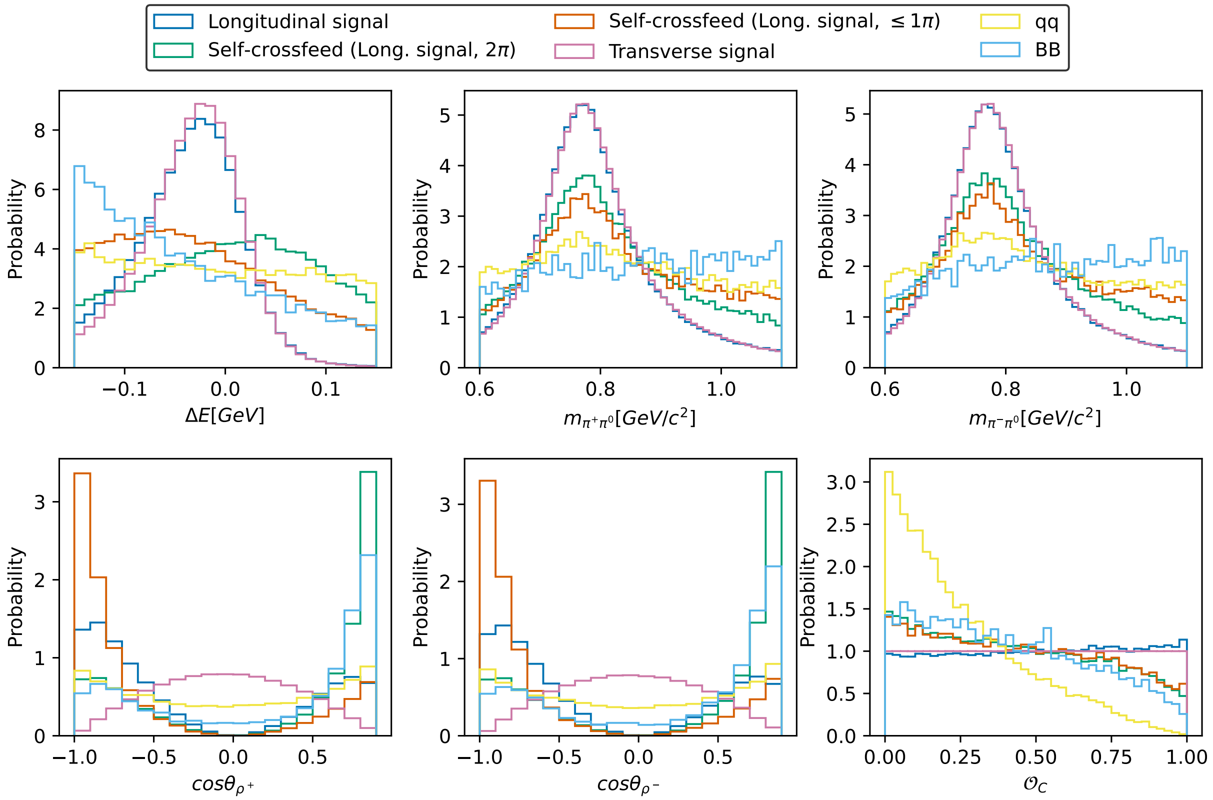


FIG. 22. Distribution of the fitting variables for $B^0 \to \rho^+\rho^-$ decays reconstructed in simulated samples after the best candidate selection.

4.3. Peaking backgrounds

Some B decays have a different distribution from other $B\bar{B}$ backgrounds that are treated as peaking backgrounds. Firstly, some B decays have the same final state particles as the signal mode, such as $B^0 \to \rho^\pm(\to \pi^\pm\pi^0)\pi^\mp\pi^0$. In this case, all particles from the peaking background decay are reconstructed as $B \to \rho^+\rho^-$, causing them to peak in the fitting region. In addition, $K-\pi$ misidentification can contribute to peaking backgrounds. In case where $K-\pi$ misidentification occurs in $K^\pm\pi^\mp\pi^0\pi^0$ final states, the resulting decay may resemble a signal. For example, $B^0 \to K(892)^{*\pm}(\to K^\pm\pi^0)\rho^\mp(\to \pi^\mp\pi^0)$ is a peaking background decay mode due to $K-\pi$ misidentification. Another cause of the peaking backgrounds is combinatorial effects. For example, $B^0 \to \pi^\pm\pi^\mp\pi^0$ is likely to be reconstructed as a signal with one π^0 from the tag side.

All decay modes included in the MC samples are checked, listed, and categorized to reduce bias arising from the peaking backgrounds. Table 14 and Table 15 show the peaking backgrounds, and Figure 23 and Figure 24 show the ΔE distributions of the peaking backgrounds. The peaking backgrounds are categorized into three types. Category A consists of decay modes that make a clear peak in ΔE distribution, with many remaining after the selection. As they can significantly affect the signal yield, they are individually modeled and fitted mode by mode. Category B consists of decay modes that do not make clear peaks in the ΔE distributions and do not remain too much but potentially have CP violation. These modes do not affect the signal yield too much, but could potentially bias the CPV parameters. They are also treated mode by mode in this analysis, and the systematic uncertainties due to their CPV parameters are considered. Category C consists of decay modes that don't remain much, don't make a peak in the signal region, and don't have CPV effects. In the prior studies, such as those conducted by Belle and BABAR, they were treated as peaking backgrounds. In contrast, thanks to the improved selection performance, their effects are much smaller in this analysis. Consequently, these decays are summed up and treated as a single "rare" peaking background category, rather than on a mode-by-mode basis.

Table 14 and Table 15 also numbers of remaining events and those uncertainties used in the fit. For modes with measured branching fractions, the yields are fixed. On the other hand, the yields are floated for the modes without measured branching fractions, and the modes without measured branching fractions in category C are summed up and floated as "rare peaking". Only $B^0 \to a_1^{\pm}\rho^{\mp}$ and $B^{\pm} \to a_1^{-}\rho^{\pm}$ are exceptions to this approach. Although the branching fraction of $B^0 \to a_1^{\pm}\rho^{\mp}$ has not been observed, an upper limit has been set to be $\mathcal{B}(B^0 \to a_1^{\pm}\rho^{\mp}) < 61 \times 10^{-6}$ by [15]. We assumed the branching fraction for this mode to be $(30 \pm 30) \times 10^{-6}$. The branching fraction for $B^{\pm} \to a_1^0\rho^{\pm}$ has not been measured. We approximate $\mathcal{B}(B^{\pm} \to a_1^0\rho^{\pm}) \simeq (f_\rho/f_\pi)^2\mathcal{B}(B^{\pm} \to a_1^0\pi^{\pm})$, where $f_{\pi,\rho} = m_{\pi,\rho}$. The systematic uncertainty is calculated to be within $\pm100\%$ of the estimated value.

TABLE 14. Peaking backgrounds from B^0. Modes with a branching fraction assigned a 100% uncertainty are unmeasured.

Decay mode	$\mathcal{B} \times 10^{-6}$	N_{bg}	category	Floated in the fitting
$B^0 \to \rho^+\pi^-\pi^0$	10 ± 10	79.2	ΔE peak	yes
$B^0 \to \pi^+\pi^-\pi^0\pi^0$	10 ± 10	11.5	ΔE peak	yes
$B^0 \to a_1(1260)^+\rho^-$	30 ± 30	3.8	ΔE peak	no
$B^0 \to K^\star(892)^+\rho^-$	10.3 ± 2.6	16.3	ΔE peak	no
$B^0 \to K_0^\star(1430)^+\rho^-$	28 ± 12	15.1	ΔE peak	no
$B^0 \to \pi^+\pi^-\pi^0$	30 ± 30	30.2	no peak, CPV	yes
$B^0 \to a_1(1260)^+\pi^-$	26 ± 5	14.3	no peak, CPV	no
$B^0 \to a_1(1260)^0\pi^0$	10 ± 10	3.5	no peak, CPV	no
$B^0 \to K_S^0\pi^0\pi^0$	23 ± 23	0.0	no peak, no CPV	yes
$B^0 \to K_2^\star(1430)^+\rho^-$	10 ± 10	0.0	no peak, no CPV	yes
$B^0 \to \rho^+\pi^-$	23.0 ± 2.3	0.0	no peak, no CPV	no
$B^0 \to \rho^0\pi^0\pi^0$	10 ± 10	2.3	no peak, no CPV	yes
$B^0 \to f_0(980)\pi^0\pi^0$	10 ± 10	1.3	no peak, no CPV	yes
$B^0 \to \omega\pi^0$	10 ± 10	0	no peak, no CPV	yes

TABLE 15. Peaking backgrounds from $B^{\pm}$. Modes with a branching fraction assigned a 100% uncertainty are unmeasured.

Decay mode	$\mathcal{B}\times 10^{-6}$	$N_{\rm bg}$	category	Floated in the fitting
$B^+ \to \rho^+\pi^0$	10.9 ± 1.4	84.7	ΔE peak	no
$B^+ \to a_1(1260)^0\pi^+$	20 ± 6	9.6	ΔE peak	no
$B^+ \to \rho^+\rho^0$	24 ± 1.9	53.8	no peak, CPV	no
$B^+ \to a_1(1260)^+\pi^0$	26 ± 7	37.7	no peak, CPV	no
$B^+ \to \rho^+ a_1(1260)^0$	50 ± 50	0.0	no peak, CPV	no
$B^+ \to K^+\pi^-\pi^+$	51 ± 2.9	3.5	no peak, no CPV	no
$B^+ \to K_0^{\star}(1430)^+\pi^0$	$11.9^{+2.0}_{-2.3}$	8.5	no peak, no CPV	no
$B^+ \to K_2^{\star}(1430)^0\rho^+$	10 ± 10	0.0	no peak, no CPV	yes
$B^+ \to \pi^+\pi^0\pi^0$	10 ± 10	7.3	no peak, no CPV	yes
$B^+ \to \rho^+\pi^+\pi^-$	10 ± 10	7.0	no peak, no CPV	yes
$B^+ \to \rho^-\pi^+\pi^+$	10 ± 10	1.0	no peak, no CPV	yes
$B^+ \to \rho^+\pi^0\pi^0$	10 ± 10	7.3	no peak, no CPV	yes

4.4. Fit model

We determine the signal branching fraction, the longitudinal polarization fraction, and the yields of the components with an extended maximum likelihood fit of the unbinned distributions of the fitting variables. Table 16 summarises all the fit models we apply, and the detailed descriptions for each component are in Appendix 2.

Biases can happen if the correlation of the variables in the six-dimensional plane is not considered correctly. Figure 25 shows the correlation in $\cos\theta_{\rho^{\pm}}$ and ΔE. To consider this, we sliced the simulated samples with ΔE and then made a PDF for $\cos\theta_{\rho^{\pm}}$ distributions for each slice. We evaluate two-dimensional correlations in each combination of the fit variables and fit components, and the correlation is considered in the same way only if there are large correlations.

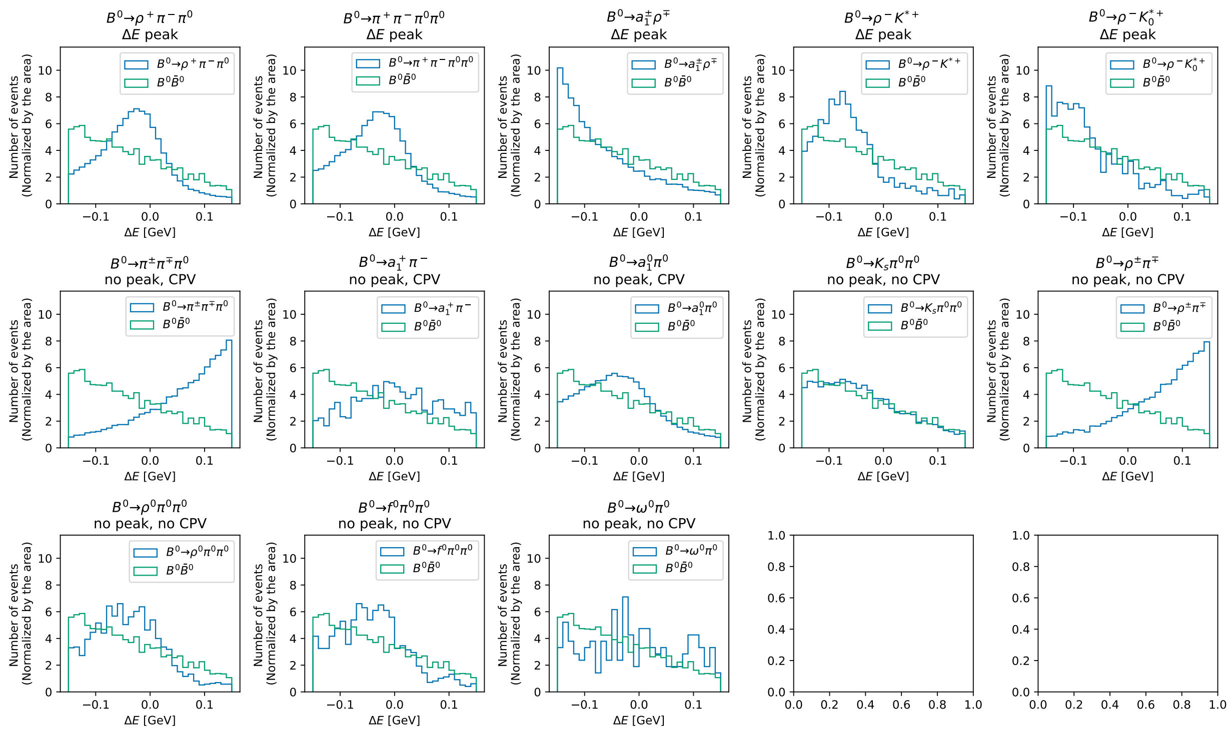


FIG. 23. ΔE distribution of B^0 peaking backgrounds in the simulated samples. The number of events is normalized to the area.

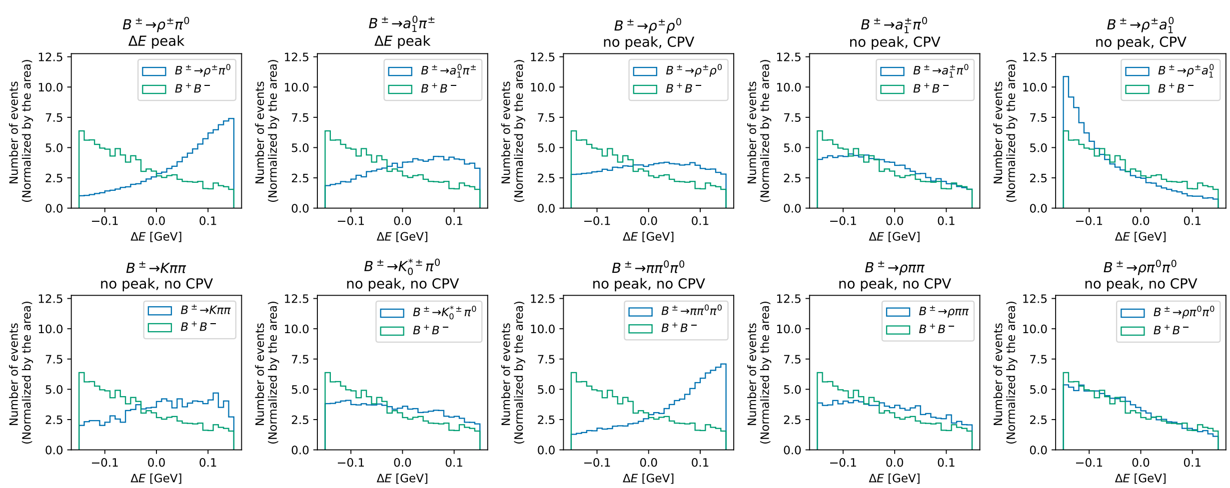


FIG. 24. ΔE distribution of $B^\pm$ peaking backgrounds in the simulated samples. The number of events is normalized to the area.

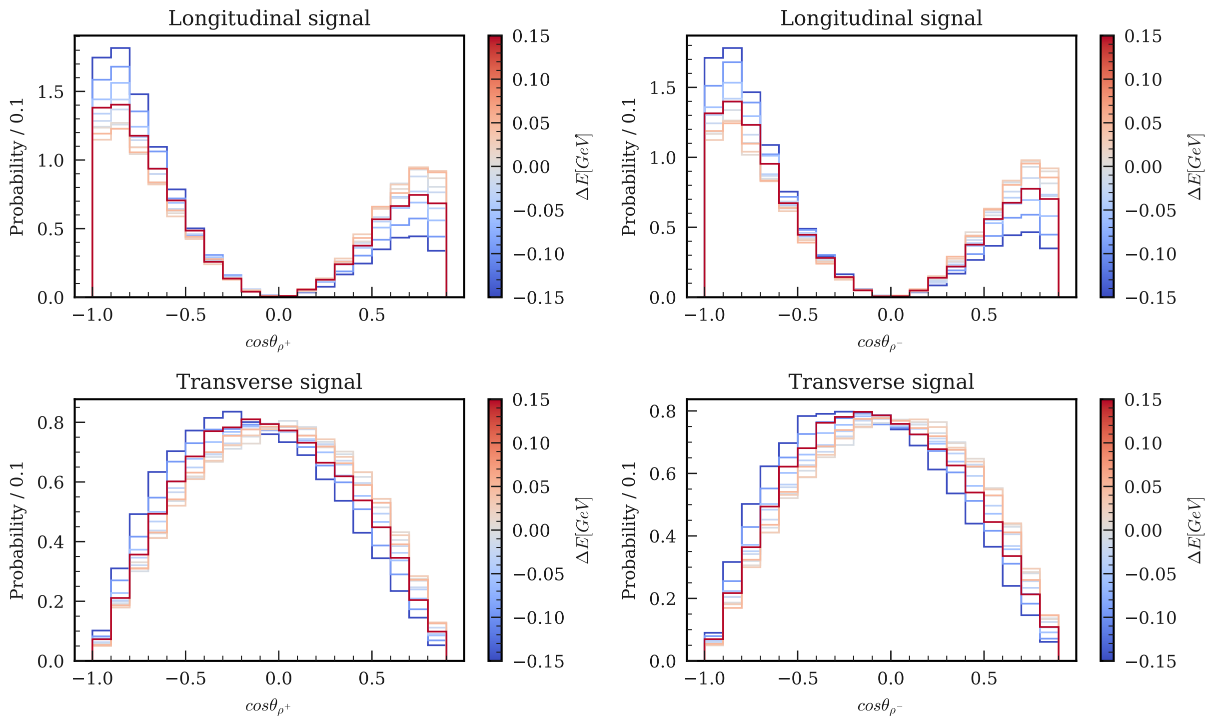


FIG. 25. Correlation between ΔE and $cos\theta_{\rho^{\pm}}$ for longitudinal signal transverse signals

TABLE 16. Summary of fit models for each event type. bG: bifurcated-Gaussian, BW: Breit-Wigner formula, CX: X-th order Chebyshev polynomial function, L: Linear function, C: Constant, and d: double. $|_X$ implies that the correlation with X is considered. Yields of the peaking backgrounds whose $\mathcal{B}$ is already measured are fixed, and the yield of the others is floated. For the self-crossfeed events, the ratio to the signal is fixed, but the total yield is floated.

	ΔE	$m(\pi^{\pm}\pi^0)$	CSMVA	$\cos\theta_{\rho^{\pm}}$	Constraint on the yield
Longitudinal signal	dbG	BW	C1	Template$\|_{\Delta E}$	Free
SxF (Long. signal, 2π)	bG	BW+ C1	C1	Template$\|_{cos\theta_{\rho^{\mp}}}$	Fixed fraction
SxF (Long. signal, not 2π)	bG	BW+ C1	C1	Template$\|_{cos\theta_{\rho^{\mp}}}$	Fixed fraction
Transverse signal	dbG	BW	C1	Template$\|_{\Delta E}$	Free
SxF (Trans. signal)	bG	BW+ C1	C1	Template$\|_{cos\theta_{\rho^{\mp}}}$	Fixed fraction
$q\bar{q}$	C2	BW+ C1	Template$\|_{\Delta E}$	Template$\|_{m_{\pi^{\pm}\pi^0}}$	Free
$B\bar{B}$	C2	C2	Template$\|_{\Delta E}$	Template$\|_{m_{\pi^{\pm}\pi^0}}$	Free
$\tau\tau$	C2	BW+ C1	exp(ax+b)+c	Template	Fixed to MC expectation
$B^0 \to \rho^+\pi^-\pi^0$	dbG	Template$\|_{m_{\pi^{\pm}\pi^0}}$	Template$\|_{\Delta E}$	Template$\|_{cos\theta_{\rho^{\mp}}}$	Free
$B^0 \to \pi^+\pi^-\pi^0\pi^0$	dbG	C1	Template$\|_{\Delta E}$	Template$\|_{\Delta E}$	Free
$B^0 \to a_1^{\pm}\rho^{\mp}$	exp(ax)+bx+c	BW+ C1	Template$\|_{\Delta E}$	Template	External input
$B^0 \to \rho^- K^{*+}$	dbG	Template$\|_{m_{\pi^{\pm}\pi^0}}$	C1	Template	External input
$B^0 \to \rho^- K_0^{*+}$	C2	Template$\|_{m_{\pi^{\pm}\pi^0}}$	C1	Template	External input
$B^0 \to \pi^{\pm}\pi^{\mp}\pi^0$	C2	BW+ C1	Template$\|_{\Delta E}$	Template	Free
$B^0 \to a_1^+\pi^-$	dbG	Template$\|_{m_{\pi^{\pm}\pi^0}}$	C1	Template	External input
$B^0 \to a_1^0\pi^0$	dbG	Template$\|_{m_{\pi^{\pm}\pi^0}}$	Template$\|_{\Delta E}$	Template$\|_{cos\theta_{\rho^{\mp}}}$	Free
$B^0 \to \rho^{\pm}\pi^{\mp}$	C2	BW+ C1	C1	Template$\|_{m_{\pi^{\pm}\pi^0}}$	External input
$B^{\pm} \to \rho^{\pm}\pi^0$	C2	Template$\|_{m_{\pi^{\pm}\pi^0}}$	C1	Template$\|_{m_{\pi^{\pm}\pi^0}}$	External input
$B^{\pm} \to a_1^0\pi^{\pm}$	dbG	BW+ C1	C1	Template$\|_{m_{\pi^{\pm}\pi^0}}$	External input
$B^{\pm} \to \rho^{\pm}\rho^0$	dbG	BW+ C1	C1	Template$\|_{m_{\pi^{\pm}\pi^0}}$	External input
$B^{\pm} \to a_1^{\pm}\pi^0$	dbG	BW+ C1	C1	Template	External input
$B^{\pm} \to \rho^{\pm}a_1^0$	exp(ax)+bx+c	BW+ C1	Template$\|_{\Delta E}$	Template	External input
$B^{\pm} \to K\pi\pi$	C2	BW+ C1	C1	Template	External input
$B^{\pm} \to K_0^{*\pm}\pi^0$	bG	C1	C1	Template$\|_{m_{\pi^{\pm}\pi^0}}$	External input
Rare peaking backgrounds	dbG	BW+ C1	C1	Template$\|_{cos\theta_{\rho^{\mp}}}$	Free

4.5. Likelihood

The likelihood we use is given by

$$
\begin{aligned}
\mathcal{L}(\Delta E,\ m_{\pi^\pm\pi^0},\ \cos\theta_{\rho^\pm},\ \mathcal{O}_C|\mathcal{B}, f_L, N_{B\bar{B}}, N_{q\bar{q}}, N_{\rho^+\pi^-\pi^0}, N_{\pi^+\pi^-\pi^0\pi^0}, N_{a_1^0\pi^0}, N_{\text{rarepeaking}})& \\
\equiv \frac{\prod_j e^{-N_j}}{N_{tot}!}\prod_{i=1}^{N_{tot}}[\mathcal{B}f_L\epsilon_L n_{B\bar{B}}P_L^i + \mathcal{B}(1-f_L)\epsilon_T n_{B\bar{B}}P_T^i +& \\
\mathcal{B}n_{B\bar{B}}f_L\epsilon_L k_{scfL2\pi}P^i_{scfL2\pi} + \mathcal{B}n_{B\bar{B}}f_L\epsilon_L k_{scfL1\pi}P^i_{scfL1\pi}& \\
+\mathcal{B}n_{B\bar{B}}(1-f_L)\epsilon_T k_{scfT}P^i_{scfT}+& \\
N_{q\bar{q}}P^i_{q\bar{q}} + N_{B\bar{B}}P^i_{B\bar{B}} + N_{\tau\bar{\tau}}P^i_{\tau\bar{\tau}}+& \\
\sum_{l=peak} N_l^i P_l^i& \\
],&
\end{aligned}
$$

where

$\mathcal{B}$: Branching fraction,

f_L: Longitudinal polarization fraction,

n_{BB}: Number of the B^0 and $\bar{B}^0$,

N_j: Number of events in category j,

ϵ_j: Total efficiency

k_j: Signal self-crossfeed ratio,

P_j^i: j's probability distribution function, and

subscript j

- L: longitudinal-polarized signal
- T: transverse-polarized signal
- $scfL2\pi$: self crossfeed longitudinal with two correctly reconstructed $\pi^\pm$
- $scfL1\pi$: self crossfeed longitudinal with one or less correctly reconstructed $\pi^\pm$
- $scfT$: self crossfeed transverse
- $B\bar{B}$: $B\bar{B}$ background
- $q\bar{q}$: $q\bar{q}$ background
- $\tau\bar{\tau}$: $\tau\tau$
- $peak$: peaking backgrounds.

In fitting, we fix k_i from the simulated samples, the signal efficiencies are shown in Table 17 and the ratio is shown in Table 18. $\mathcal{B}$, f_L, and N_is are floating parameters. The free parameters in the fits are $\mathcal{B}$, f_L, number of BB, qq, $B^0 \to \pi^+\pi^-\pi^0\pi^0$, $B^0 \to \rho\pi\pi^0$, $B^0 \to \pi^+\pi^-\pi^0$, $B^0 \to a_1^0\pi^0$, and rare peaking backgrounds. The shape parameter of ΔE distribution for BB backgrounds is also free to float.

Components	Efficiency
Longitudinal Signal	3.93%
Transverse Signal	7.08%

TABLE 17. Signal efficiency after the optimized selection

Components	Ratio
SxF (Long. 2π)	0.19
SxF (Long. $< 1\pi$)	0.16
SxF (Trans.)	0.08

TABLE 18. Ratio of the self-crossfeed signals to the correctly reconstructed signal

4.6. Estimator properties

4.6.1. Result of the fitting to MC samples

We sampled the simulated samples randomly with replacement and conducted fitting. The statistics of the simulated samples correspond to 363 fb^{-1}. Figure 26 shows the fit results, and Figure 27 shows the signal-enhanced plot derived from the fitting. Table 19 shows the selection criteria for each variable to enhance signal visibility. The dataset and PDF in the fit are well-matched. Table 20 lists the fitting results. All the fitted values are consistent with the input values.

Variable	Selection
ΔE[GeV]	0.6 $< m_{\pi^\pm\pi^0}$[GeV/c^2] $<$0.9
	$\mathcal{O}_C$ $>$0.96
$m_{\pi^\pm\pi^0}$[GeV/c^2]	-0.08 $< \Delta E$[GeV] $<$0.04
	0.6 $< m_{\pi^\mp\pi^0}$[GeV/c^2] $<$0.9
	$\mathcal{O}_C$ $>$0.96
$\mathcal{O}_C$	-0.08 $< \Delta E$[GeV] $<$0.04
	0.6 $< m_{\pi^\mp\pi^0}$[GeV/c^2] $<$0.9
$cos\theta_{\rho^\pm}$	-0.08 $< \Delta E$[GeV] $<$0.04
	0.6 $< m_{\pi^\mp\pi^0}$[GeV/c^2] $<$0.9
	$\mathcal{O}_C$ $>$0.96

TABLE 19. Signal-Enhance selection for the plot

TABLE 20. Result of fitting to MC samples

Parameter	Fit result	Input
$\mathcal{B}(\times 10^{-6})$	27.1 $\pm$ 2.5	27.7
f_L	0.989 $\pm$ 0.033	0.99
$N_{\mathrm{q\bar{q}}}$	2318.7 $\pm$ 83.1	2265.5
$N_{\mathrm{B\bar{B}}}$	819.6 $\pm$ 76.1	906.2
$N_{\mathrm{B^0\to\rho^+\pi^-\pi^0}}$	153.0 $\pm$ 88.6	23.9
$N_{\mathrm{B^0\to\pi^+\pi^-\pi^0\pi^0}}$	-116.5 $\pm$ 53.5	7.9
$N_{\mathrm{B^0\to\pi^\pm\pi^\mp\pi^0}}$	23.3 $\pm$ 23.7	6.6
$N_{\mathrm{B^0\to a_1^0\pi^0}}$	56.2 $\pm$ 32.1	29.4
N_{Rare}	23.6 $\pm$ 50.0	30.2

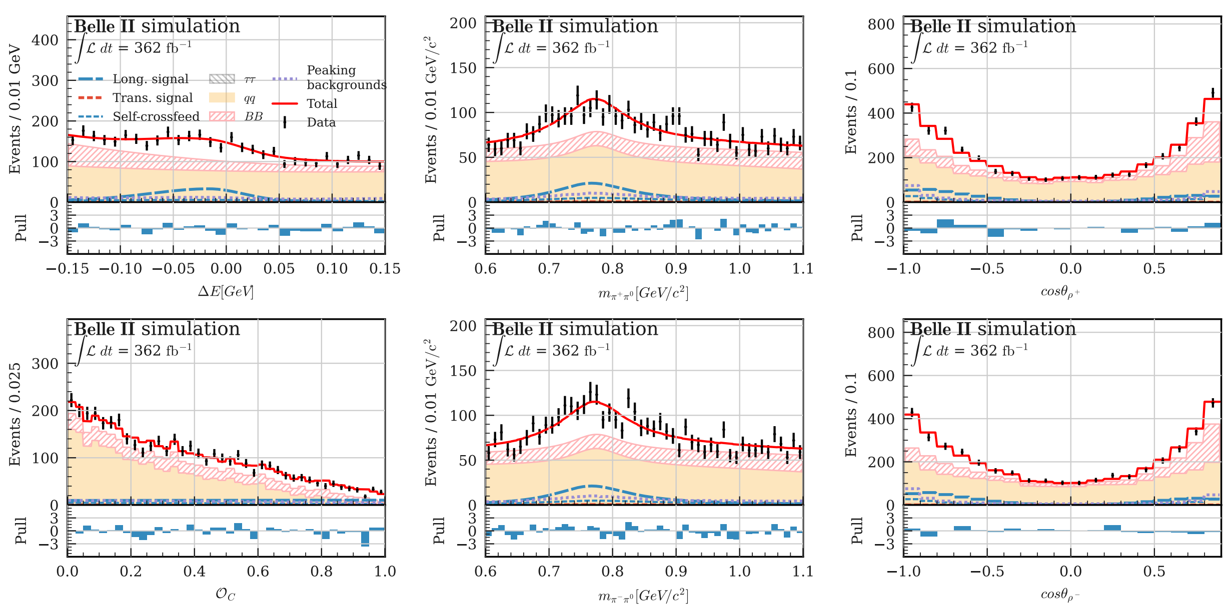


FIG. 26. Result of the fitting to 363 fb^{-1} of MC samples

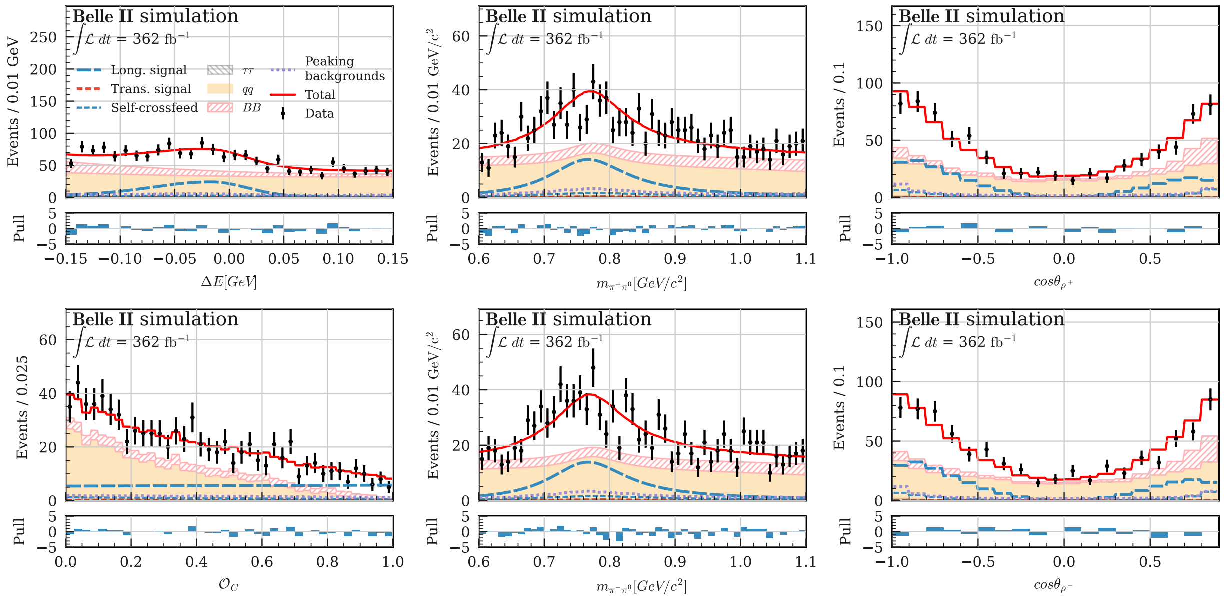


FIG. 27. Signal-Enhance plot of the fitting to 363 fb^{-1} of MC samples

4.7. Fitter validation

4.7.1. Bootstrap study

To validate the biases of the fitting procedure, we repeat the fit to samples that are randomly sampled from the simulated samples many times. That is called a "bootstrap" study. Table 21 shows the size of the MC samples used in the bootstrap study. All samples are generated considering the run condition differences in the data. The sample size for each iteration corresponds to 365 fb^{-1} which is the same size as the data. We performed 500 iterations to estimate the bias.

Table 22 and Figure 28 show the results of the bootstrap study, and Figure 29 shows the correlation among the fit parameters. While some of the fitted results show slight biases, the biases of the physics parameters are smaller than statistical uncertainty, which is acceptable.

Sample	Size of the total samples
Signal (Long.)	4033842 events (141 ab^{-1})
Signal (Trans.)	9904845 events (34000 ab^{-1})
$q\bar{q}$,$B\bar{B}$	1443.999 fb^{-1}
Peaking backgrounds(Fixed)	495261 events (17-50 ab^{-1} depending on the mode)
Peaking backgrounds(Floated)	4033842 events (70-500 ab^{-1} depending on the mode)

TABLE 21. Size of the MC samples used in the bootstrap study

Parameter	Mean	Input	Mean (Pull)	Std. (Pull)
$\mathcal{B}(\times 10^{-6})$	27.4 ± 0.1	27.7	-0.108 ± 0.047	1.033 ± 0.033
f_L	1.002 ± 0.001	0.990	0.306 ± 0.047	1.042 ± 0.034
$N_{q\bar{q}}$	2275.4 ± 2.6	2265.5	0.116 ± 0.045	1.003 ± 0.032
$N_{B\bar{B}}$	853.9 ± 2.5	906.2	-0.697 ± 0.046	1.014 ± 0.033
$N_{B^0\to\rho^+\pi^-\pi^0}$	56.1 ± 3.4	23.9	0.336 ± 0.055	1.210 ± 0.039
$N_{B^0\to\pi^+\pi^-\pi^0\pi^0}$	-7.6 ± 2.3	7.9	-0.322 ± 0.057	1.253 ± 0.041
$N_{B^0\to\pi^\pm\pi^\mp\pi^0}$	19.1 ± 0.8	6.6	0.472 ± 0.046	1.022 ± 0.033
$N_{B^0\to a_1^0\pi^0}$	38.6 ± 1.1	29.4	0.249 ± 0.046	1.025 ± 0.033
N_{rare}	42.6 ± 1.7	30.2	0.195 ± 0.048	1.055 ± 0.034

TABLE 22. Result of the bootstrap study. Uncertainties come from statistical uncertainty of the number of generated bootstrap samples.

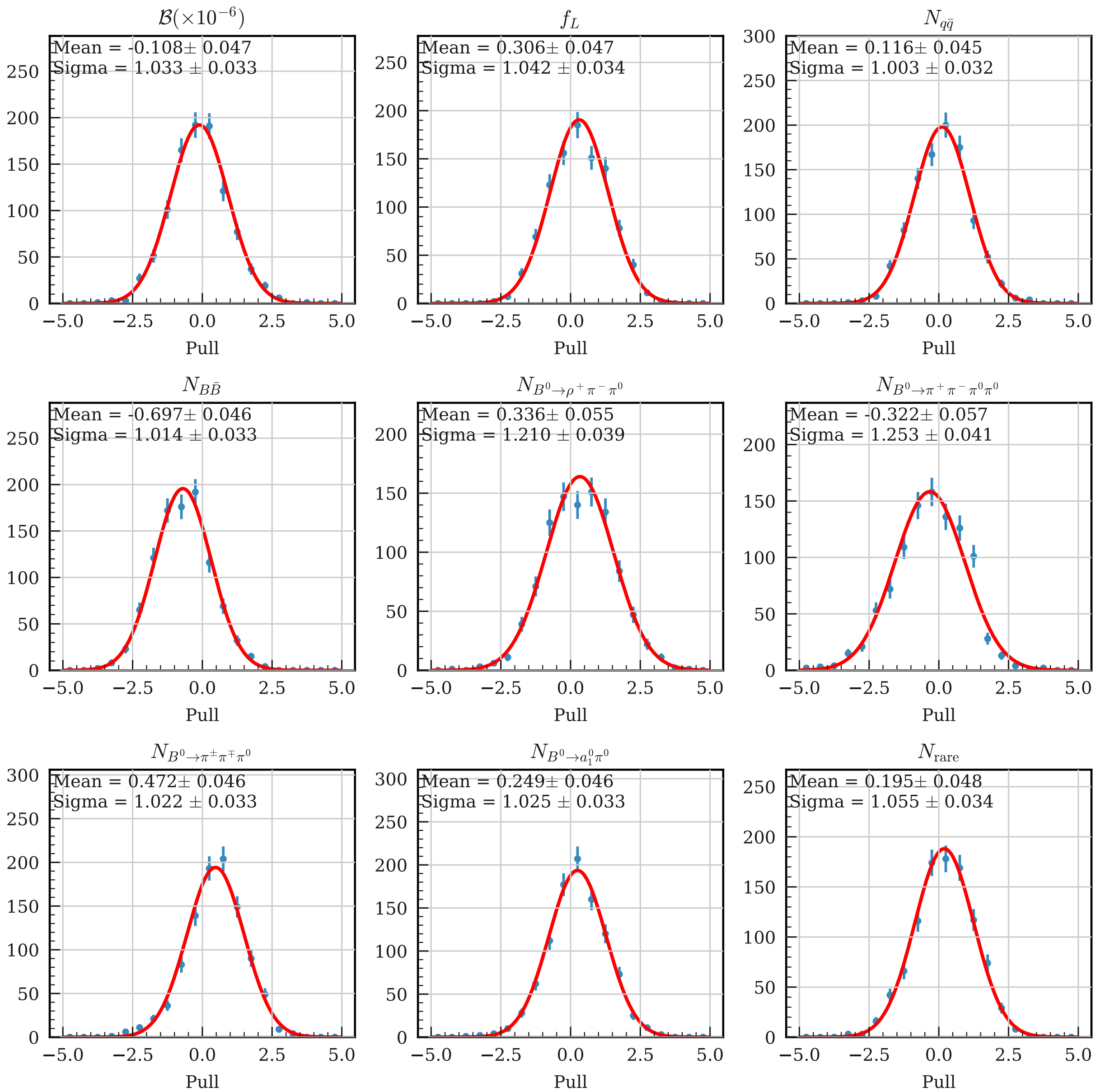


FIG. 28. Pull of the bootstrap study

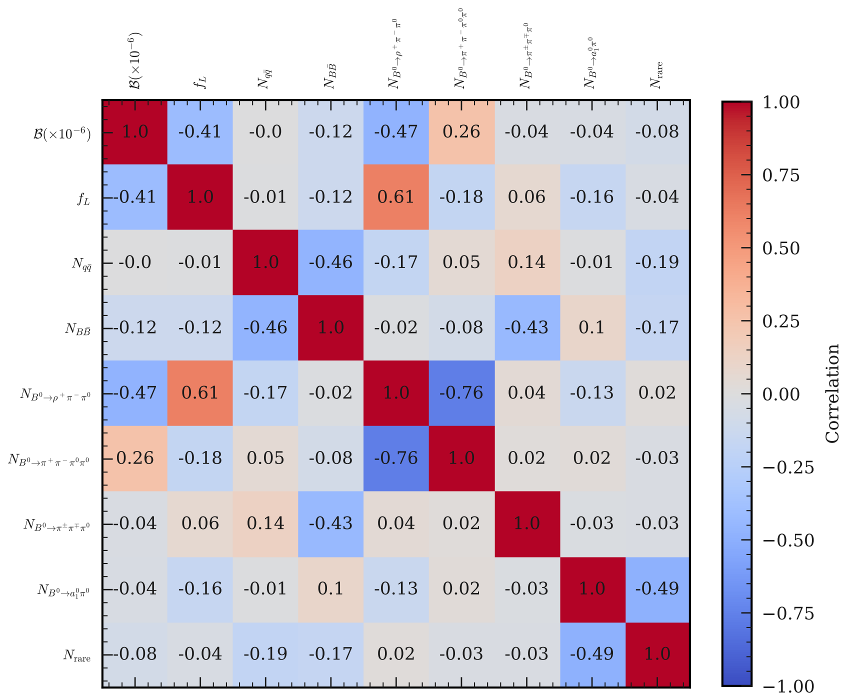


FIG. 29. Correlation of the fit parameters

4.7.2. Linearity test

A linearity test to physics observables was performed to understand the biases due to the physics parameters. In this test, the sample was the same as for the bootstrap study, and the number of signal events was changed depending on $\mathcal{B}$ and f_L. Figure 30 shows the results of the linearity test. Although the bootstrap study indicates that the fitted results are a little biased, it was observed that the sizes of these biases do not depend on the $\mathcal{B}$ and f_L so much. The observed biases are included in the systematic uncertainty and are not major sources of it, as described in the dedicated section on the systematic uncertainties.

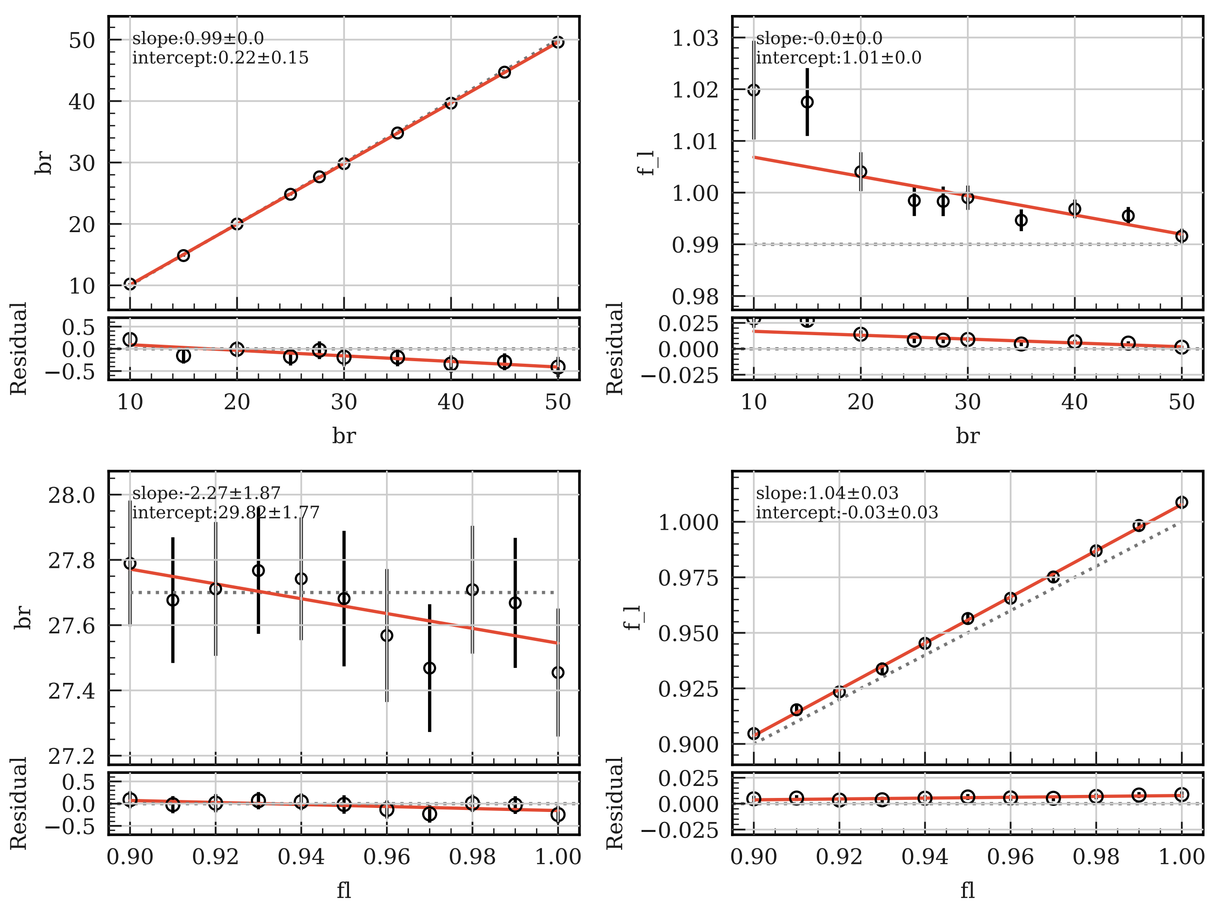


FIG. 30. Result of the linearity test for $\mathcal{B}$ and f_L using the simulated samples.

5. CP FITTING

5.1. Δt resolution of the signal

The measured Δt is affected by the detector response. To take this into account, the Δt resolution needs to be understood and should be considered in CP fitting. Firstly, we model the Δt resolution using the simulated samples. We tested $\Delta t - \Delta t_{\rm MC}$ distribution where $\Delta t_{\rm MC}$ is a decay-time difference of $B^0_{\rm sig}$ and $B^0_{\rm tag}$ without detector resolution. The Δt resolution function used in this analysis is similar to that in Ref. [16] and is given as a function of $\delta\Delta t \equiv \Delta t - \Delta t^{\rm MC}$ with a parameter σ, as follows.

$$
\begin{aligned}
\mathcal{R}(\delta\Delta t;\sigma) &= (1-f_{\rm OL})\,\mathcal{R}_{\rm core}\,(\delta\Delta t;\sigma) + f_{\rm OL}\mathcal{R}_{\rm OL}(\delta\Delta t;\sigma)\\
\\
\mathcal{R}_{\rm core}\,(\delta\Delta t;\sigma) &= (1-f_{\rm tail}\,)\cdot G\,(\delta\Delta t;\mu_{\rm main}\,\cdot\sigma, s_{\rm main}\,\cdot\sigma)\\
&\quad + f_{\rm tail}((1-f_{\rm exp})G\,(\delta\Delta t;\mu_{\rm tail}\,\cdot\sigma, s_{\rm tail}\,\cdot\sigma)\\
&\quad + f_{\rm exp}\,\cdot G\,(\delta\Delta t;\mu_{\rm tail}\,\cdot\sigma, s_{\rm tail}\,\cdot\sigma)\\
&\quad \otimes \big((1-f_{\rm R})\exp_{-}(\delta\Delta t/c\cdot\sigma) + f_{\rm R}\exp_{+}(-\delta\Delta t/c\cdot\sigma)\big)\\
f_{\rm tail}\,(\sigma) &= f_{\rm max}\cdot {\rm erf}\,(\sigma,\mu=f_{\mu},\sigma=f_{\sigma})\\
\\
\mathcal{R}_{\rm OL}(\delta\Delta t;\sigma) &= G\,(\delta\Delta t, 0, \sigma_{\rm OL})
\end{aligned}
\tag{27}
$$

where

- $\sigma \equiv \sigma_{\Delta t}$, i.e., the uncertainty on Δt from the vertex fitting
- $G(x,\mu,\sigma)$ is the Gaussian distribution of mean μ and width σ
- $\exp_{+}(x) = \exp(x)$ where $x > 0$, and $\exp_{+}(x) = 0$ otherwise. Similarly for $\exp_{-}(x)$
- $\mu_{\rm main}$, $s_{\rm main}$ are the mean and width (in units of $\sigma_{\Delta t}$) of the main Gaussian
- $\mu_{\rm tail}$, $s_{\rm tail}$ are the mean and width of the tail Gaussian
- c is an exponential constant describing the length of the exponential tails
- $f_{\rm tail}\,(\sigma)$ is the fraction of events belonging to the tail part of the distribution (tail Gaussian or exponential tails), as opposed to events belonging to the main Gaussian. $f_{\rm tail}$ is itself a function of $\sigma_{\Delta t}$ as explained in more detail below
- $f_{\rm exp}$ is the fraction of events belonging to the exponential part of the resolution function (as opposed to the purely Gaussian part).
- $f_{\rm R}$ is the fraction of events in the exponential tails that belong in the right tail (as opposed to the left tail).
(This equation and text above are cited from Ref. [17].)
- $f_{\rm OL}$ is the fraction of the outlier part of the resolution.

- $\mathcal{R}_{\rm OL}$: The outlier part of the resolution. The mean is fixed to 0, and the $\sigma_{\rm OL}$ is fixed to 200 ps.

This resolution function is a standard and widely used in the Belle II analysis. However, some differences were found in the $B^0 \to \rho^+\rho^-$ samples, and the fit failed with this resolution function. This is because of the difference in the decay and the analysis software. The cause of the issue mainly comes from the modeling for the fraction of the tail. Figure 31 shows the difference in the f_{tail} modelling between $B^0 \to \rho^+\rho^-$ and $B^0 \to D^{*-}\pi^-$. In this validation, all resolution function parameters were fixed except for f_{tail}, and the fitting was performed to the samples in each σ region. f_{tail} was modeled using the error function in the nominal modeling, but both $B \to \rho^+\rho^-$ and $B^0 \to D^{*-}\pi^-$ show the discrepancy from the function when σ is large. In addition, in the $B \to \rho^+\rho^-$ analysis, the vertex resolution is worse than $B^0 \to D^{*-}\pi^-$ in average because the vertex position is determined using only two charged $\pi^{\pm}$. It means that the effects in the large σ region become significant, leading to the failure of the fitting. It may not be a major issue with the current statistics, but it could potentially become an issue with larger statistics in the future analysis. So, we modified the modeling.

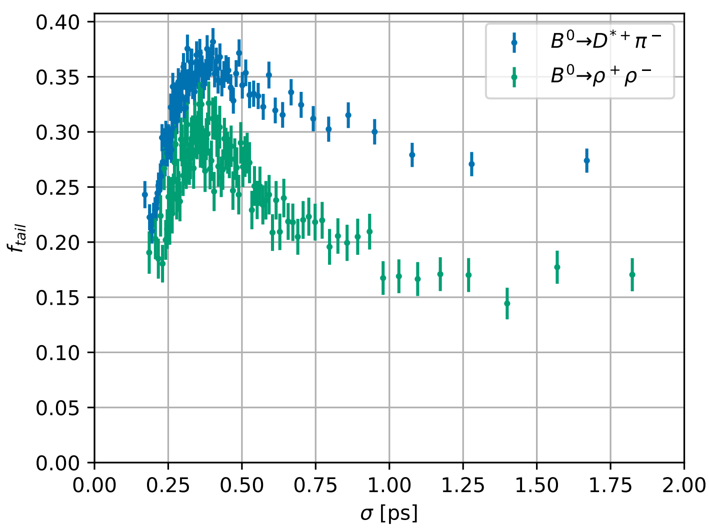


FIG. 31. The difference in the f_{tail} modeling between $B \to \rho^+\rho^-$ and $B^0 \to D^{*-}\pi^-$

To improve the modeling for the large σ region, the f_{tail} modeling is updated using a Bifurcated Gaussian function and Offset. Figure 32 shows the f_{tail} fitting for $B \to \rho^+\rho^-$ samples. Equation 28 shows the updated f_{tail} modeling.

$$f_{\rm tail}\,(\sigma) = f_{max} b(\sigma, \mu = f_{\mu}, \sigma_R = f_{\sigma R}, \sigma_L = f_{\sigma L}) + C_{f_{\rm tail}} \tag{28}$$

where

- f_{tail}: Fraction of the tail component.
- f_{max}: Constant parameter to scale the bifurcated Gaussian term.
- b: Bifurcated Gaussian function.

- f_μ: Mean of the Gaussian. It is shared in both the right and left parts.
- $f_{\sigma R}, f_{\sigma L}$: σ of the Gaussian for the right and left parts, respectively.
- $C_{f_{\rm tail}}$: Constant.

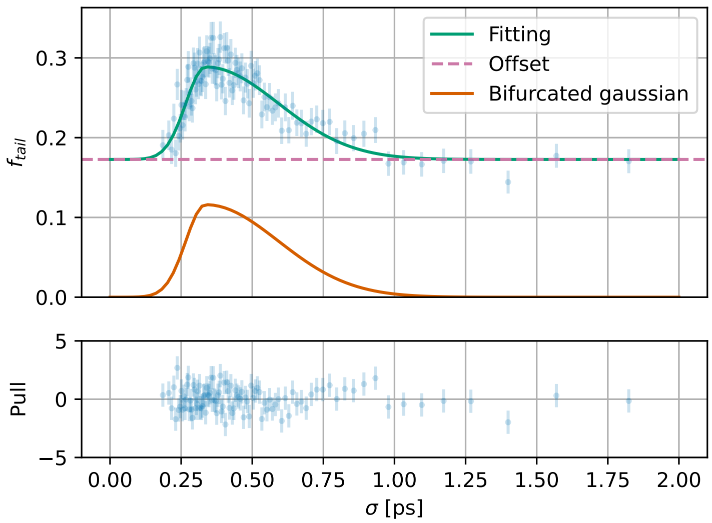


FIG. 32. Updated f_{tail} modeling for $B \to \rho^+\rho^-$

Resolution function fitting was performed to the simulated signal samples using the updated resolution function model for $B \to \rho^+\rho^-$. Figure 33 shows the fitted resolution function, and the model describes the sample well. Table 23 lists the parameters obtained from the fitting. The performance of the resolution function is also checked in each σ region and qr bin, and Figure 34 and Figure 35 show the results. The resolution function works well in all σ and qr bin regions. To validate the Δt resolution function shape in the data, we used $B^0 \to D^*\pi$. It is difficult to determine all the resolution function parameters simultaneously, and some of the parameters are fixed in the validation. The systematic uncertainty of the resolution function is also assigned for the fixed parameters in the fitting. It is discussed in Section 8. We also performed studies about the bias in CP parameters due to the difference in the resolution shape between the signal and control mode, and due to the method to validate the Δt resolution using control mode. It is shown in the Appendix 7

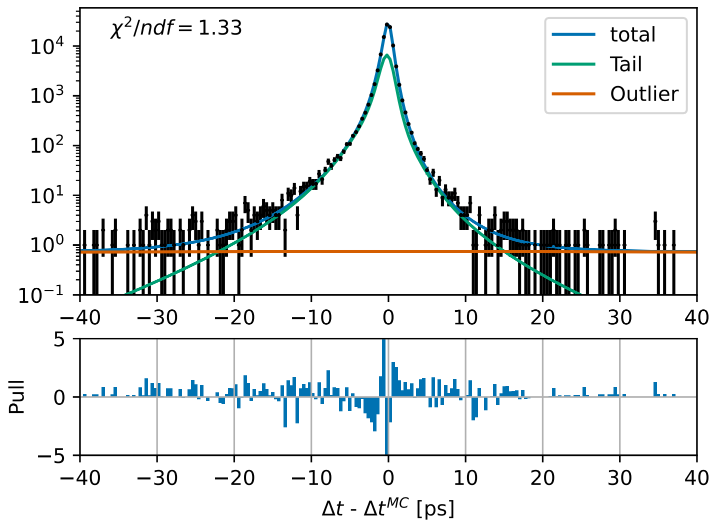


FIG. 33. Fit result of the resolution function to $B^0 \to \rho^+\rho^-$ signalMC samples. The number of signal MC samples is more than 200 times the expected signal yield with 363fb^{-1}

.

Parameter	Fit result	Floating in data
f_{OL}	0.0015 ± 0.0002	No
c	4.72 ± 0.19	No
f_{R}	0.23 ± 0.01	No
f_{exp}	0.19 ± 0.01	No
μ_{main}	-0.14 ± 0.01	Yes
μ_{tail}	-0.68 ± 0.06	Yes
s_{main}	1.15 ± 0.02	Yes
s_{tail}	2.0 ± 0.07	Yes
σ_{OL}	200.0 ± 2.0	No
c_6	4.94 ± 0.39	No
f_{R6}	0.26 ± 0.03	No
f_{exp6}	0.11 ± 0.01	No
μ_{main6}	-0.08 ± 0.02	Yes
μ_{tail6}	-0.34 ± 0.05	Yes
f_{μ}	0.34 ± 0.02	No
$f_{\sigma R}$	0.24 ± 0.03	No
$f_{\sigma L}$	0.12 ± 0.03	No
f_{max}	0.12 ± 0.01	No
$C_{f_{\mathrm{tail}}}$	0.26 ± 0.03	No
N	100002.3 ± 316.24	Yes

TABLE 23. Fit result of the resolution function to $B^0 \to \rho^+\rho^-$ signalMC samples.

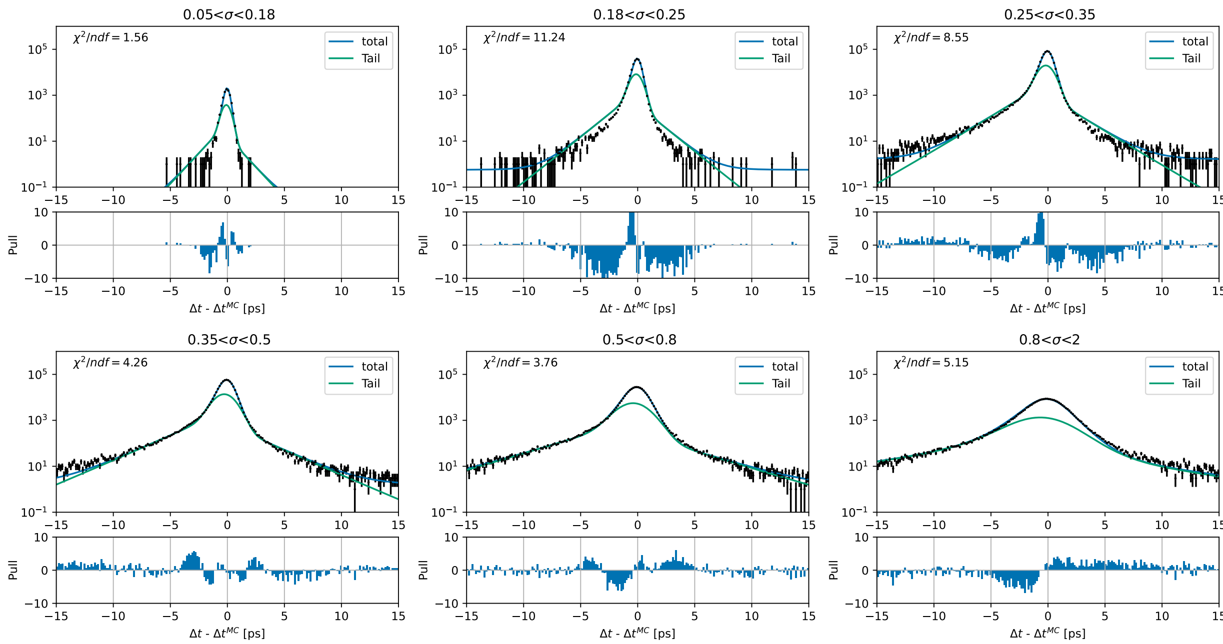


FIG. 34. σ dependence of the resolution function fit to signalMC samples

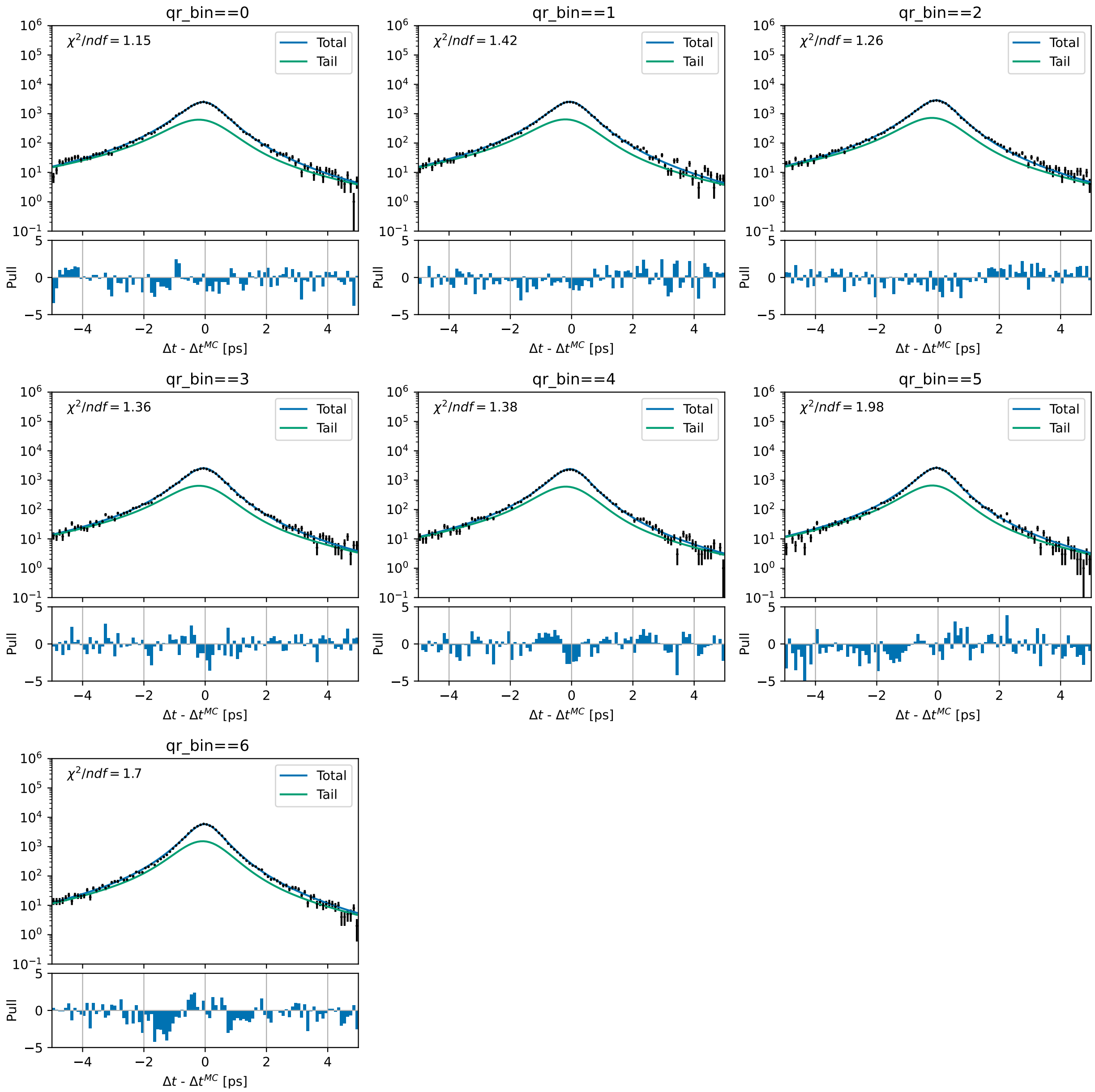


FIG. 35. *qr* bin dependence dependence of the resolution function fit to signalMC samples

5.2. Δt modeling for the signal

5.2.1. Longitudinal signal modelling

Time-dependent CP violation parameters, $S_{\rm CP}$ and $A_{\rm CP}$, are measured using the decay-time difference between the signal $B \to \rho^+\rho^-$ decay and the tag side B^0 decay in $\Upsilon(4S) \to B^0\bar{B}^0$ events. Equation 29 shows the PDF for the decay-time difference, with the time difference defined as $\Delta\tau = \tau_{\rm sig} - \tau_{\rm tag}$, where $\tau_{\rm sig}$ and $\tau_{\rm tag}$ are the decay times of signal and tag side B mesons, respectively. τ represents the lifetime of B^0, Δm_d is B^0 mixing frequency, q is the flavor of the tag-side B meson, with $q = 1$ for B^0 and $q = -1$ for $\bar{B}^0$, and $S_{\rm CP}$ and

A_{CP} is the time-dependent CP violation parameters for $B \to \rho^+\rho^-$.

$$\mathcal{P}(\Delta\tau, q) = \frac{\exp(-|\Delta\tau|/\tau)}{4\tau}(1 + q(S_{\mathrm{CP}}\sin(\Delta m_d\Delta\tau) + A_{\mathrm{CP}}\cos(\Delta m_d\Delta\tau))) \tag{29}$$

In the real data, it is impossible to measure $\Delta\tau$ directly. Instead, Δt is the decay time difference measured via the decay vertex difference along the boost axis. Δt is measured from the vertex position and is affected by the Δt resolution described in section 5.1. Furthermore, even if there are no detector resolution effects, $\Delta\tau$ and Δt are different values because the calculation of Δt is projected onto the boost direction, and it does not take into account the motion of B meson in Υ(4S) rest frame. Equation 30 shows Δt_{MC}, that is the measured Δt without detector response, where $K = \frac{\beta_B}{\beta} \simeq 0.225$, with β_B being the velocity of the B meson in $\Upsilon(4S)$ frame and β being the velocity of the $\Upsilon(4S)$, $\Sigma_\tau = \tau_{sig} + \tau_{tag}$ is the sum of the decay times of the signal and tag side B mesons, and θ^* is the polar angle of B^0_{sig} momentum with respect to the boost direction in the $\Upsilon(4S)$ frame. This effect is called kinematic approximation.

$$\Delta t_{\mathrm{MC}} = \Delta\tau + K\Sigma_\tau cos\theta^* \tag{30}$$

Equation 31 shows the PDF for the $\cos\theta^*$, which is derived from the spin=1 nature of $\Upsilon(4S)$. Figure 36 shows the distribution of $\cos\theta^*$ of the signal MC and the shape of the PDF, indicating that both share a similar shape.

$$\mathcal{P}(\cos\theta^*) = \frac{3}{4}(1 - \cos^2\theta^*) \tag{31}$$

This kinematic-approximation approach has been validated [17].

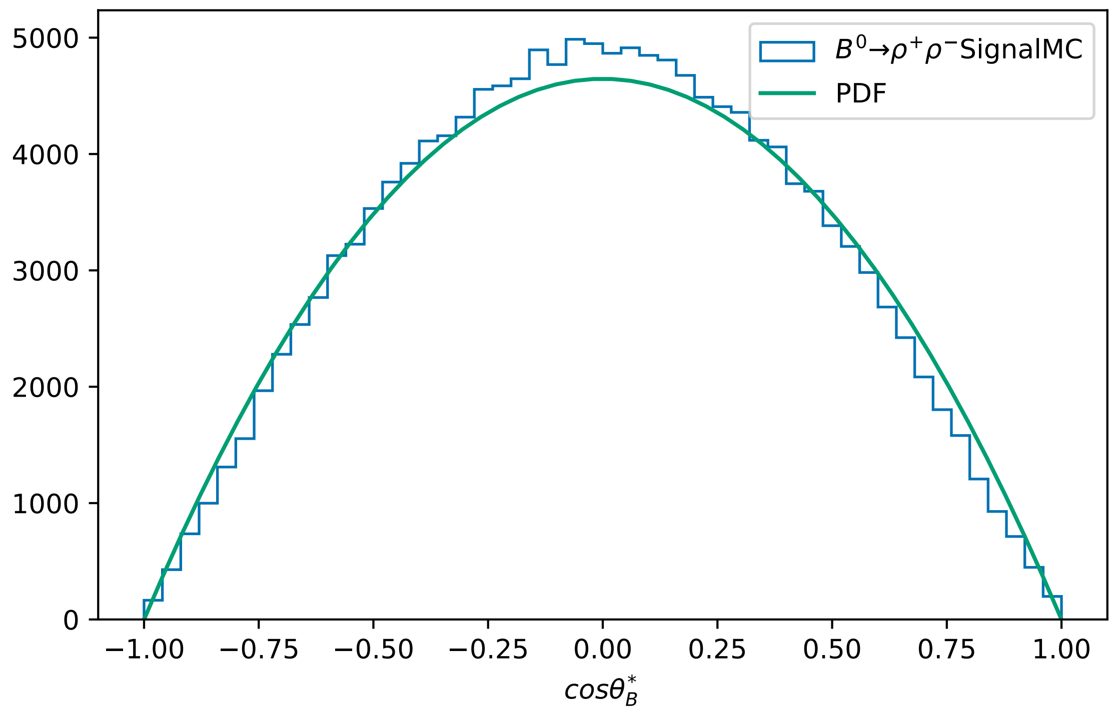


FIG. 36. $\cos\theta^*$ distribution of the signal MC and PDF

Another source of the discrepancy between the ideal case and real data is the misidentification of the flavor of the tag-side B^0. To account for this misidentification effect, the following parameters are adopted for the PDF used in the time-dependent fit: the wrong tag fraction $\omega = \frac{1}{2}(\omega_{B^0_{tag}} + \omega_{\bar{B}^0_{tag}})$, the wrong tag fraction difference $\Delta\omega = (\omega_{B^0_{tag}} - \omega_{\bar{B}^0_{tag}})$, and the efficiency difference $\mu = \frac{\varepsilon_{B^0_{\rm tag}} - \varepsilon_{\bar{B}^0_{\rm tag}}}{\varepsilon_{B^0_{\rm tag}} + \varepsilon_{\bar{B}^0_{\rm tag}}}$, where $\omega_{B^0_{tag}}$ and $\omega_{\bar{B}^0_{tag}}$ denote the wrong tag fractions for B^0 and $\bar{B}$ respectively, and $\varepsilon_{B^0_{\rm tag}}$ and $\varepsilon^0_{\bar{B}^0_{\rm tag}}$ represent the tagging efficiencies for B^0 and $\bar{B}^0$, respectively. We fixed the wrong tag fraction parameters to the results of the GFlaT paper [10]. Figure 37 compares the signal with the reference, demonstrating good consistency between them.

Equation 32 shows the Δt distribution taking into account the kinematic approximation, Δt resolution function $\mathcal{R}(\delta\Delta t;\sigma)$, wrong tag fraction ω, wrong tag fraction difference $\Delta\omega$, and efficiency difference μ,

$$
\begin{aligned}
&PDF_{obs}(\Delta t, \sigma, q) = \\
&\int_{-1}^{1} d\cos\theta^* \left[\frac{3}{4}(1-\cos^2\theta^*)\right] \int_{-\infty}^{\infty} d\Delta t_{\rm MC} \int_{\frac{|\Delta t_{\rm MC}|}{1+K\cdot(sign)\Delta t_{\rm MC}\cos\theta^*}}^{\infty} d\Sigma_\tau \\
&\frac{1}{2\tau}\exp\left(\frac{-\Sigma_\tau}{\tau}\right)\Big\{1 - q\Delta\omega + q\mu(1-2\omega) \\
&+ [q(1-2w) + \mu(1-q\Delta w)][S_{CP}\sin(\Delta m_d\Delta\tau) + A_{CP}\cos(\Delta m_d\Delta\tau)]\Big\} \otimes \mathcal{R}(\delta\Delta t, \sigma)
\end{aligned}
\tag{32}
$$

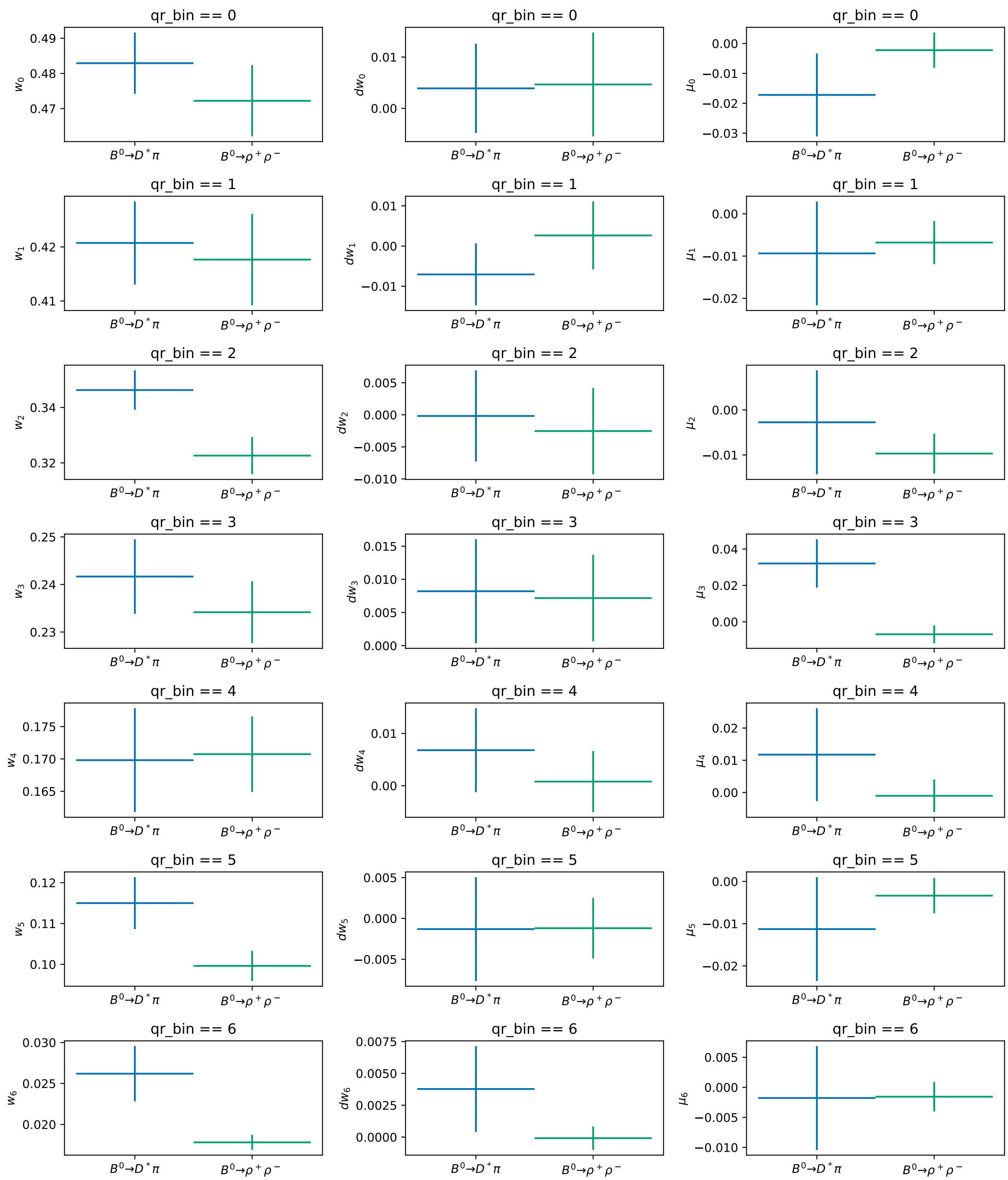


FIG. 37. Comparison of the wrong tag fraction parameters between $B^0 \to \rho^+\rho^-$ signal MC and reference value [10]. The first column shows the wrong tag fraction, the second shows the wrong fraction difference between B^0 tag and $\bar{B}^0$ tag, and the third shows the efficiency difference.

5.2.2. Lifetime fitting

We validated the fitter before applying it to CP violation parameters by fitting to Δt distribution without flavor information. In this validation, fit parameters are the lifetime of B^0 and the resolution function parameters, which would be determined in the fitting to the control samples. The PDF utilized in the fitting is the same as Equation 32 except for the flavor information. Equation 33 shows the PDF for the fitting.

$$
\begin{aligned}
PDF_{obs}(\Delta t, \sigma, q) = \int_{-1}^{1} d\cos\theta^* \left[\frac{3}{4}(1-\cos^2\theta^*)\right] \int_{-\infty}^{\infty} \Delta t_{\mathrm{MC}} \int_{\frac{\Delta t_{\mathrm{MC}}}{1+K\cdot(sign)\Delta t_{\mathrm{MC}}\cos\theta^*}}^{\infty} d\Sigma_\tau \\
\frac{1}{2\tau}\exp\left(\frac{-\Sigma_\tau}{\tau}\right)\mathcal{R}(\delta\Delta t, \sigma)
\end{aligned}
\tag{33}
$$

The fitting is performed to a simulated sample of the signals, and the fitted results are shown in Table 24 and Figure 38. The fit agrees with the data very well, and the fitted resolution function parameters are consistent with the fit results to the $\Delta t - \Delta t_{\mathrm{MC}}$ distributions. In addition, the fitted B^0 lifetime, τ, is consistent with the MC truth value, and the resolution function parameter does not introduce bias into the estimation of the B^0 lifetime. As shown above, the fitter works well.

Parameter	Fit result	$\Delta t - \Delta t_{\mathrm{MC}}$ fit results or MC truth	Floated in the fit
τ	1.513 ± 0.008	1.519	Yes
f_{OL}	0.0015	0.0015	No
c	4.72	4.72	No
f_{R}	0.23	0.23	No
f_{exp}	0.19	0.19	No
μ_{main}	-0.05 ± 0.03	-0.14	Yes
μ_{tail}	-0.75 ± 0.08	-0.68	Yes
s_{main}	1.13 ± 0.03	1.15	Yes
s_{tail}	2.29 ± 0.1	2.0	Yes
σ_{OL}	200.0	200.0	No
c_6	4.94	4.94	No
f_{R6}	0.26	0.26	No
f_{exp6}	0.11	0.11	No
μ_{main6}	-0.1 ± 0.08	-0.08	Yes
μ_{tail6}	-0.06 ± 0.2	-0.34	Yes
f_μ	0.34	0.34	No
$f_{\sigma R}$	0.24	0.24	No
$f_{\sigma L}$	0.12	0.12	No
f_{max}	0.12	0.12	No
$C_{f_{\mathrm{tail}}}$	0.26	0.26	No
N	100000.13 ± 316.23	100000	Yes

TABLE 24. Fit result to the Δt distribution of the simulated signal samples

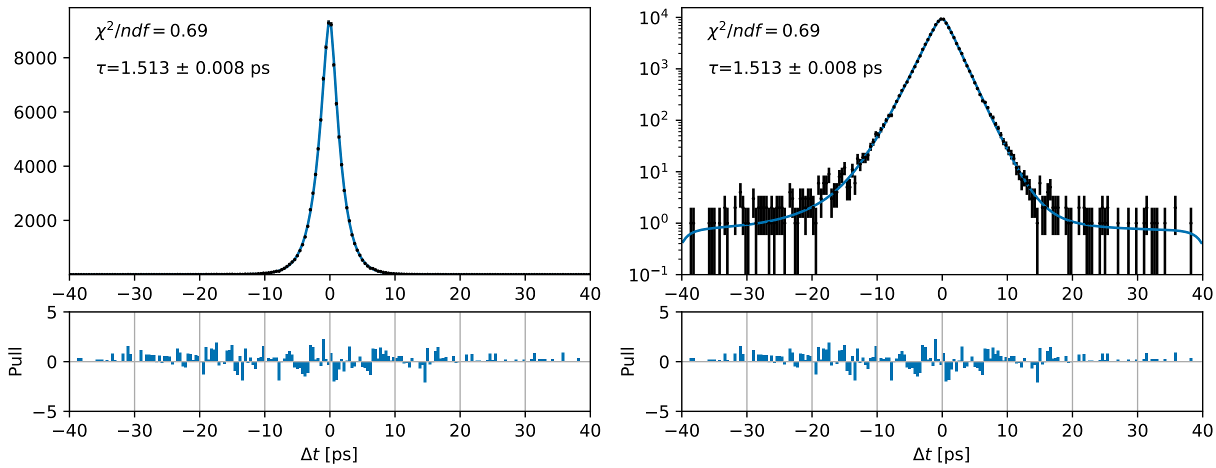


FIG. 38. Fit projection of the Δt fitting to signalMC samples

5.2.3. CP fitting

CP fitting to the simulated samples of the signal was performed. The CP violation parameters for the simulated parameters are set to $S_{\mathrm{CP}} = -0.14$ and $A_{\mathrm{CP}} = 0.00$ which is the same as the world average. The PDF for the fitting is described in Equation 32, with all the resolution function parameters fixed to the fit results for the $\Delta t - \Delta t_{\mathrm{MC}}$ distribution. The shape of the resolution function and the parameters of the wrong tag fraction were adjusted according to the qr bin value. Table 25 and Figure 39 show the fitting results, and the fitted values are consistent with the input values. The fitted distribution agrees with the data. Figure 40 shows the qr-bin dependence of the fitted distributions. The fitter works well in all the qr bins.

Parameter	Fit result	MC truth
A_{CP}	-0.004 $\pm$ 0.005	0.00
S_{CP}	-0.144 $\pm$ 0.007	-0.14
N	183251.3 $\pm$ 428.1	183251

TABLE 25. CP fit results to signal MC samples

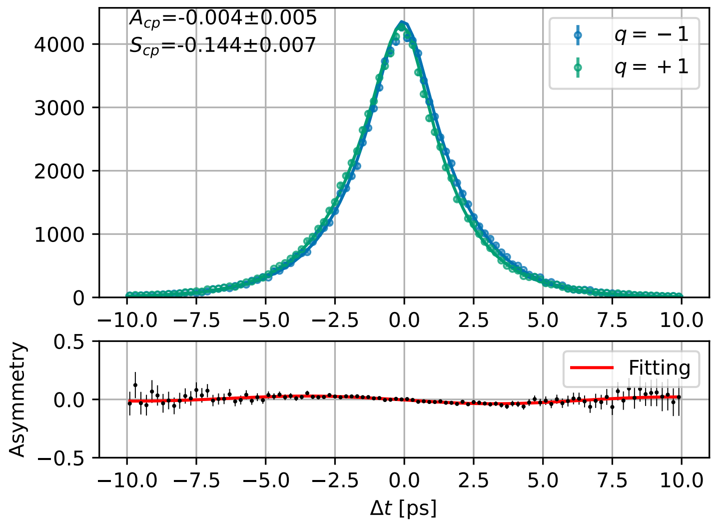


FIG. 39. CP fit projection to signalMC samples. $S_{\rm CP} = -0.14$ and $A_{\rm CP} = 0.00$.

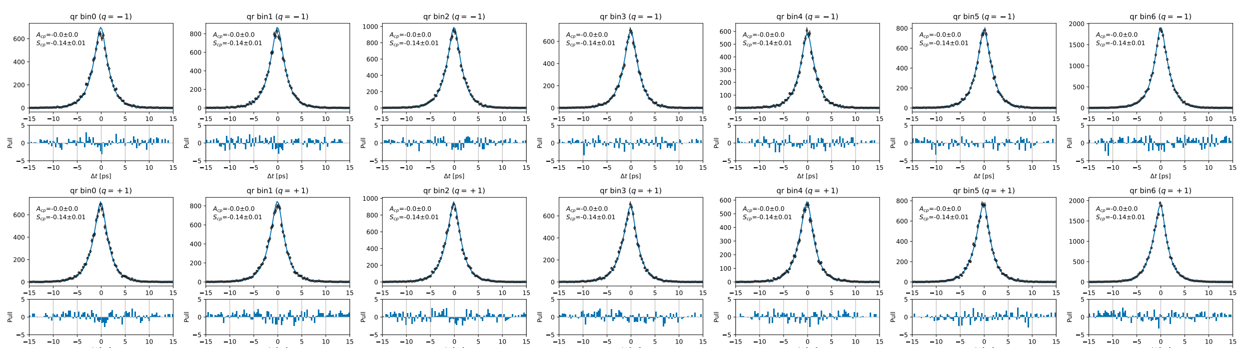


FIG. 40. qr bin dependence of the CP fit to signalMC samples.

To validate the fitter, a bootstrap study was performed. Figure 41 shows the result of the bootstrap study for CP fitting to the simulated samples of the longitudinal signal. The CP violation parameters are set to be the same as the world average, which is $S_{\rm CP} = -0.14$ and $A_{\rm CP} = 0.00$, and the number of the events in each bootstrap sample is set to 500, that is the expected signal yield in a dataset of 365 fb^{-1}. The results show that the fitter is not biased, and error estimation is accurate. Figure 42 shows the linearity test for the CP fitting. The same bootstrap study was performed in the test while changing the CP violation parameters. The fitter shows good linearity in all the CP violation parameter regions. The pull value of $S_{\rm CP}$ in Figure 41 appears slightly biased, but it can be attributed to the statistical uncertainty inherent in the total signalMC samples because the results demonstrate the good linearity of $S_{\rm CP}$.

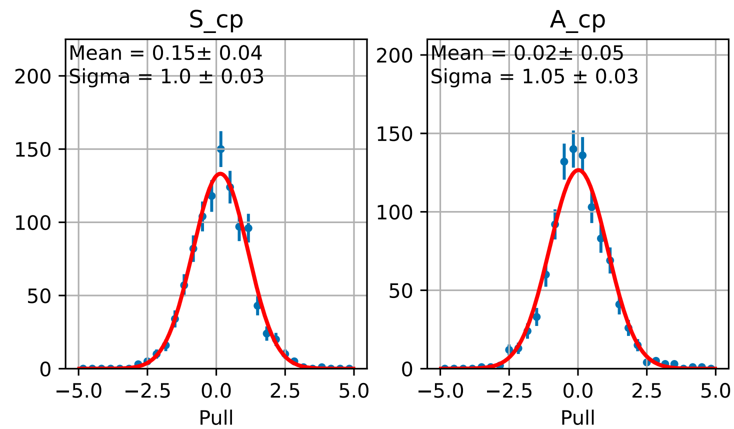


FIG. 41. Bootstrap test of the CP fitting to signalMC samples

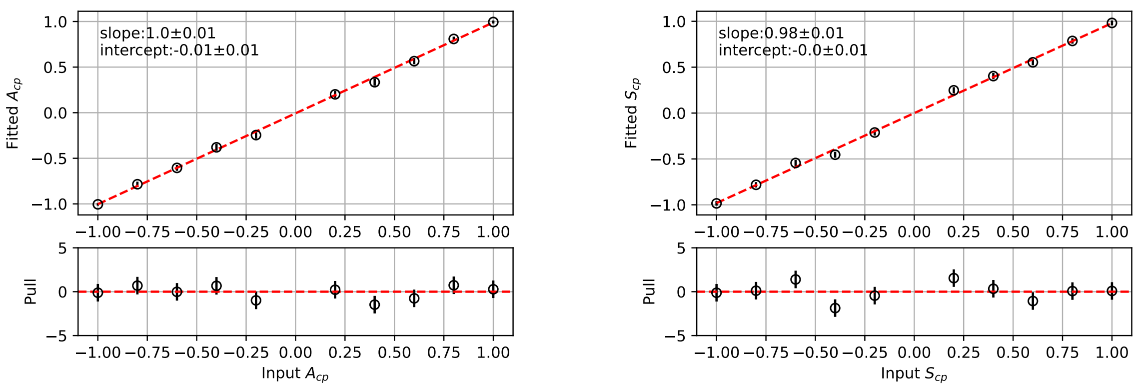


FIG. 42. Linearity test of the CP fitting to signalMC samples

5.2.4. Transverse signal modelling

Transversely polarized $B \to \rho^+\rho^-$ decays contain both CP-Even and CP-Odd contributions and do not show CP violating effects. Therefore, the transverse signal must be modelled separately from the longitudinal signal in both the signal extraction fitting and the time-dependent CP violation fitting. The vertex of the transverse signal is correctly reconstructed, thus the Δt distribution follows the B^0 lifetime distribution that is affected by the detector resolution and kinematic approximation. Considering then, the fit PDF for the transverse signal is Equation 33. Figure 43 shows the fitting result, and Table 26 shows the fitted and MC truth values. The fitter works well, and the fitted values are consistent with MC truth. When fitting to real data, the lifetime is fixed to the PDG values, and the resolution function parameters are fixed to the fit results from the control-mode study.

Parameter	Fit result	$\Delta t - \Delta t_{\mathrm{MC}}$ fit results or MC truth	Floated in the fit
τ	1.530 ± 0.008	1.519	Yes
f_{OL}	0.0	0.0	No
c	4.67	4.67	No
f_{R}	0.24	0.24	No
f_{exp}	0.21	0.21	No
μ_{main}	-0.14 ± 0.04	-0.16	Yes
μ_{tail}	-1.18 ± 0.12	-0.63	Yes
s_{main}	1.14 ± 0.04	1.15	Yes
s_{tail}	2.66 ± 0.14	2.09	Yes
σ_{OL}	200.0	200.0	No
c_6	5.14	5.14	No
f_{R6}	0.11	0.11	No
f_{exp6}	0.09	0.09	No
μ_{main6}	-0.04 ± 0.09	-0.11	Yes
μ_{tail6}	-0.43 ± 0.25	-0.33	Yes
f_{μ}	2.0	2.0	No
$f_{\sigma R}$	0.27	0.27	No
$f_{\sigma L}$	0.07	0.07	No
f_{max}	0.18	0.18	No
$C_{f_{\mathrm{tail}}}$	0.3	0.3	No
N	100000.13 ± 316.23	100000	Yes

TABLE 26. Δt fit result to the transverse signal samples

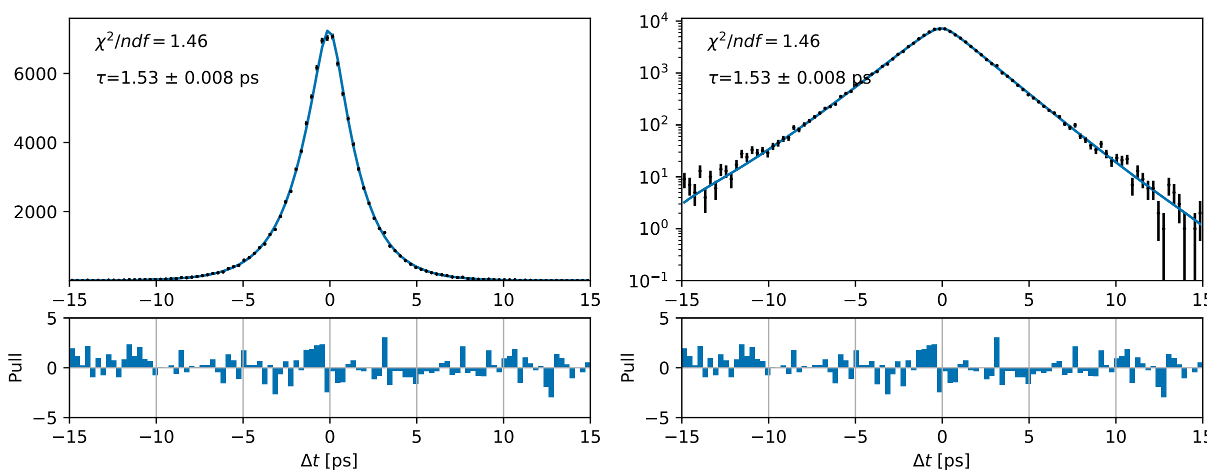


FIG. 43. Δt fit projection for the transverse signal using the simulated samples.

5.2.5. Self-crossfeed signal modelling

As described in Section 2, the self-crossfeed signals, which are mis-reconstructed signal events, are categorized into 2π Sxf and $< 2\pi$ Sxf. The 2π Sxf is expected to have the same Δt distribution as the signal because it has two correctly reconstructed $\pi^{\pm}$ and has the same vertex position as the correctly reconstructed signal. In contrast, $< 2\pi$ Sxf is expected to have a different Δt distribution from the signal because only one $\pi^{\pm}$ from the signal is used in the vertexing. However, it still has a part of the CP violating effects from the signal. Figure 44 shows the Δt distribution difference between the signal and self-crossfeed. The 2π self-crossfeed has a distribution similar to that of the signal, with any small differences potentially attributed to the differences of the qr-bin distributions. Figure 45 shows the differences in the Δt resolution between self-crossfeed and signals, while Figure 46 shows the difference in the qr distribution between signals and self-crossfeed. On the other hand, $< 2\pi$ self-crossfeed events show a distribution difference from the signal as expected because of the mis-reconstructed vertex position.

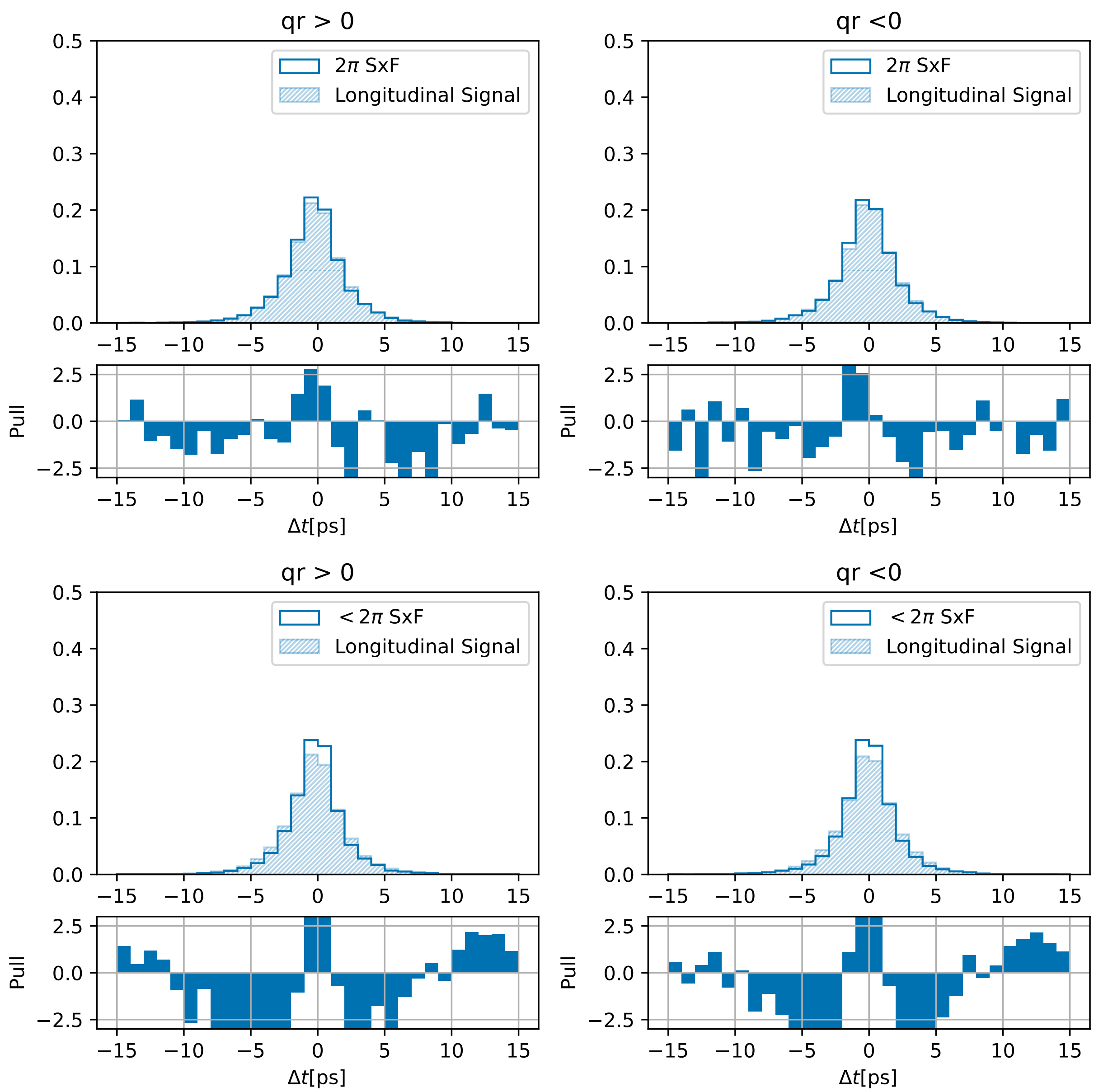


FIG. 44. Comparison of the Δt distribution between signal and self-crossfeed in the simulated samples.

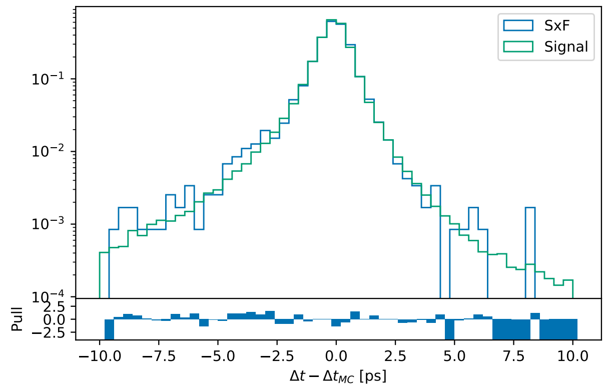


FIG. 45. Δt resolution difference between the longitudinal signal and 2π sxf using signal MC

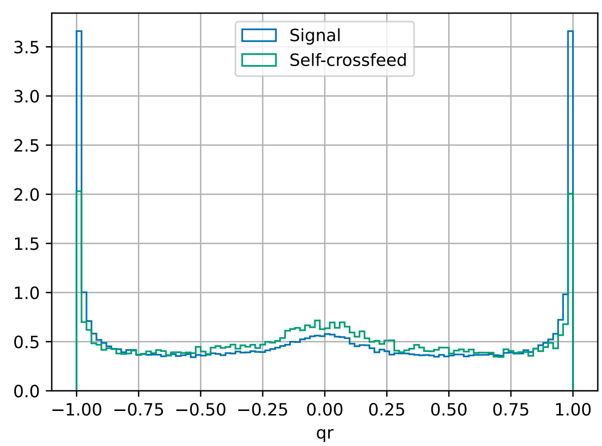


FIG. 46. qr difference between the longitudinal signal and 2π sxf using signal MC

Figure 47 and Table 27 show the CP fitting results to the 2π SxF samples. The CP parameters are consistent with the MC truth, and the fitted distributions match the data well. The 2π SxF samples display the distribution different from those in the signal extraction fitting, but they can still be utilized as a signal in the Δt fitting.

Parameter	Fit result	MC truth
$A_{\rm CP}$	-0.012 ± 0.011	0
$S_{\rm CP}$	-0.134 ± 0.016	-0.14
N	42124.0 ± 205.2	42124

TABLE 27. CP Fit result to run-dependent simulated sample of 2π SxF events

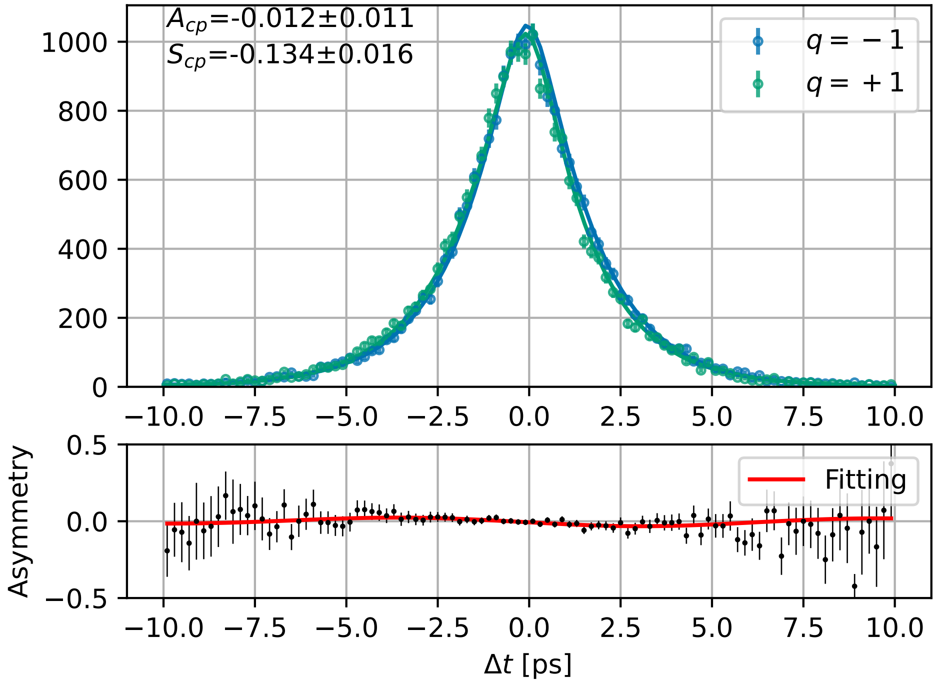


FIG. 47. CP fit projection for 2π self-crossfeed signal in the simulated samples.

The distributions for $< 2\pi$ SxF samples are different from the signal distributions because at least one charged track from the signal is missing in the reconstruction. To account for this difference, it is necessary to revise the lifetime parameter, τ, which originally corresponds to the B^0 lifetime. In the $< 2\pi$ SxF, the vertex position is determined using one or fewer $\pi^{\pm}$ from the signal combined with one or more charged particles from the tag side decay. This results in the B^0 lifetime appearing different. Table 28 and Figure 48 show the CP fit results with fixed τ and floated τ, respectively. In both cases, there is no significant bias on the CP parameters. However when the fit is performed with a fixed lifetime, the fitted result shows a discrepancy. This discrepancy is migrated by applying an effective lifetime. In this analysis, we used the modelling with an effective lifetime.

Parameter	Fit result	MC truth
A_{CP}	-0.014 ± 0.012	0
S_{CP}	-0.132 ± 0.022	-0.14
τ	1.259 ± 0.01	1.519
N	31805.0 ± 178.3	31805

Parameter	Fit result	MC truth
A_{CP}	-0.014 ± 0.013	0
S_{CP}	-0.128 ± 0.021	-0.14
τ	1.519	1.519
N	31805.0 ± 178.3	31805

TABLE 28. CP Fit result to run-dependent simulated samples of $< 2\pi$ SxF events (Left: With effective lifetime, Right: With fixed lifetime)

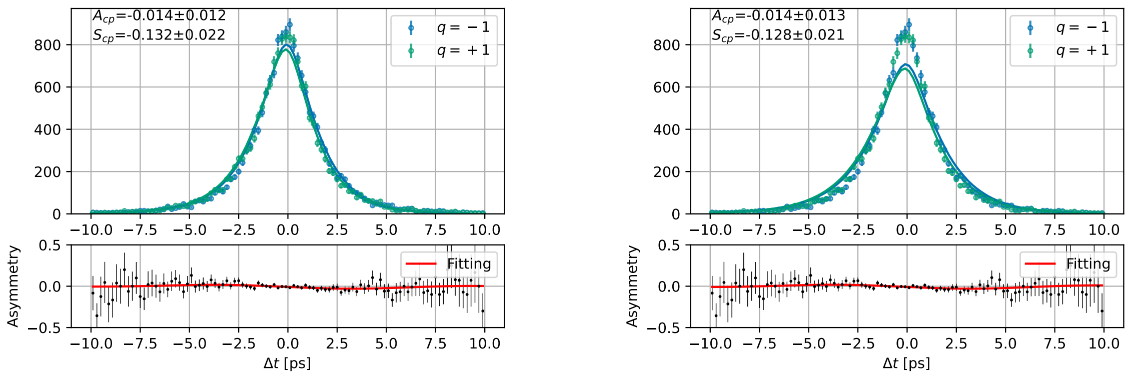


FIG. 48. CP fit projection for the simulated samples of $< 2\pi$ self-crossfeed signal. (Left: With effective lifetime, Right: With fixed lifetime

5.3. Δt modeling for the backgrounds

5.3.1. $q\bar{q}$ modelling

The PDF used for the $q\bar{q}$ is shown in Equation 34. Table 29 and Figure 49 show the results of the fitting to the Δt distribution for $q\bar{q}$ samples using Equation 34. The fitted distribution matches the samples well. The shape will be calibrated using $M_{\rm bc}$ sideband data and fixed in the fitting to the signal region.

$$\mathcal{P}_{qq}(\Delta t) =(1 - f_{\rm tail})G(\Delta t, \mu, s_{\rm main}\sigma) + f_{\rm tail}G(\Delta t, \mu, s_{\rm tail}\sigma) \tag{34}$$

Parameter	Fit result
μ	0.002 ± 0.003
$f_{\rm tail}$	0.05 ± 0.0
$s_{\rm main}$	1.38 ± 0.01
$s_{\rm tail}$	9.11 ± 0.21
N	24970.0 ± 158.0

TABLE 29. Fit result to the Δt distribution for $q\bar{q}$ samples using the simulated samples.

5.3.2. $B\bar{B}$ modelling

The $B\bar{B}$ backgrounds consist of the B^0 decay, and the Δt distribution looks similar to the B^0-lifetime distribution. However, most of the B candidates are combinatorial, so the lifetime appears to be shifted, akin to the behavior observed in $< 2\pi$ SxF (Section 5.2.5). The PDF to model the Δt distribution for $B\bar{B}$ is the same as Equation 33 with the B^0 lifetime as a floating parameter. All the resolution function parameters are fixed in the fitting, and the fit parameters are τ and N. Figure 50 and Table 30 show the results of the fitting, and Figure 51 shows the qr-bin dependency of the Δt PDF. The fitted distribution matches the samples well, and the the estimated lifetime is slightly shorter than the actual

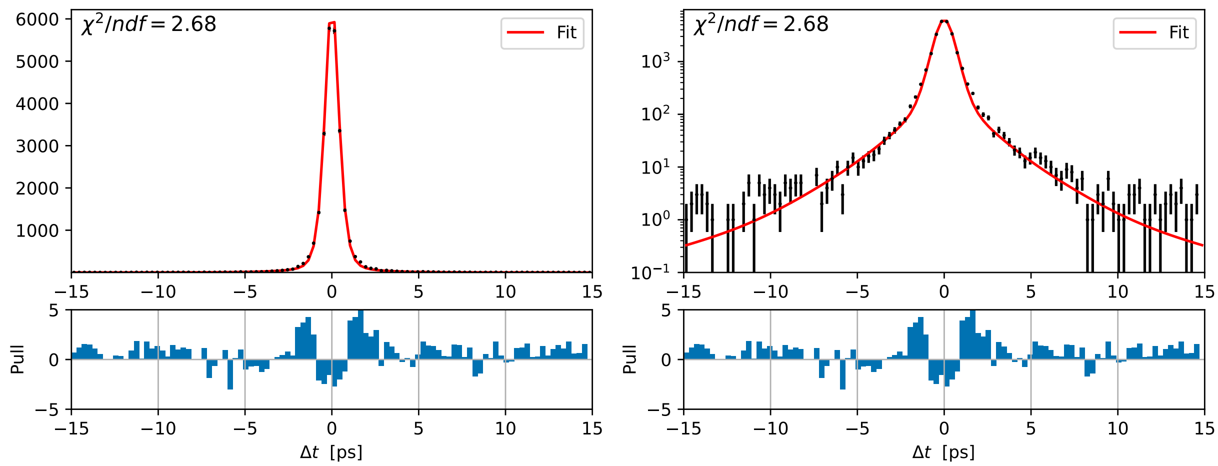


FIG. 49. Fit projection to the Δt distribution of $q\bar{q}$ samples in the simulated samples.
(Left: Linear scale, Right: Log scale)

B^0 lifetime. The fitted effective lifetime differs from the results from the $< 2\pi$ SxF fitting, and therefore an effective lifetime fitting is necessary for each type of sample.

Parameter	Fit result
τ	1.39 ± 0.02
N	5747.0 ± 75.8

TABLE 30. Fit result to the Δt distribution of $B\bar{B}$ in the simulated samples.

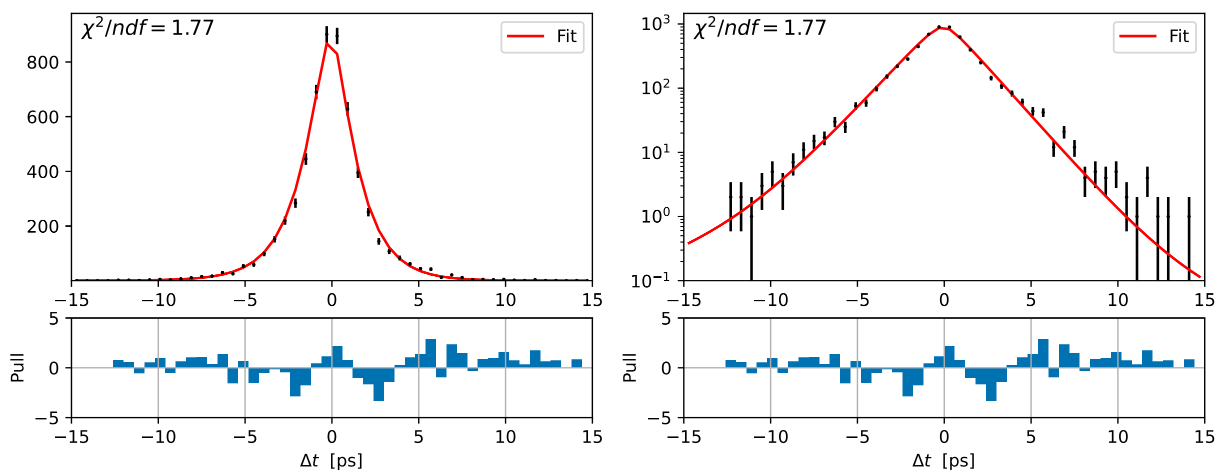


FIG. 50. Fit projection to the Δt distribution of $B\bar{B}$ in the simulated samples.
(Left: Linear scale, Right: Log scale)

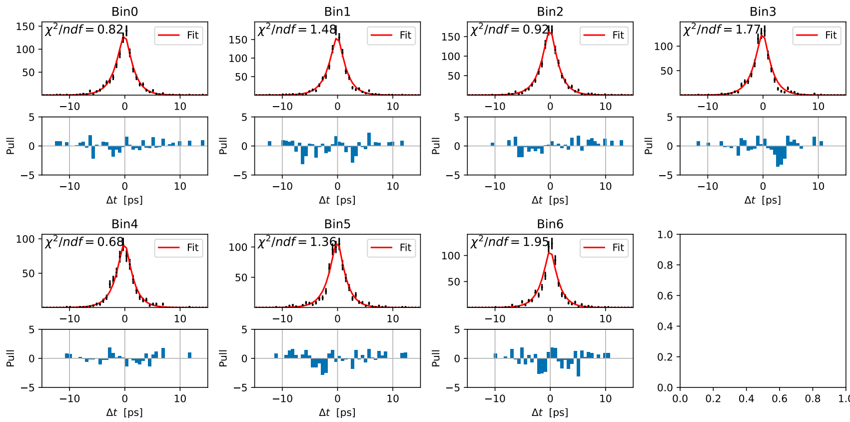


FIG. 51. qr bin dependency of Δt PDF for BB background using the simulated samples.

5.4. $\tau\tau$ Modeling

Equation 35 shows the Δt PDF modeling for the $\tau\tau$ backgrounds. The parametrization is the same PDF which is for the qq backgrounds. Figure 52 shows the fit result to Δt distribution using the equation, and it agrees with MC well.

$$\mathcal{P}_{\tau\tau}(\Delta t) = (1 - f_{\mathrm{tail}})G(\Delta t, \mu, s_{\mathrm{main}}\sigma) + f_{\mathrm{tail}}G(\Delta t, \mu, s_{\mathrm{tail}}\sigma) \tag{35}$$

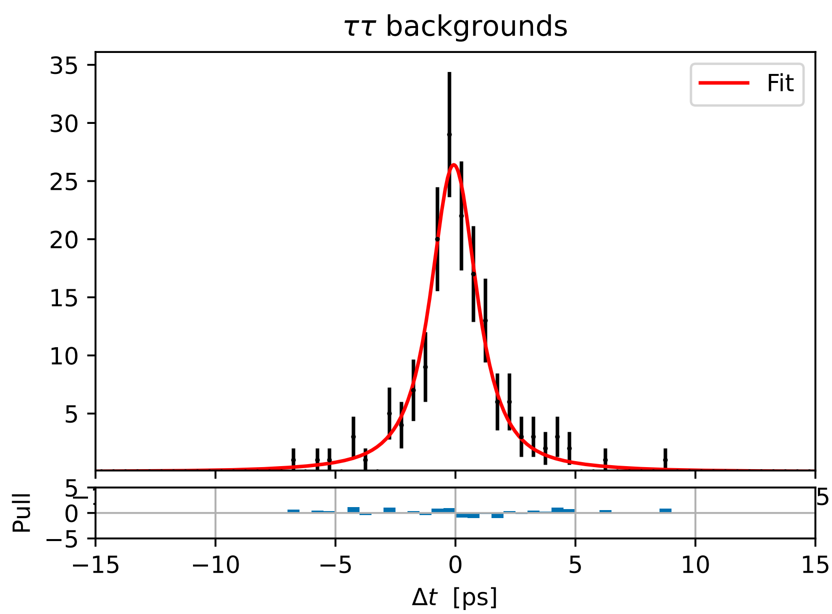


FIG. 52. Δt modeling for $\tau\tau$ background using simulated samples.

5.5. Δt modeling for the peaking backgrounds

The peaking background modes provide a different Δt distribution, which needs to be modeled separately from the generic $B\bar{B}$ backgrounds. Although the shape of these peaking backgrounds is similar to the generic $B\bar{B}$ backgrounds, the effective lifetime is different. To consider this, Δt fitting to each peaking background mode was performed. The PDF used for this fitting is the same as that described in Equation 33, and the fit parameters are N and τ. Table 31 and Figure 53 show the fitting results. The effective lifetime varies depending on the decay modes and needs to be modelled for each mode. Figure 54 shows the fitted distribution for each peaking decay mode, and all fitted distributions agree with the dataset.

Mode	Fitted lifetime
$B^0 \to \rho^+\pi^-\pi^0$	1.546 ± 0.007
$B^0 \to \pi^+\pi^-\pi^0\pi^0$	1.540 ± 0.011
$B^0 \to a_1^\pm\rho^\mp$	1.375 ± 0.016
$B^0 \to \rho^- K^{*+}$	1.480 ± 0.038
$B^0 \to \rho^- K_0^{*+}$	1.446 ± 0.064
$B^0 \to \pi^\pm\pi^\mp\pi^0$	1.396 ± 0.016
$B^0 \to a_1^+\pi^-$	1.541 ± 0.066
$B^0 \to a_1^0\pi^0$	1.443 ± 0.006
$B^0 \to \rho^\pm\pi^\mp$	1.423 ± 0.017
$B^\pm \to \rho^\pm\pi^0$	1.462 ± 0.006
$B^\pm \to a_1^0\pi^\pm$	1.505 ± 0.017
$B^\pm \to \rho^\pm\rho^0$	1.430 ± 0.009
$B^\pm \to a_1^\pm\pi^0$	1.403 ± 0.012
$B^\pm \to \rho^\pm a_1^0$	1.399 ± 0.015
$B^\pm \to K\pi\pi$	1.324 ± 0.048
$B^\pm \to K_0^{*\pm}\pi^0$	1.466 ± 0.016
Rare peaking backgrounds	1.558 ± 0.009

TABLE 31. Fitted lifetime to simulated samples for the peaking background modes

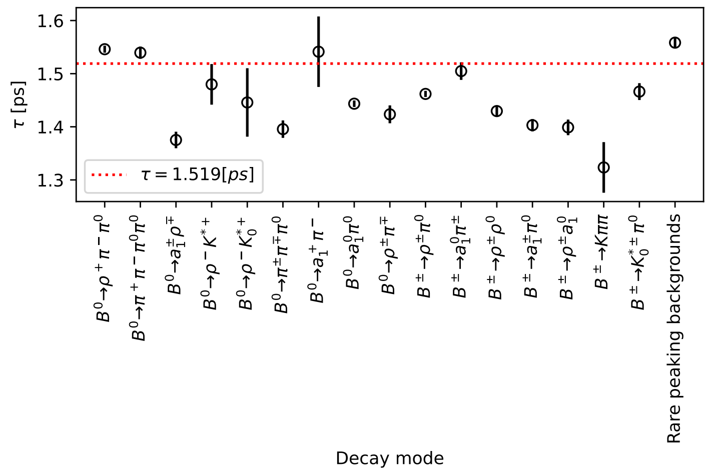


FIG. 53. Fitted lifetime to simulated samples for the peaking background modes

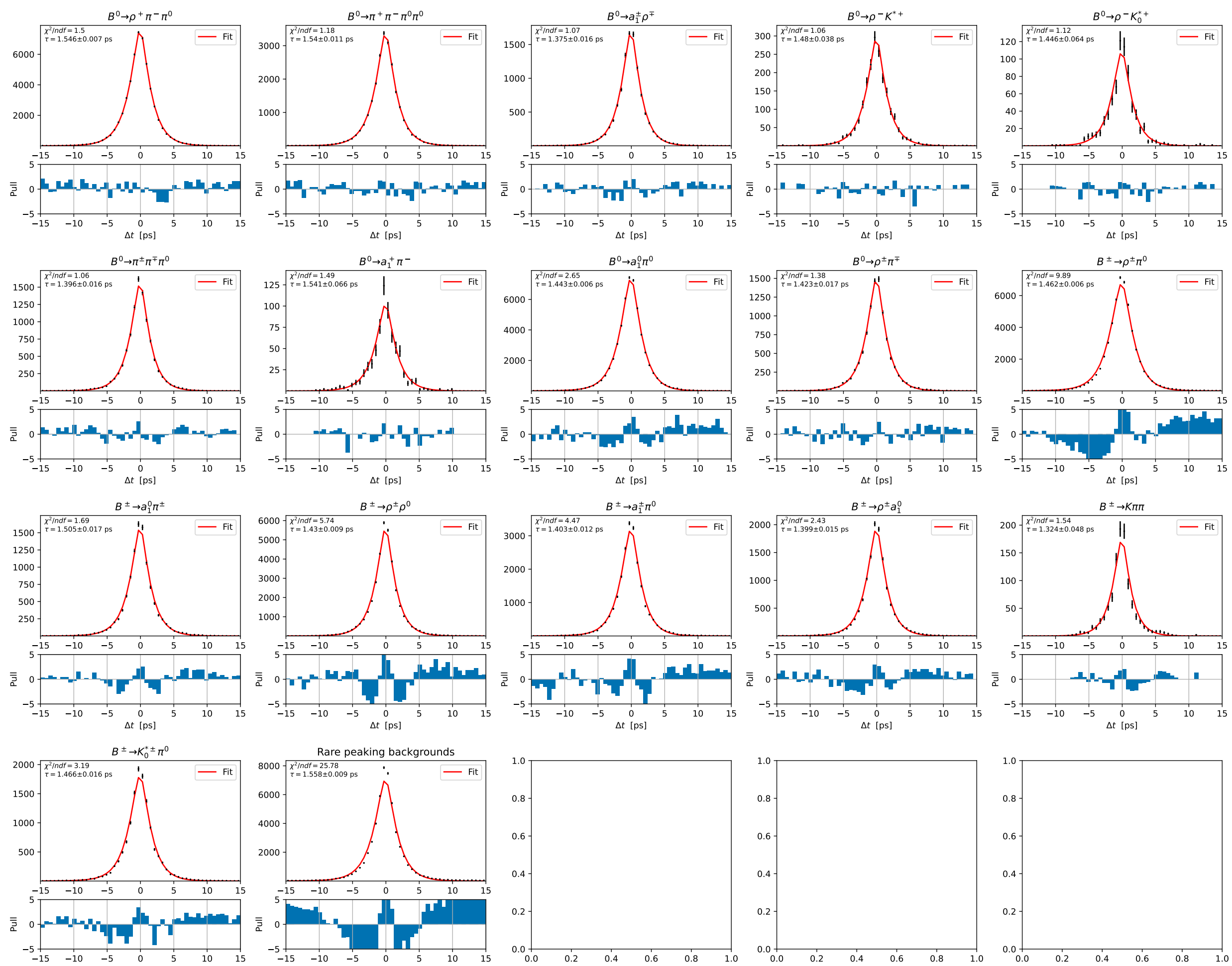


FIG. 54. Δt modelling for the peaking background modes using simulated samples.

5.6. Fit model

Equation 36 shows the PDF used in the fitting. i, j run over fit components, which include Longitudinal-signal, Transverse-signal, self-crossfeed (2π and $< 2\pi$), $q\bar{q}$, $B\bar{B}$, and the peaking backgrounds. $\mathcal{P}_i$ is the PDF for each components, and f_i is the event fraction for each fit component. Utilizing the PDF for the signal extraction described in Section 4, the probability associated with each fit component of the event was calculated. The CSMVA output is not utilized in this calculation due to its significant correlation with the qr bin. $\mathcal{P}(qr)$ represents the event fraction within each qr bin. Figure 55 displays the event fraction for each qr bin, which is incorporated into the calculation of the event fraction. $\mathcal{P}(r)$ demonstrates the power to separate signal from background. The shape of the Δt PDF and the resolution function are determined by the qr bin. Therefore, to mitigate the risk of bias due to the Punzi effect [18], the qr bin must be included in the process of signal extraction.

$$\begin{aligned} \mathcal{P}(\Delta t) &= \sum_i N_i f_i \mathcal{P}_i(\Delta t) \\ f_i &= \frac{\mathcal{P}_i(\Delta E, m_{\pi^+\pi^0}, m_{\pi^-\pi^0}, \cos\theta_\rho^+, \cos\theta_\rho^-)\mathcal{P}_i(r)}{\sum_j \mathcal{P}_j(\Delta E, m_{\pi^+\pi^0}, m_{\pi^-\pi^0}, \cos\theta_\rho^+, \cos\theta_\rho^-)\mathcal{P}_j(r)} \end{aligned} \tag{36}$$

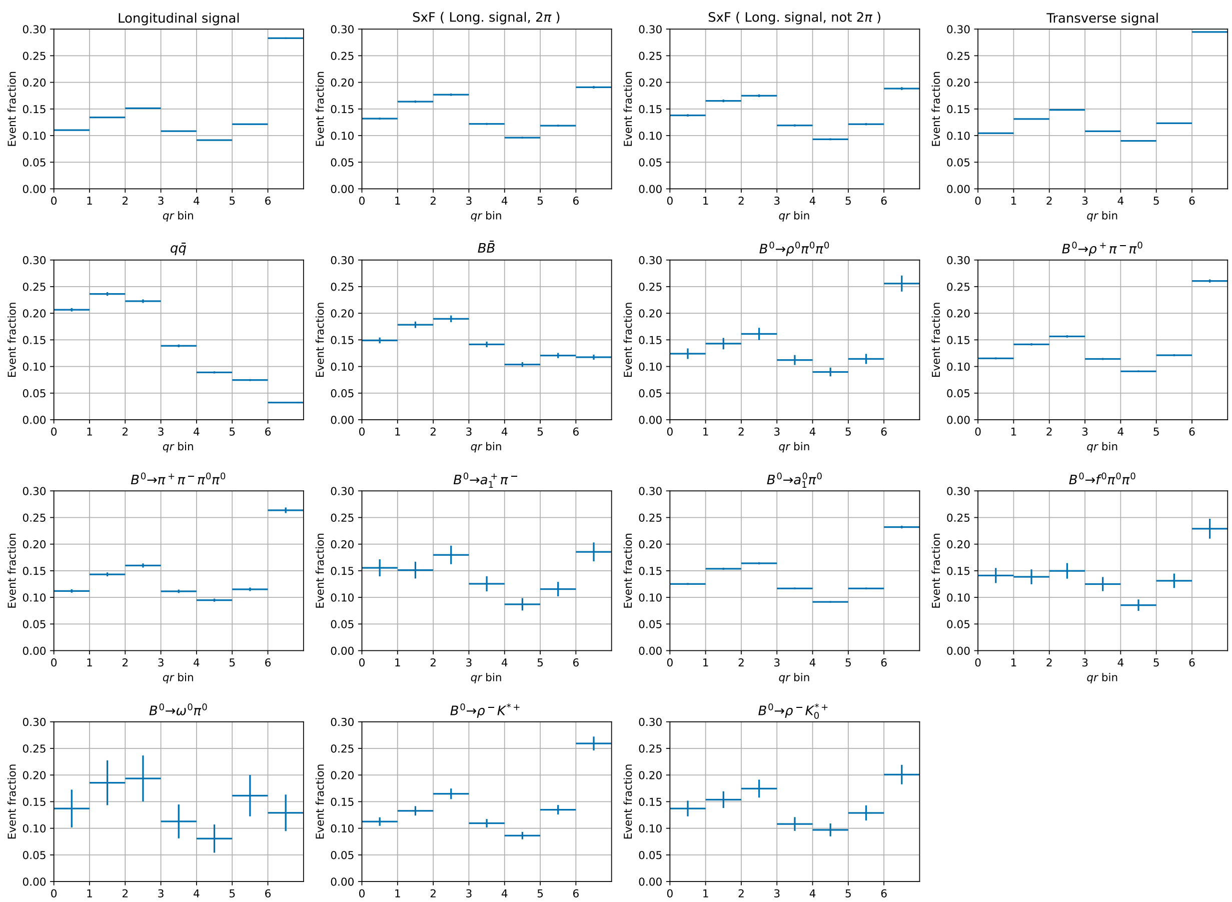


FIG. 55. Event fraction in each qr bin estimated by the simulated samples.

Table 32 and Figure 56 show the results of Δt fitting to MC samples, and Figure 57 shows the likelihood scan result for S_{CP} and A_{CP}. The dataset utilized here is identical to that described in Section 3. The obtained results are in good agreement with the inputs, and the fitted distribution closely matches the dataset. Figure 58 shows the S-plot of the fitting. Here, only the longitudinal-signal contribution and the fit results are plotted. The S-plot is compared to the signal MC distribution in Figure 59. The signal distribution extracted through the fitting agrees with that of the MC truth.

Parameter	Fit result	Input
A_{cp}	0.087 ± 0.122	0.000
S_{cp}	0.080 ± 0.170	-0.14

TABLE 32. Result of the CP fitting to simulated samples corresponds to 362 fb^{-1}.

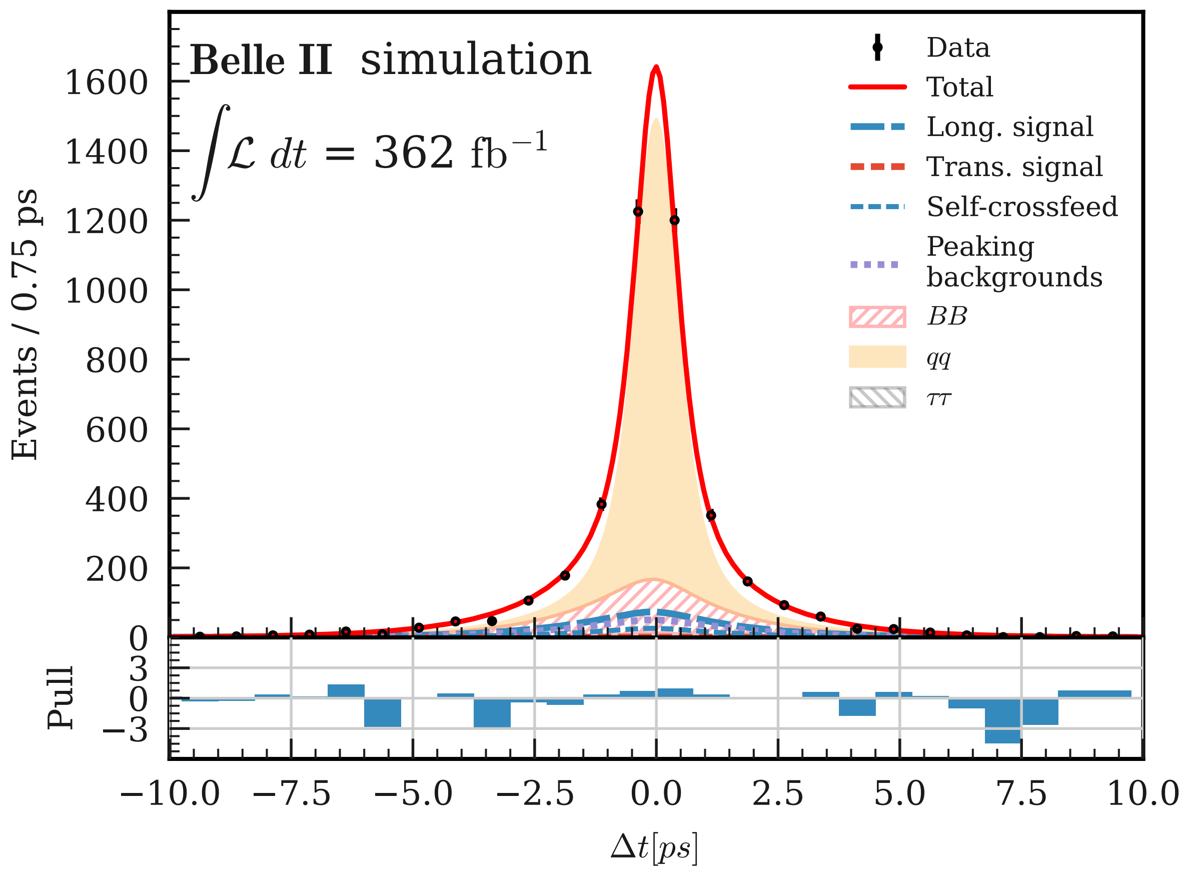


FIG. 56. Result of the Δt fitting to simulated samples corresponds to 362 fb^{-1}

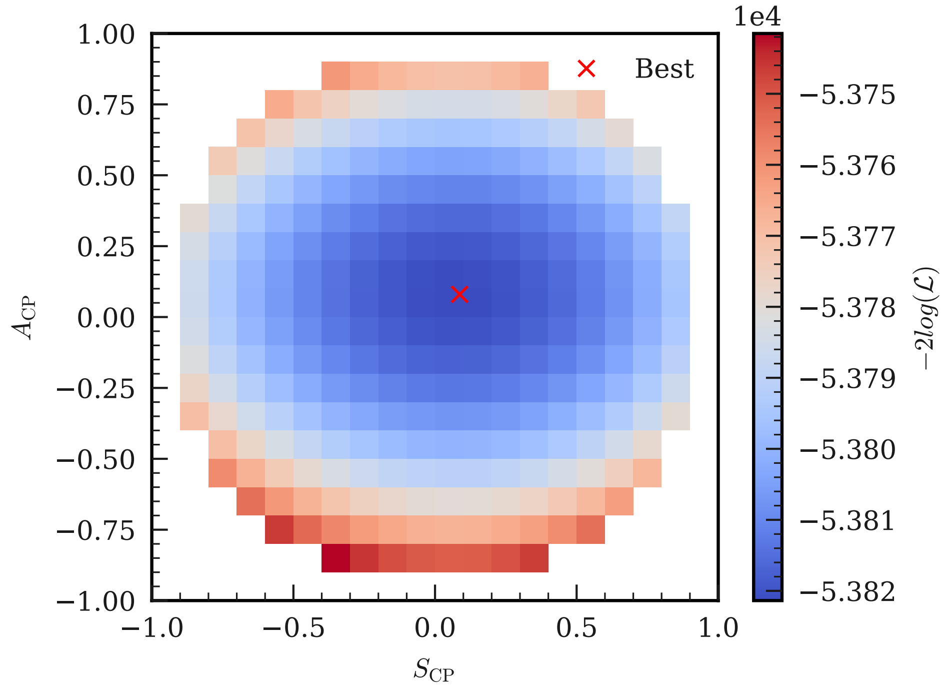


FIG. 57. Likelihood scan for $\S_{\rm CP}$ and $A_{\rm CP}$.

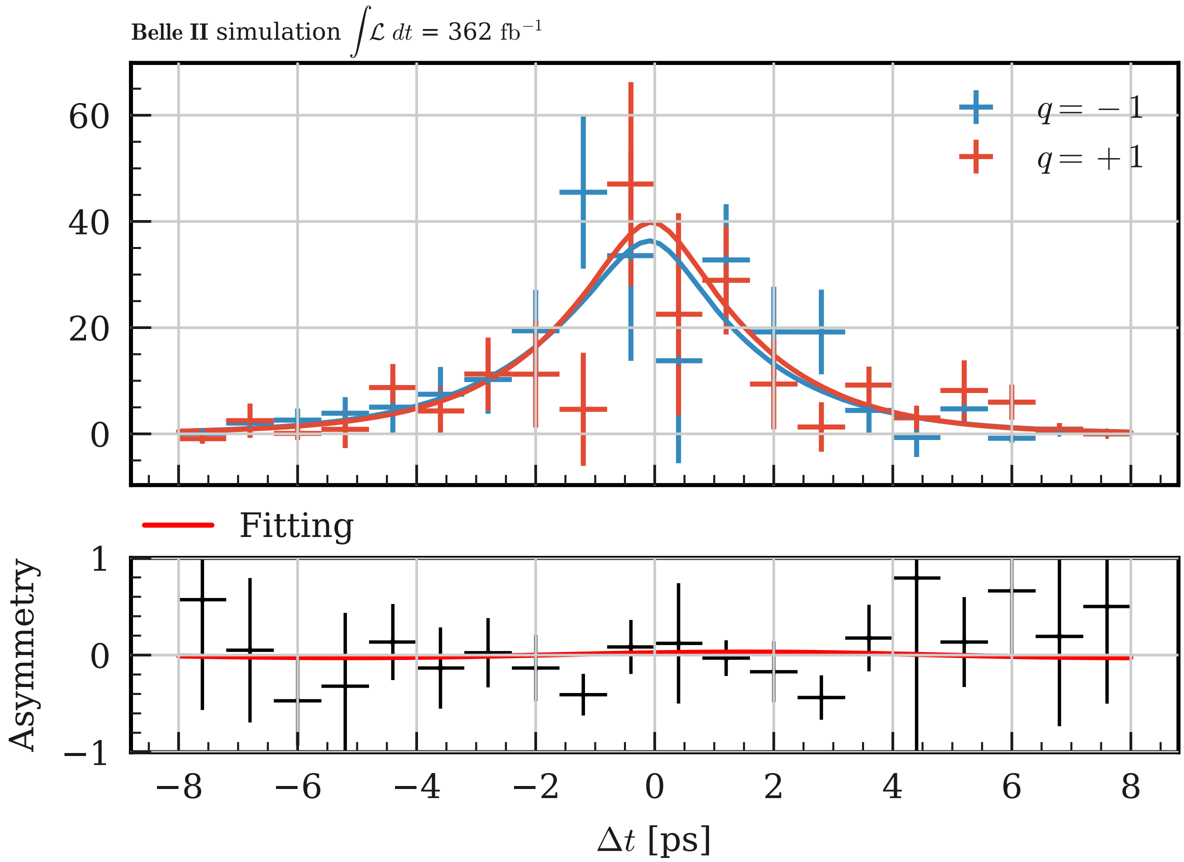


FIG. 58. S-plot of Δt fitting to simulated samples corresponds to 362 fb^{-1}

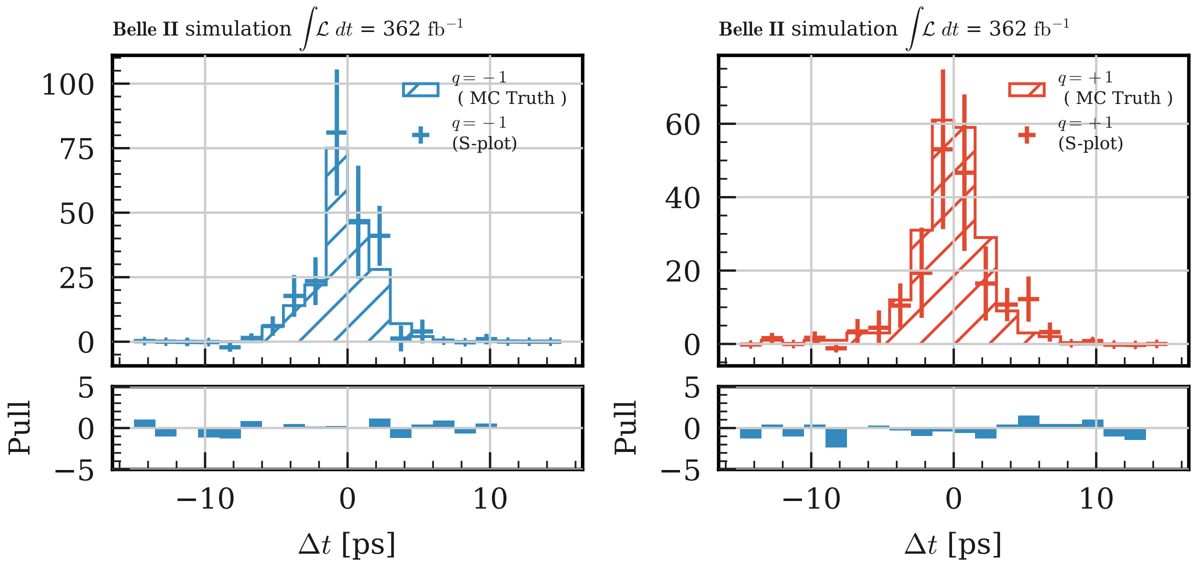


FIG. 59. Comparison of MC distribution and S-plot obtained from the signal extraction fitting to simulated samples corresponds to $362 fb^{-1}$

5.7. Fitter validation

5.7.1. Bootstrap study

A bootstrap study is conducted to validate the CP fitter. The dataset and the sampling method in this study are the same as in Section 4.7.1. To avoid the overestimation of the fit bias, we performed the study while changing the sign of qr randomly for qq and BB samples. It is because they have a qr asymmetry within the statistical uncertainty and it can make a bias in CP parameters. The details are discussed in Section 6 Figure 60 and Figure 61 show the results of the bootstrap studies. Table 33 lists the values of the bootstrap results.

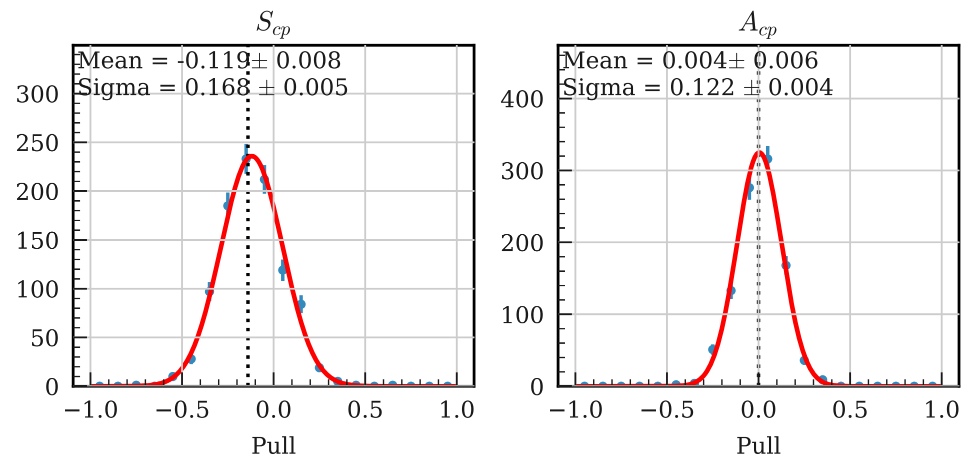


FIG. 60. Fitted S_{CP} and A_{cp} in the bootstrap study using simulated samples. The statistics of each sample correspond to 362 fb^{-1}. The dashed lines represent PDG values of S_{CP} and A_{cp}.

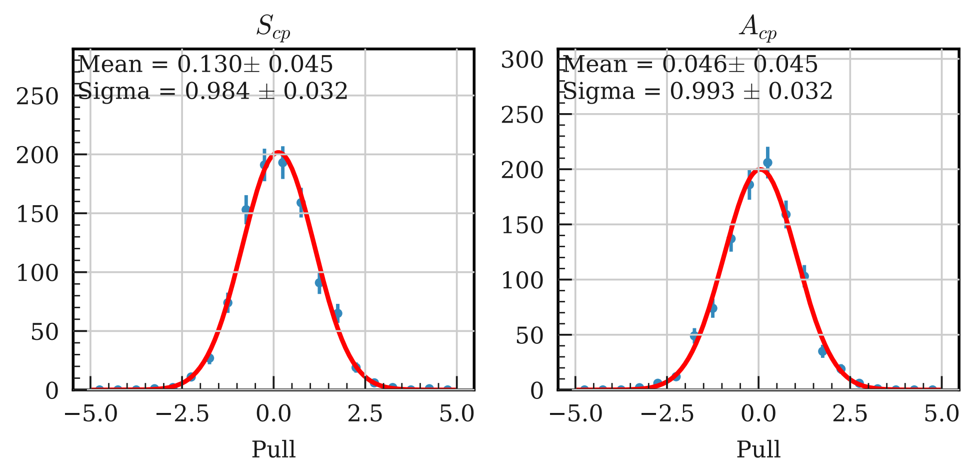


FIG. 61. Result of the Bootstrap study for CP fitting using simulated samples. The statistics of each sample correspond to 362 fb^{-1}

Parameter	Mean	Input	Mean (Pull)	Std. (Pull)
S_{cp}	-0.118 ± 0.005	-0.14	0.130 ± 0.045	0.984 ± 0.032
A_{cp}	0.006 ± 0.004	0.000	0.046 ± 0.045	0.993 ± 0.032

TABLE 33. Result of the Bootstrap study for CP fitting. Uncertainties come from the statistics of the number of bootstrap samples.

5.7.2. Linearity test

A linearity test was performed to check a bias depending on the CP parameters. Samples applied for this test were the same as the bootstrap study described in Section 5.7.1, with the exception of the signal samples. SignalMC was generated and used only for the longitudinal signals. Figure **??** shows the results of the linearity test. The fitted results are slightly biased due to the correlation with the peaking backgrounds, but the magnitude of the bias is acceptable in all the CP parameter regions.

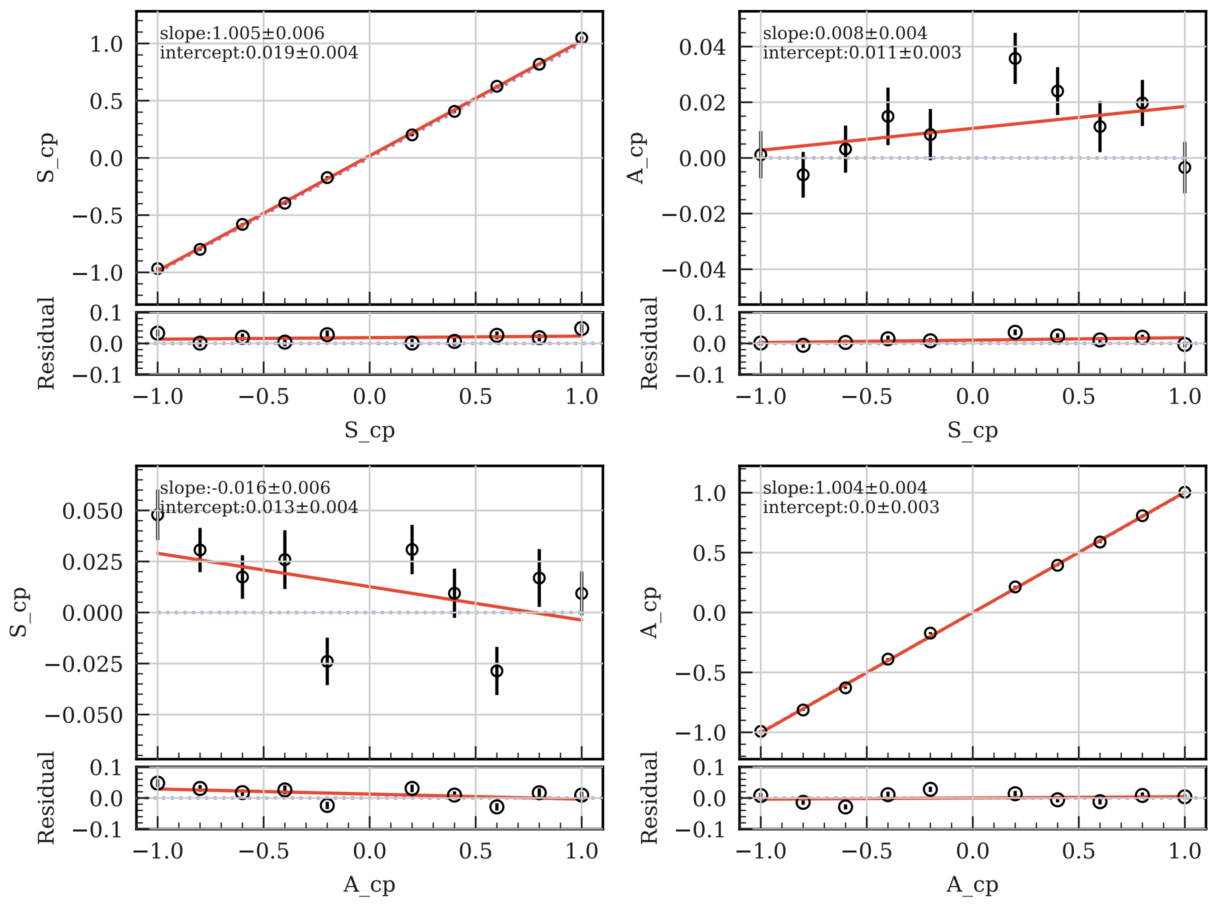


FIG. 62. Result of the linearity test for the CP fitting

As illustrated above, the methods for CP fitting are ready and are validated using MC samples.

6. ANALYSIS VALIDATION

Before the data fitting, we validated and estimated data-MC discrepancies for signal-extraction fitting and CP fitting.

6.1. ΔE and $m_{\pi^{\pm}\pi^0}$ shape

We perform Data/MC comparison by analyzing the control channel, $B^- \to D^0(\to K^-\pi^+\pi^0)\rho^-(\to \pi^-\pi^0)$. We use the simulated samples that have twice larger statistics than the data. The decays are reconstructed by applying the selections shown in Table 34 to reduce the signal significance. The applied selection of photon, π^0, $\pi^{\pm}$, and $\rho^{\pm}$ is the same as in case of $B^0 \to \rho^+\rho^-$. We apply the same corrections for tracks and photons as in the case of $B^0 \to \rho^+\rho^-$. We extract a reconstructed candidate with the smallest sum of reduced χ^2 in a collision, as given by,

$$\text{sum of reduced}\,\chi^2 = \frac{\chi^2_{\mathrm{D}^0}}{\mathrm{ndf}_{\mathrm{D}^0}} + \sum_{\pi^0_{\mathrm{D}^0},\,\pi^0_{\rho^{\pm}}} \frac{\chi^2_{\pi^0}}{\mathrm{ndf}_{\pi^0}}.$$

We check the distributions of the fitting variables as shown in Figure 63. In the ΔE and $m_{\pi^{\pm}\pi^0}$ distributions, the Data-MC discrepancies are clearly seen. Here, we investigate the data-MC discrepancy for any shifts and/or variations in the distributions of each variable. The Data-MC discrepancies of CSMVA output are discussed in Section 6.3.

The assessment of the Data-MC discrepancy was carried out through a systematic approach. Initially, MC samples were employed to model the distributions of ΔE, $m_{\pi^{\pm}\pi^0}$, and transformed CSMVA output for $D^0\rho$, self-crossfeed, $B\bar{B}$, and continuum samples. The fit models are summarized in Table 35. We verified the feasibility of accurately estimating the yields via a simultaneous fit of the three variables using MC samples. The fit results are shown in Figure 64 and Table 36. The self-crossfeed ratio was determined from MC truth information and was fixed in the fit. Next, we introduced shift and fudge factors that alter the shapes of ΔE and $m_{\pi^{\pm}\pi^0}$. For all Gaussian functions included in the models of ΔE, a common fudge factor and shift width are applied as follows: $\mu + w_{\Delta E}$ and $\sigma \times k_{\Delta E}$.

TABLE 34. Summary table of the selection for charged kaons, neutral D mesons, and charged B mesons in reconstruction.

target particle	selection and fitting
$K^{\pm}$	all
	$\lvert dr \rvert$ <0.5 cm
	$\lvert dz \rvert$ <2 cm
	theta in CDC acceptance
	chiProb>0
	PID(Binary)>0.6
D^0	1.8 GeV< M <1.95 GeV/c^2
	mass-constraint kFit
$B^{\pm}$	M_{bc} >5.274 GeV/c^2
	$\lvert\Delta E\rvert$ <0.15 GeV

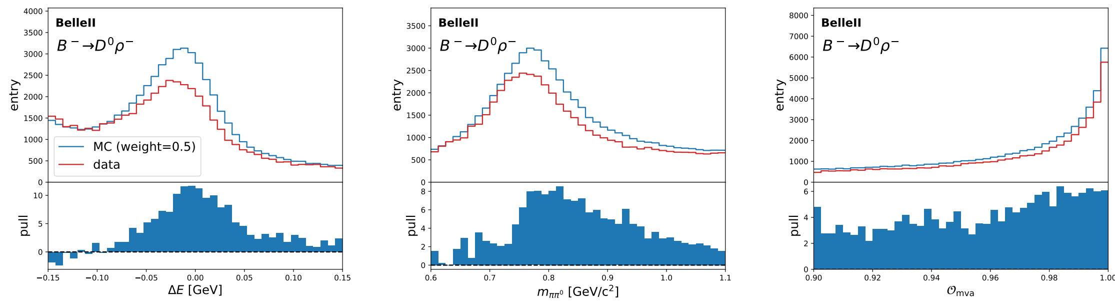


FIG. 63. Data-MC differences in $B \to D^0\rho$. The statistics of the MC samples are twice as large as that of the data, therefore it's scaled.

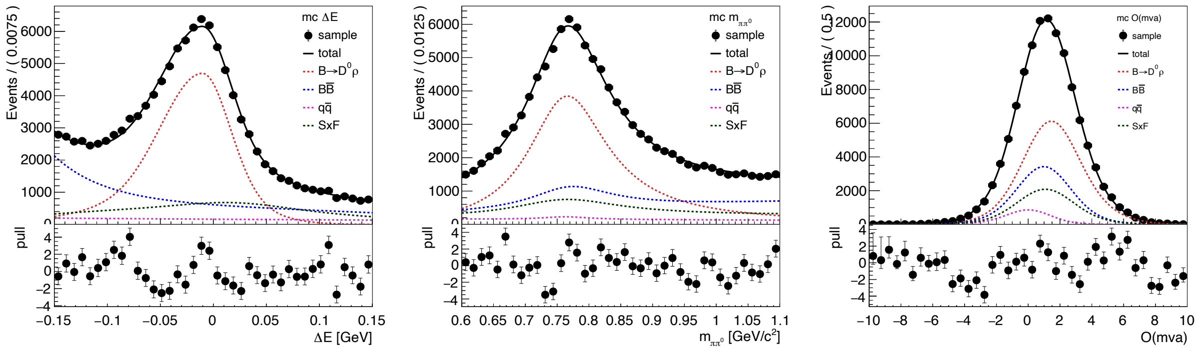
TABLE 35. Summary of fit models for each event type in the control study. bG: bifurcated-Gaussian, BW: Breit-Wigner formula, L: Linear function, and d: double.

	ΔE	$m_{\pi^\pm\pi^0}$	CSMVA
$D^0\rho$	dbG	BW	dG
SxF	dG	BW+L	dG
$q\bar{q}$	$\exp\{(ax)\} + bx + c$	BW+L	dG
$B\bar{B}$	$\exp\{(ax)\}$	BW+L	dG

Here, $w_{\Delta E}$ represents the shift width, $k_{\Delta E}$ is the fudge factor, and μ and σ are the mean and standard deviation of the Gaussian. For all Breit-Wigner formulas, a common fudge factor and shift width are also applied as follows: $m_0 + w_{m_{\pi\pi^0}}$ and $\Gamma_0 \times k_{m_{\pi\pi^0}}$. Here, $w_{m_{\pi\pi^0}}$ represents the shift width, $k_{m_{\pi\pi^0}}$ is the fudge factor, m_0 is the resonance mass, and Γ_0 is the nominal width. μ and k, along with the yields, were treated as floating parameters in fits. We determined the fudge factors and shift widths through a simultaneous fit against the data samples. The fit results are summarized in Figure 65 and Table 37. We concluded that $w_{\Delta E}$ is (-7.60 ± 0.46) MeV, $k_{\Delta E}$ is 1.141 ± 0.015, w_{m_ρ} is (-9.82 ± 0.67) MeV/c^2, and k_{m_ρ} is 1.025 ± 0.014.

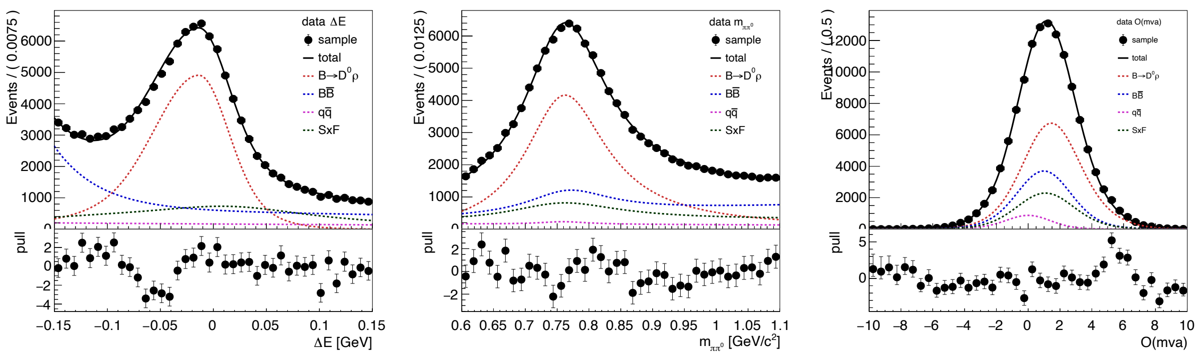
TABLE 36. Summary of MC fit results.

Parameter	Fit result	MC truth
$N_{D^0\rho}$	59025.8 ± 303.4	59102
$N_{q\bar{q}}$	6519.53 ± 387.2	5066
$N_{B\bar{B}}$	30803.4 ± 466.8	32156



(a) ΔE (b) $m_{\pi^\pm\pi^0}$ (b) log-transformed CSMVA

FIG. 64. Simultaneous fit results in the run-dependent simulated samples.



(a) ΔE (b) $m_{\pi^\pm\pi^0}$ (b) log-transformed CSMVA

FIG. 65. Simultaneous fit results in the data.

TABLE 37. Summary of data fit results. The units of $w_{\Delta E}$ and w_{m_ρ} are MeV and MeV/c^2, respectively.

Parameter	Fit result
$N_{D^0\rho}$	24779.8 ± 380.3
$N_{q\bar{q}}$	2063.3 ± 316.6
$N_{B\bar{B}}$	14090.9 ± 640.0
$w_{\Delta E}$	-7.60 ± 0.46
$k_{\Delta E}$	1.141 ± 0.015
w_{m_ρ}	-9.82 ± 0.67
k_{m_ρ}	1.025 ± 0.014

6.2. Resolution function

6.2.1. *Fitting procedure*

$B^0 \to D^{*-}\pi^+$ is a control mode for the resolution function utilized in the $B^0 \to \rho^+\rho^-$ analysis. According to Figure 66, the MC-based Δt resolution for $B^0 \to \rho^+\rho^-$ exhibits similarity to that for $B^0 \to D^*\pi$. To validate the shape of the resolution function, we performed a lifetime fitting to $B^0 \to D^{*-}\pi^+$ samples. The fit parameters are the B^0-lifetime and resolution-function shape parameters. The PDF applied in the fitting is identical to that specified in Equation 33. Table 38 and Figure 68 show the fit results to $B^0 \to D^{*-}\pi^+$ signalMC samples, indicating successful implementation of the fitter. Figure 67 shows the comparison between the fitted results to $B^0 \to D^{*-}\pi^+$ and the $\Delta t - \Delta t_{\rm MC}$ fit results to $B^0 \to \rho^+\rho^-$ or MC truth. The fitted results are in agreement with the truth values, suggesting this mode is suitable for use as a control mode for the resolution function.

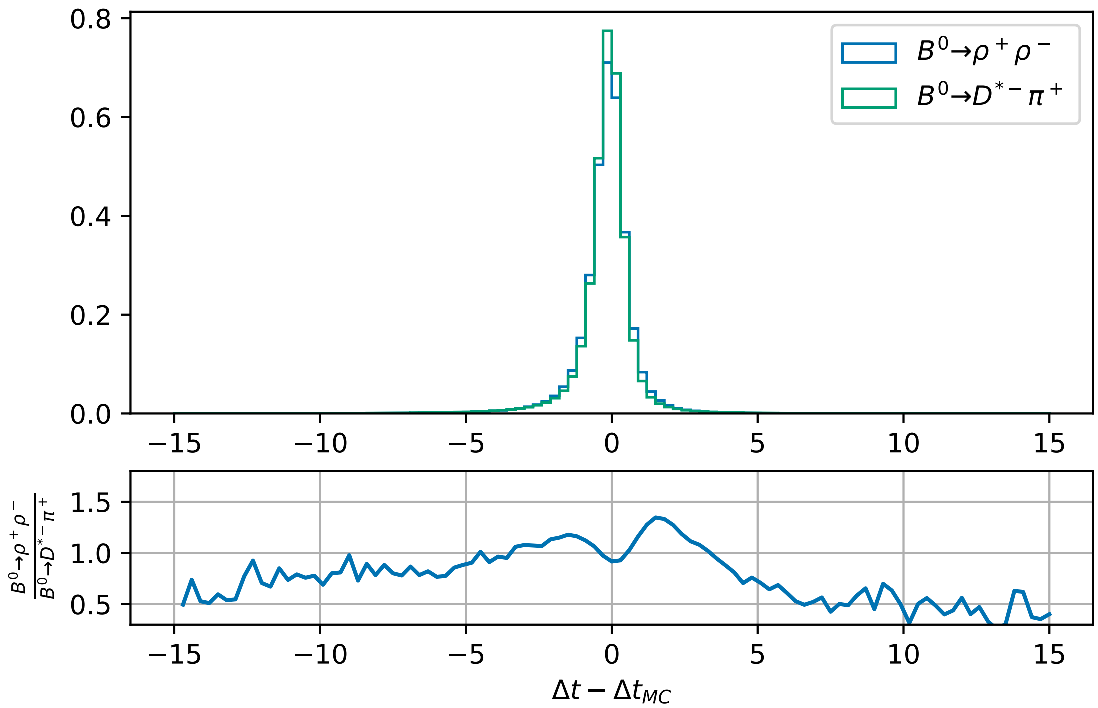


FIG. 66. Comparison of the Δt resolution between $B^0 \to \rho^+\rho^-$ signal MC and $B^0 \to D^*\pi$ signal MC

Parameter	Fit result	MC truth or $\Delta t - \Delta t_{\mathrm{MC}} fitresult$	Floated in the fit
τ	1.522 ± 0.008	1.519	Yes
f_{OL}	0.0015	0.0015	No
c	4.72	4.72	No
f_{R}	0.23	0.23	No
f_{exp}	0.19	0.19	No
μ_{main}	-0.04 ± 0.06	-0.14	Yes
μ_{tail}	-1.17 ± 0.13	-0.68	Yes
s_{main}	1.17 ± 0.05	1.15	Yes
s_{tail}	2.33 ± 0.19	2.0	Yes
σ_{OL}	200.0	200.0	No
c_6	4.94	4.94	No
f_{R6}	0.26	0.26	No
f_{exp6}	0.11	0.11	No
μ_{main6}	-0.02 ± 0.12	-0.08	Yes
μ_{tail6}	-0.51 ± 0.27	-0.34	Yes
f_{μ}	0.34	0.34	No
$f_{\sigma R}$	0.24	0.24	No
$f_{\sigma L}$	0.12	0.12	No
f_{max}	0.12	0.12	No
$C_{f_{\mathrm{tail}}}$	0.26	0.26	No
N	100000.13 ± 316.23	100000	Yes

TABLE 38. Δt resolution function fit result to the simulated samples of $B^0 \to D^{*-}\pi^+$.

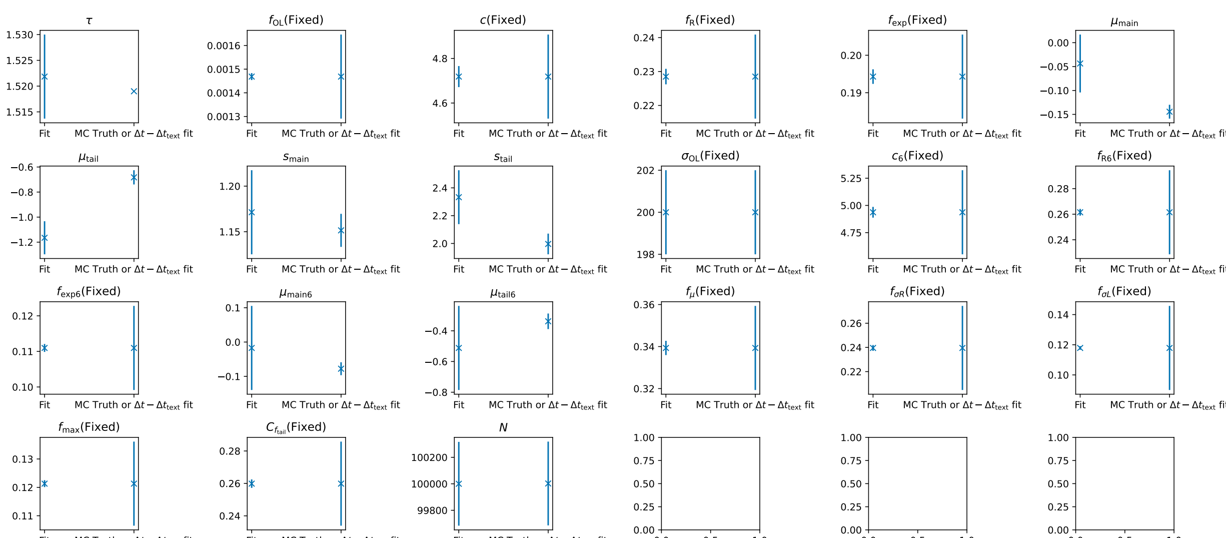


FIG. 67. Comparison between fitted and MC truth resolution function parameters.

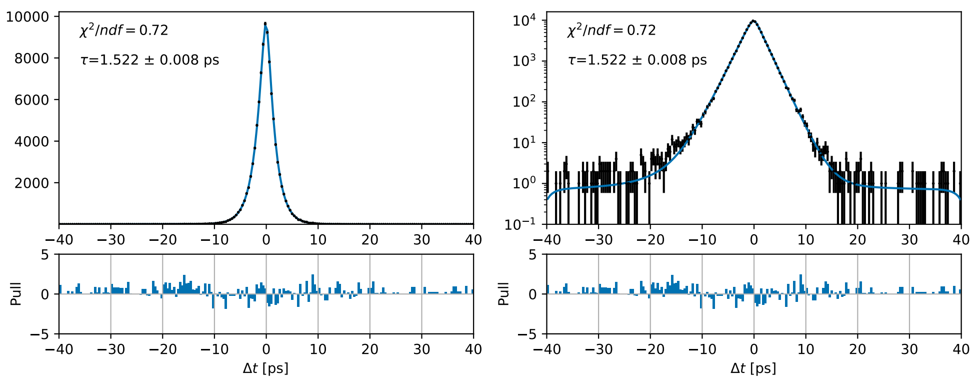


FIG. 68. Δt resolution function fit projection to MC of $B^0 \to D^{*-}\pi^+$ samples.

ΔE was used to extract the $B^0 \to D^{*-}\pi^+$ signals from backgrounds. Table 39 lists the expected event counts for each component in the fitting regions. Figure 69 shows the distributions of the variables in those regions. The PDFs for ΔE, correspoding to both the signal and background components, are summarized in Figure 70. The signal, self-crossfeed, and $B^0 \to D^{*-}K^+$ are modelled using triple bifurcated Gaussians. $B\bar{B}$ is modelled with a combination of a first-order Chebyshev polynomial and a bifurcated Gaussian. $q\bar{q}$ is modelled with a first-order Chebyshev polynomial. The PDF for Δt associated with $q\bar{q}$ and $B\bar{B}$ is the same as the PDF in the $B^0 \to \rho^+\rho^-$ signal region. The PDF for the signal, self-crossfeed, and $B^0 \to D^{*-}K^+$ is the same as the PDF for $B^0 \to \rho^+\rho^-$, excluding CP-violating effect. The floating parameters in the fitting are the lifetime of B^0, the shape parameters of the resolution function, and the number of signals, $q\bar{q}$, and $B\bar{B}$ events. The number of the $B^0 \to D^{*-}K^+$ and the ratio of the self-crossfeed events to the signals are fixed to the MC expectation.

Components	Expected number of events
Signal	4987.5
Self-crossfeed	119.7
BB	359.7
qq	13.6
$B^0 \to D^{*-}K^+$	91.8

TABLE 39. Expected number of events at $B^0 \to D^{*-}\pi^+$ fit region in the run-dependent simulated samples.

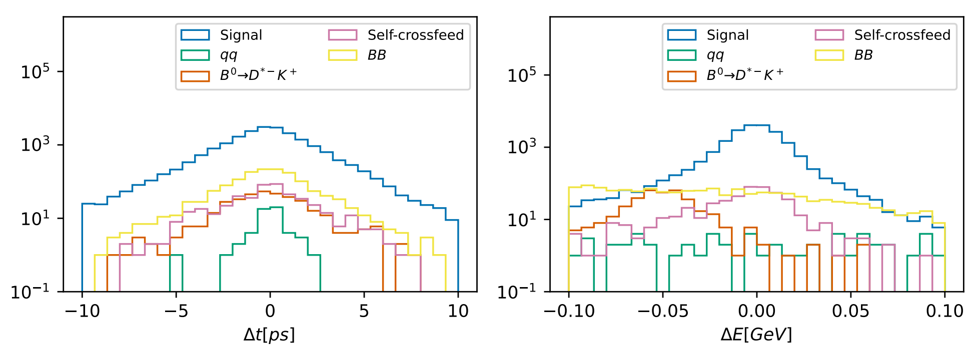


FIG. 69. Distribution of the fit variables in the fitting region in the run-dependent simulated samples

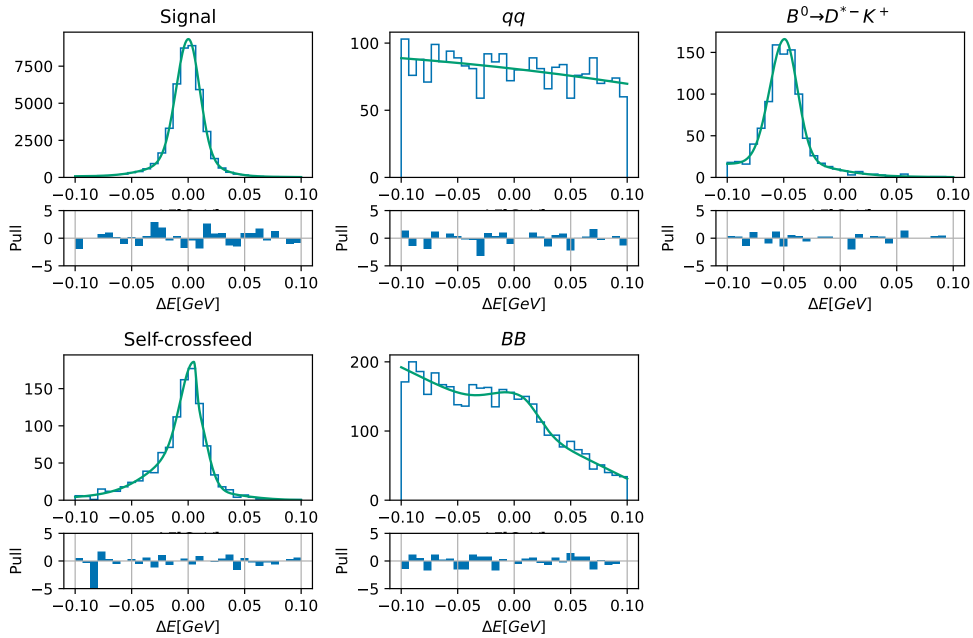


FIG. 70. ΔE PDF for $B^0 \to D^{*-}\pi^+$ and its backgrounds in the run-dependent simulated samples

6.2.2. *Fitting to the data*

Figure 71 shows the fit results for the $B^0 \to D^{*-}\pi^+$ preselected dataset, corresponding to an integrated luminosity of 362 fb^{-1}. Table 40 and Table 41 present the yields and resolution parameters obtained from the fit, respectively. These yields and parameters are consistent with the MC expectations or the fit results to MC truth values. A systematic uncertainty arising from this is discussed in Section 8.

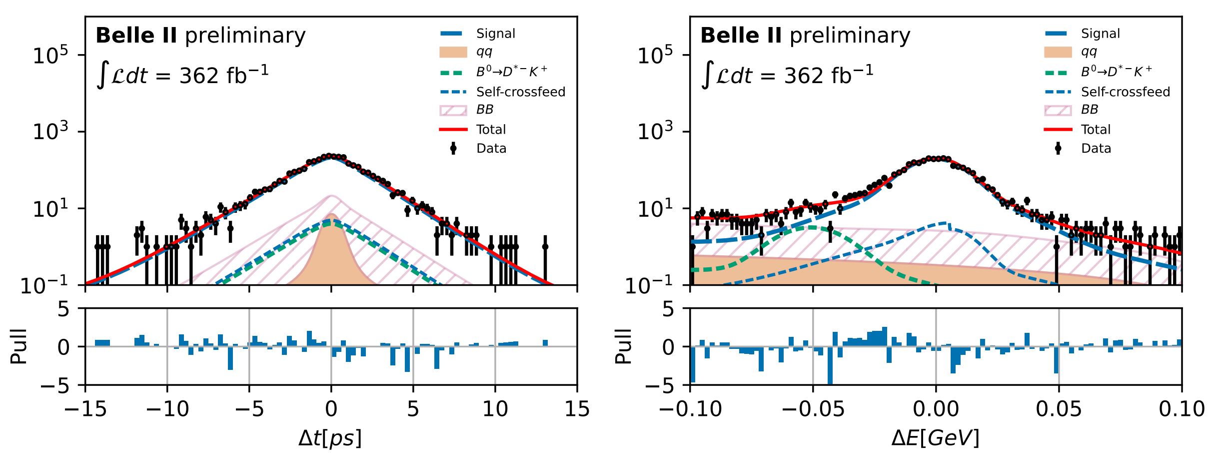


FIG. 71. Result of the fitting to the $B^0 \to D^{*-}\pi^+$ data

Parameter	Fit results	MC expectation
N_{Signal}	2987.7 ± 58.8	3234.3
$N_{q\bar{q}}$	33.0 ± 17.6	3.0
$N_{B\bar{B}}$	201.9 ± 33.4	218.0

TABLE 40. Fitted yield of the signals and backgrounds in $B^0 \to D^{*-}\pi^+$ fitting to data.

Parameter	Fit result	MC truth or $\Delta t - \Delta t_{\mathrm{MC}} fitresult$	Floated in the fit
τ	1.515 ± 0.038	1.519	Yes
f_{OL}	0.0017	0.0017	No
c	4.25	4.25	No
f_{R}	0.23	0.23	No
f_{exp}	0.21	0.21	No
μ_{main}	-0.03 ± 0.11	-0.14	Yes
μ_{tail}	-0.67 ± 0.25	-0.67	Yes
s_{main}	1.13 ± 0.2	1.15	Yes
s_{tail}	3.26 ± 0.5	1.91	Yes
σ_{OL}	200.0	200.0	No
c_6	4.66	4.66	No
f_{R6}	0.24	0.24	No
f_{exp6}	0.11	0.11	No
μ_{main6}	0.1 ± 0.28	-0.09	Yes
μ_{tail6}	-1.43 ± 0.71	-0.3	Yes
f_{μ}	0.33	0.33	No
$f_{\sigma R}$	0.27	0.27	No
$f_{\sigma L}$	0.12	0.12	No
f_{max}	0.14	0.14	No
$C_{f_{\mathrm{tail}}}$	0.26	0.26	No

TABLE 41. Δt resolution function fit result to $B^0 \to D^{*-}\pi^+$ data.

6.3. Continuum suppression efficiency

6.3.1. Fitting procedure

To validate the signal efficiency of the continuum suppression, we measured the yields of $B^0 \to D^{*-}\pi^+$ with and without the continuum suppression. The fitting with a continuum-suppression cut is detailed in Section 6.2.1, and the fitting approach without the continuum-suppression cut and the data-MC ratio of CSMVA efficiency are covered in this section. To extract the $B^0 \to D^{*-}\pi^+$ yield, we performed a fit applying the PDF parameterizations for ΔE the ΔE, which are consistent with those explained in Section 6.2.1 while determining the parameters by fitting to the run-dependent simulated samples without the CSMVA selection. Figure 72 shows the modeling of ΔE distribution for each component.

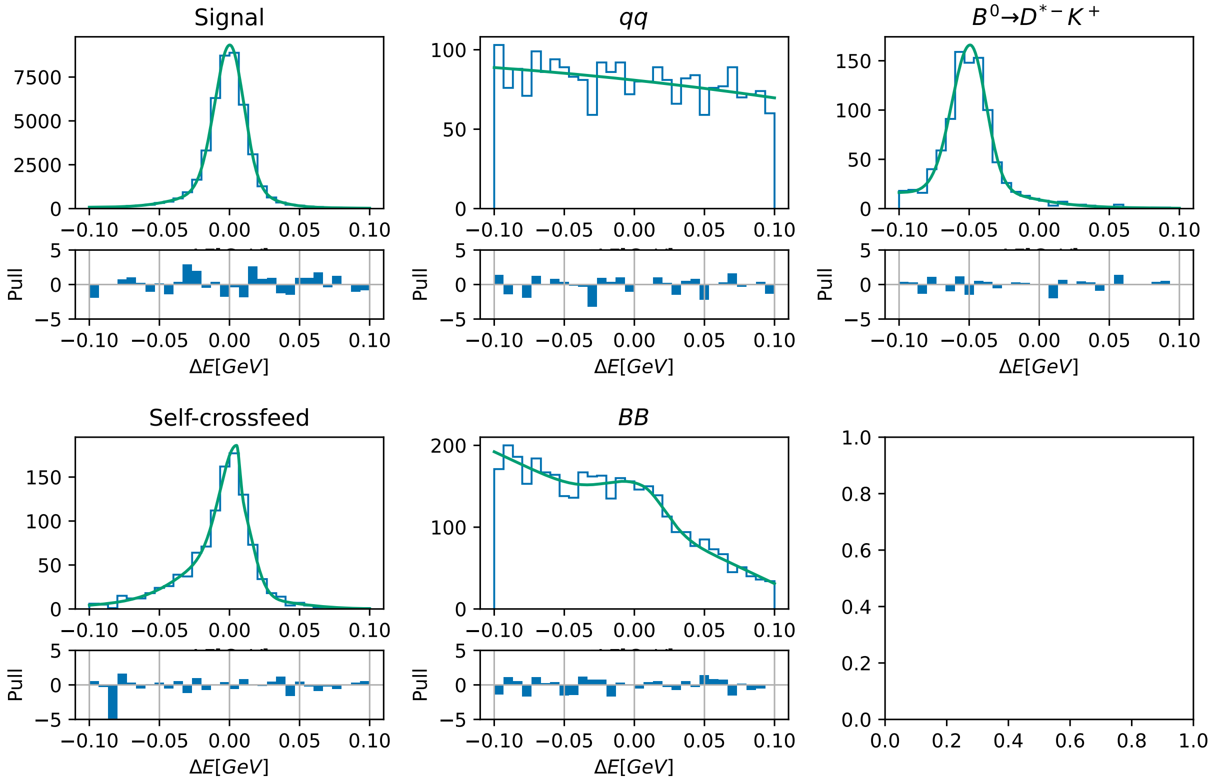


FIG. 72. ΔE PDF for $B^0 \to D^{*-}\pi^+$ and backgrounds without CSMVA selection in the run-dependent simulated samples

6.3.2. *Fitting to the data*

Figure 73 illustrates the fit results, and Table 42 lists the fitted yields without the CSMVA selection. Using the results shown in Table 40 and Table 42, the Data-MC ratio of the CSMVA efficiency is measured to be 1.08 ± 0.03. We corrected the signal efficiency with the data-MC ratio, and calculated the systematic uncertainty using the statistical uncertainty of the data-MC ratio.

Parameter	Fit results	MC expectation
N_{Signal}	9448.0 ± 110.1	11074.8
$N_{q\bar{q}}$	1032.1 ± 119.0	605.8
$N_{B\bar{B}}$	701.6 ± 135.9	926.6

TABLE 42. The fitted yields of the signals and backgrounds in $B^0 \to D^{*-}\pi^+$ fitting without CSMVA selection in the data.

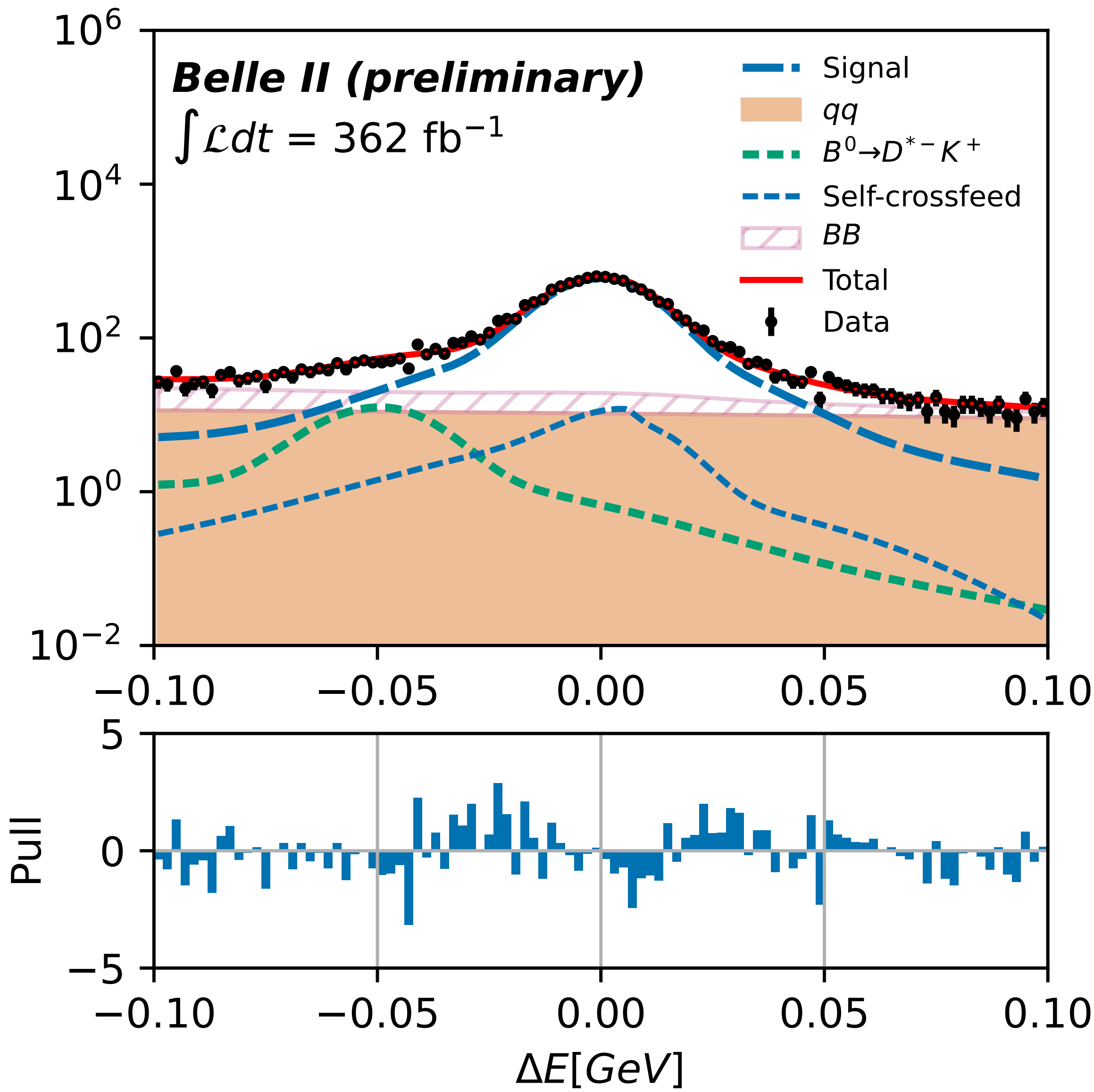


FIG. 73. Result of the fitting to the $B^0 \to D^{*-}\pi^+$ data without CSMVA selection

6.4. Background Δt PDF

The off-resonance data and the data in an $M_{\rm bc}$ sideband region are utilized to validate the shape of the Δt PDF for $B\bar{B}$ and $q\bar{q}$. Here, the sideband region is defined as $5.24 < M_{\rm bc} < 5.26$. Figure 74 compares the distributions of Δt and qr bins between the signal box and the sideband region using run-dependent simulated samples. The shape in the sideband region is similar to that in the signal box. It indicates that the sideband data is useful for the validation. The fit results to the Δt distribution of the off-resonance data are presented in Figure 75. The fitted values listed in Table 43 exhibit good agreement with expectations.

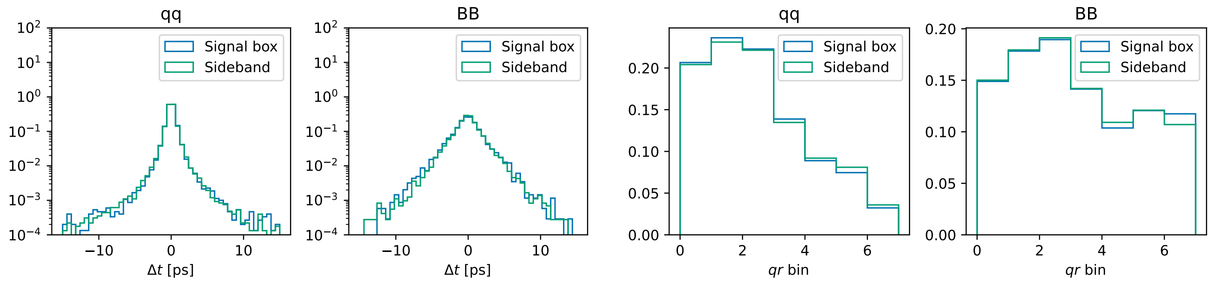


FIG. 74. Comparison of Δt and qr bin distribution of qq and BB background in Signal box and sideband region using the run-dependent simulated samples

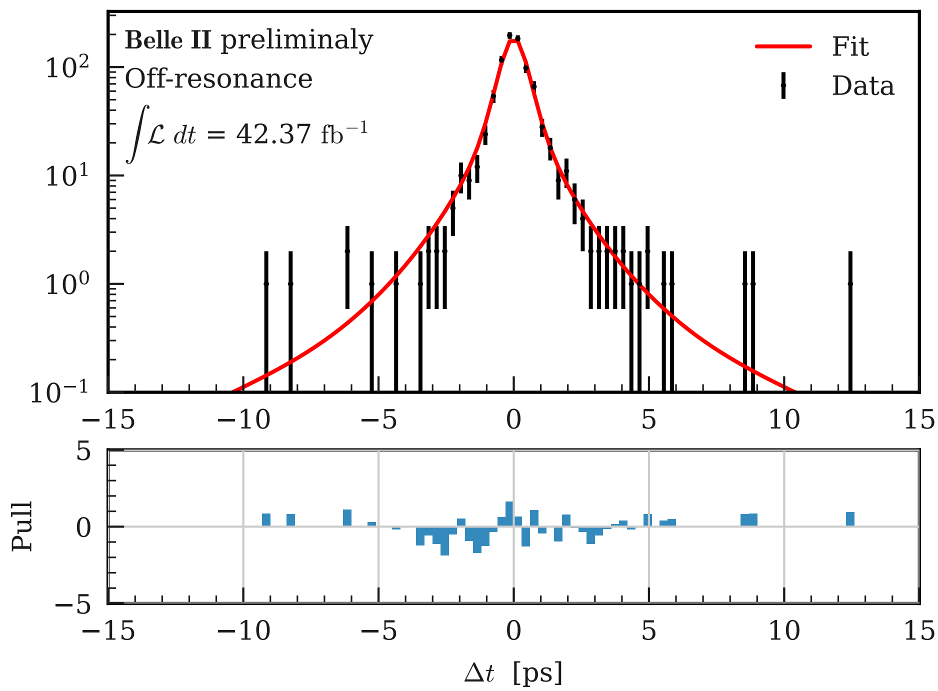


FIG. 75. Fitting to the off-resonance data

Parameter	Fit result	MC expectation
μ	0.003 ± 0.017	0.003
f_{tail}	0.1 ± 0.02	0.051
s_{main}	1.312 ± 0.052	1.382
s_{tail}	5.702 ± 0.641	9.525

TABLE 43. Result of the fitting to off-resonance data

Validation of the shape of the Δt PDF for $B\bar{B}$ was performed using the M_{bc} sideband data. The fit for this validation was conducted with Δt and ΔE as the variables. As in the $B^0 \to \rho^+\rho^-$ fitting, the fitting processes are split into two steps: signal-extraction fitting and Δt fitting. The Δt PDF shape for $q\bar{q}$ samples is fixed to the results of the fitting to the off-resonance data. Floating parameters are the number of $q\bar{q}$ and $B\bar{B}$ events and the effective lifetime of $B\bar{B}$ backgrounds. Figure 76 and Table 44 show the fit results. The systematic uncertainty in the effective lifetime due to the uncertainty of the off-resonance fitting is included. All the parameters show a good agreement with the expectations from MC simulations. The systematic uncertainty related to the precision of the $q\bar{q}$ and $B\bar{B}$ Δt PDF shape will be discussed later.

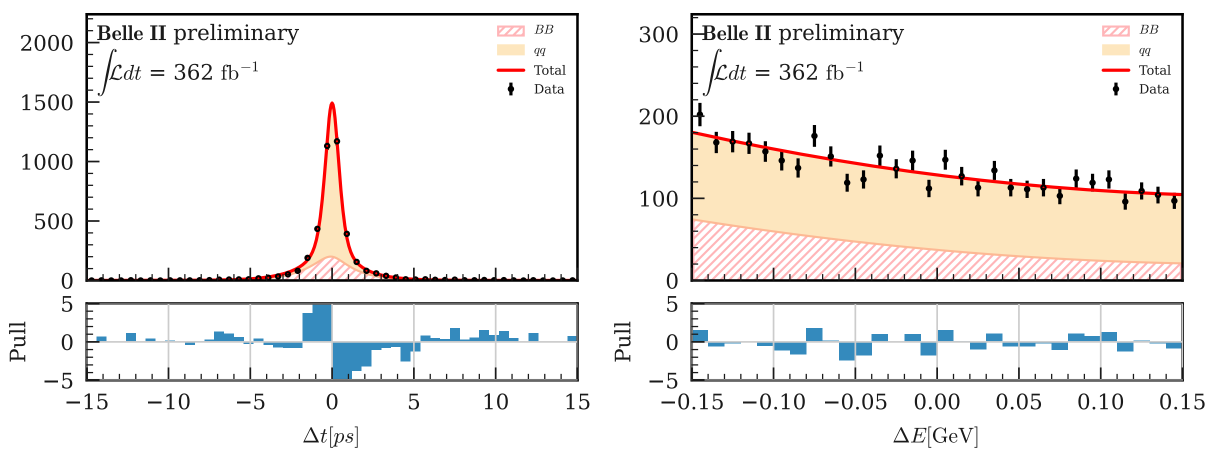


FIG. 76. Fit projection to the M_{bc} sideband data

Parameter	Fit result	MC expectation	Sample
τ_{BB}	1.289 ± 0.173	1.414	Sideband
μ	0.003 ± 0.017	0.003	Off-resonance
f_{tail}	0.1 ± 0.02	0.051	Off-resonance
s_{main}	1.312 ± 0.052	1.382	Off-resonance
s_{tail}	5.702 ± 0.641	9.525	Off-resonance
N_{qq}	2771.8 ± 199.5	2552.3	Sideband
N_{BB}	1222.2 ± 195.8	1305.7	Sideband

TABLE 44. Result of the fitting to M_{bc} sideband and off-resonance data

6.5. $\cos\theta_\rho$ PDF shape

6.5.1. $\cos\theta_\rho$ PDF shape for qq and BB backgrounds

The $\cos\theta_\rho$ PDF shape was validated by utilizing samples in $M_{\rm bc}$ sideband region defined above. The Data-MC discrepancy in $\cos\theta_\rho$ is appreciable, as illustrated in Figure 77. Similarly, large discrepancies in P_π and P_{π^0} appear in Figure 78. Both discrepancies influence $cos\theta_\rho$. Compared to the discrepancy in P_{π^0}, the discrepancy in P_π is more pronounced while it was not found in $B^0 \to D^{*-}\pi$, as shown in Figure 79. Thus, it is considered that the Data-MC discrepancy is due to the mis-modelling of $q\bar{q}$ backgrounds.

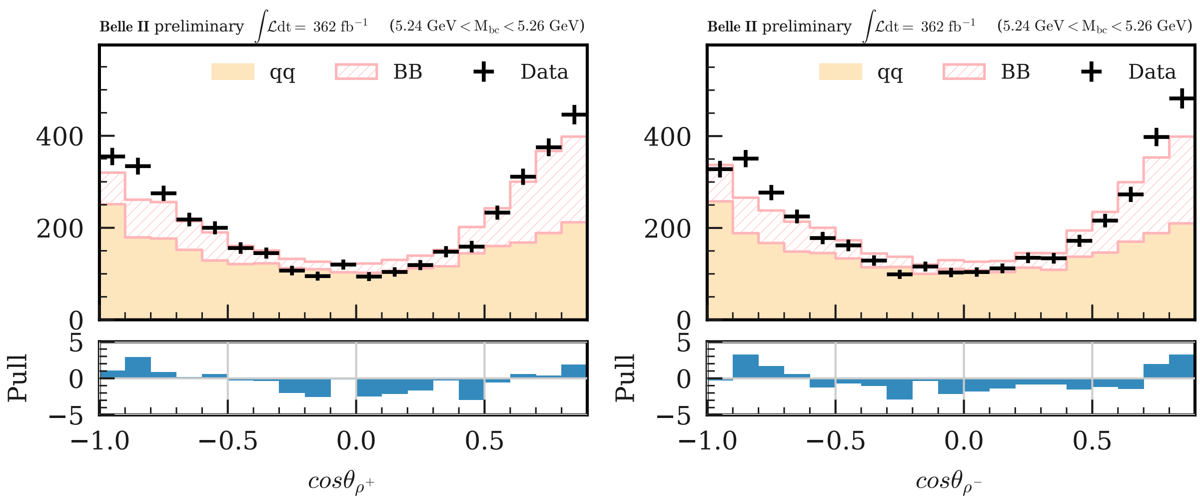


FIG. 77. Data-MC difference of $\cos\theta_\rho$ in the sideband region

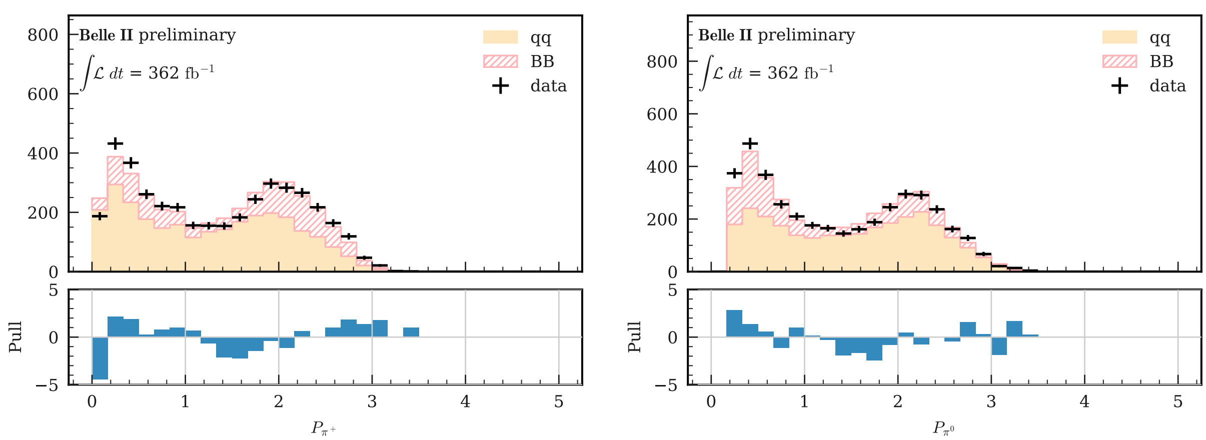


FIG. 78. Data-MC difference of P_π and P_{π^0} in the sideband region

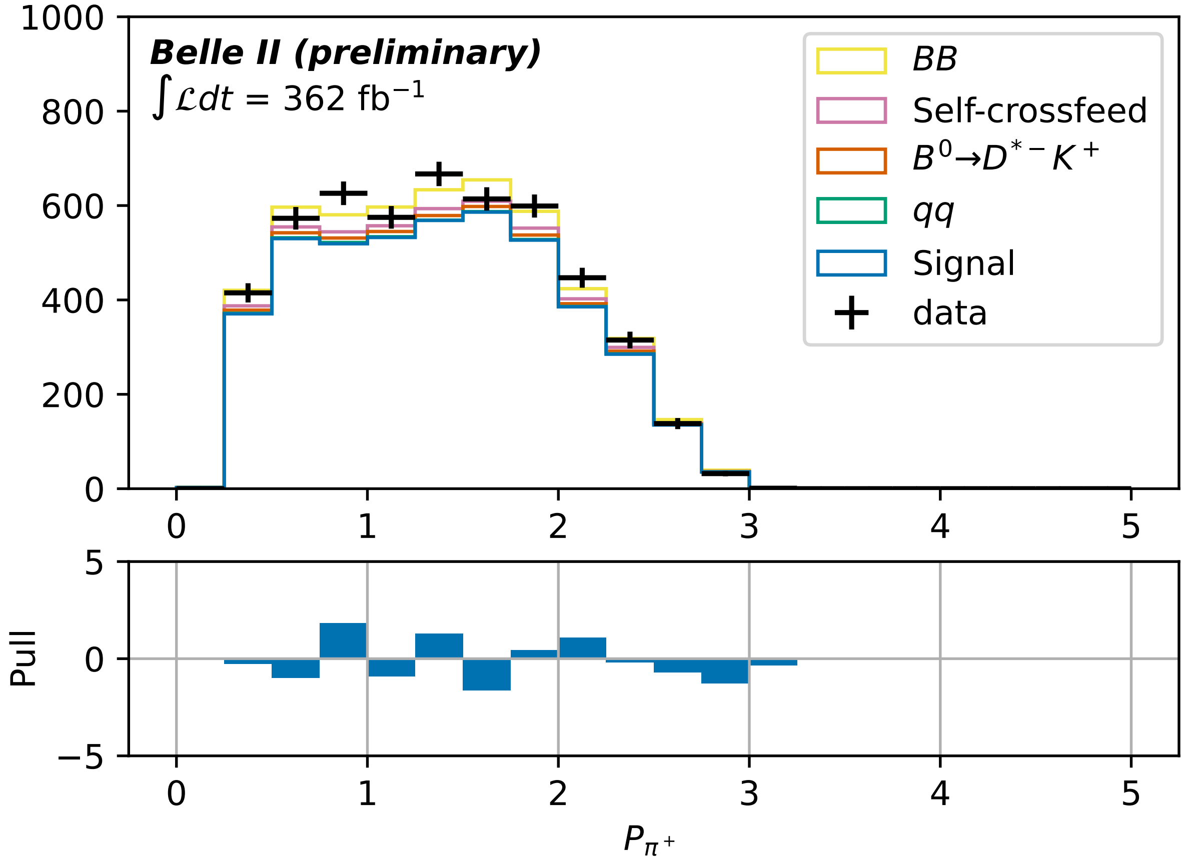


FIG. 79. Data-MC difference of P_π in the $B^0 \to D^{*-}\pi^+$ samples

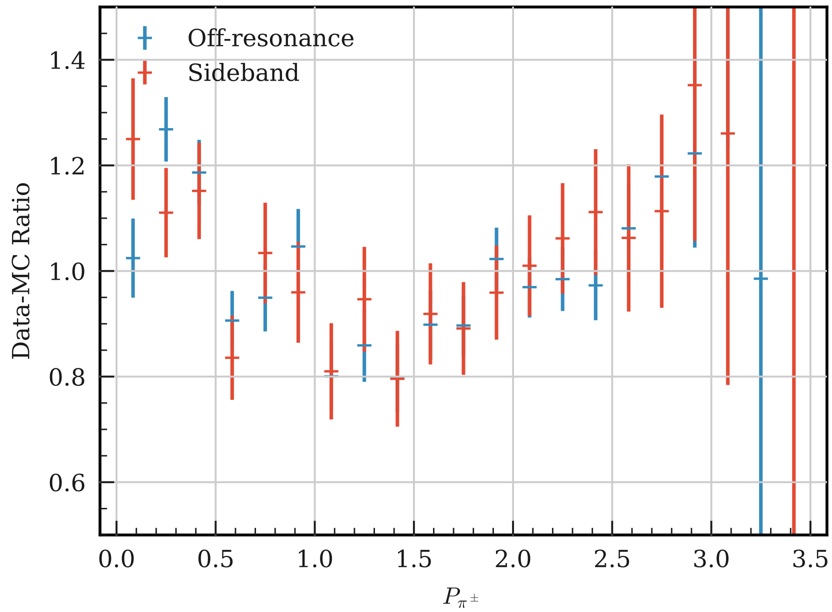


FIG. 80. Comparison of the data-MC ratio in $P_{\pi\pm}$ distribution between sideband data and off-resonance data.

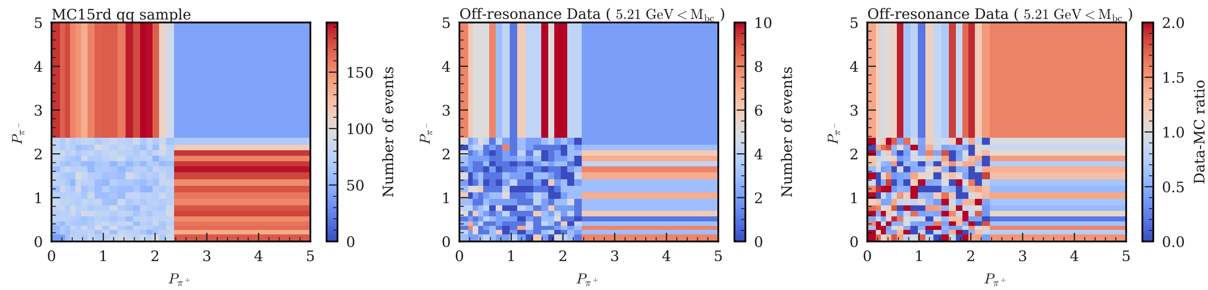


FIG. 81. Data-MC correction using 2D P_π distribution using off-resonance data

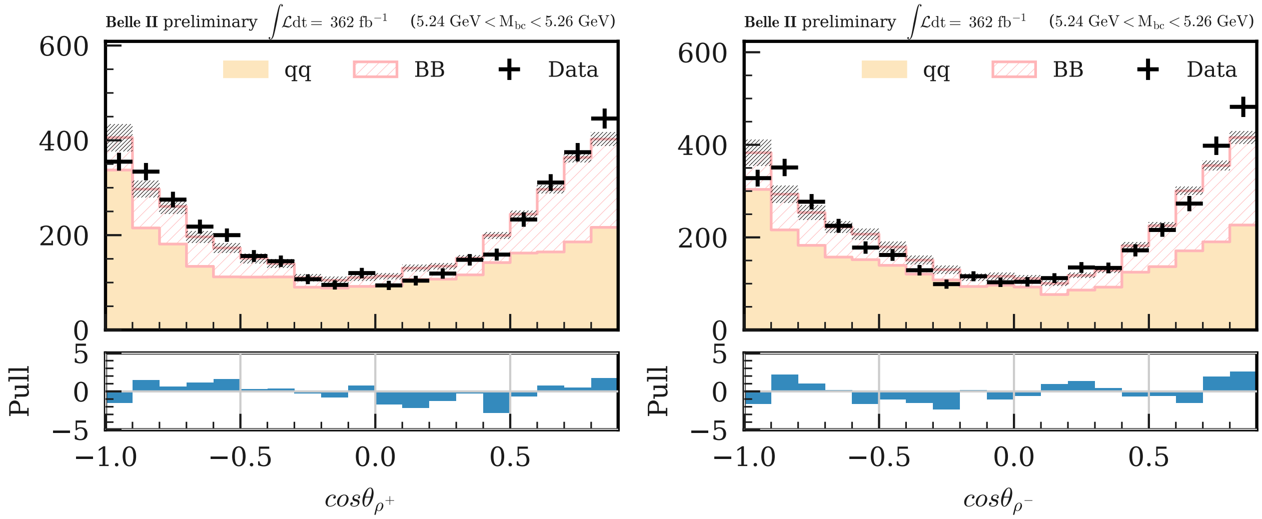


FIG. 82. Data-MC difference of $\cos\theta_\rho$ in the sideband region after the correction

Figure 80 shows a data-MC difference comparison between off-resonance data and sideband data, and both distributions demonstrate good agreement with each other. The Data-MC discrepancy in P_π can be described by the mis-modelling of $q\bar{q}$. Assuming the $B\bar{B}$ samples are accurately represented in MC simulations, we evaluated the Data-MC ratio for $q\bar{q}$ samples in the P_π distributions using off-resonance data, and Figure 81 shows the result. In consideration of the Data-MC discrepancy, we corrected $\cos\theta_\rho$-PDF shape by weighting the MC samples of $q\bar{q}$ using Data-MC ratio of P_π. The corrected $\cos\theta_\rho$ distribution, shown in Figure 82, reveals significantly enhanced concordance between the data and model predictions. The remaining Data-MC differences are included in the systematic uncertainties.

6.5.2. $\cos\theta_\rho$ PDF shape for the signal

We also evaluated the $\cos\theta_\rho$ data-MC discrepancy of the signal using $B^\pm \to D^0\rho^\pm$. The selection is the same as in 6.1. Figure 83 shows the data-MC discrepancy in $\cos\theta_\rho$ distribution. In this evaluation, the number of events is fixed by the fit result to ΔE, $m_{\pi\pi^0}$, and the transformed CSMVA outputs which are shown in Table 37. We estimated the data-MC discrepancy of the signal component by subtracting the background effects considering the data-MC discrepancy of BB and qq backgrounds shown in Section 6.5.1. Figure 84 shows the data-MC discrepancy after subtracting the background effects. There are some discrepancies, but the systematic uncertainty is acceptable. It will be discussed later.

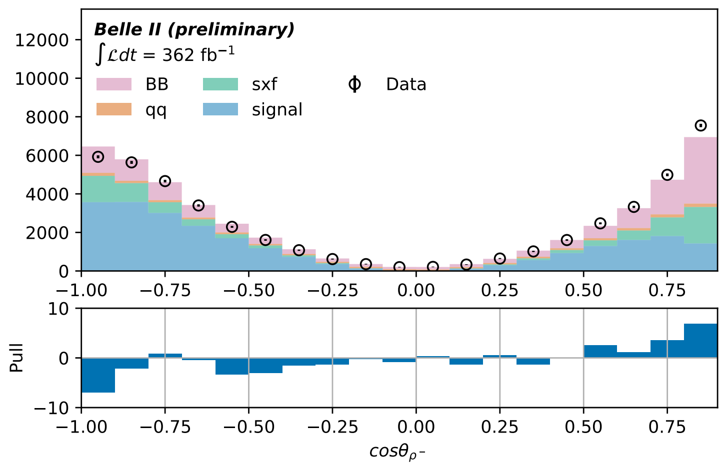


FIG. 83. data-MC discrepancy in $\cos\theta_\rho$ distribution in $B \to D^0\rho$ data.

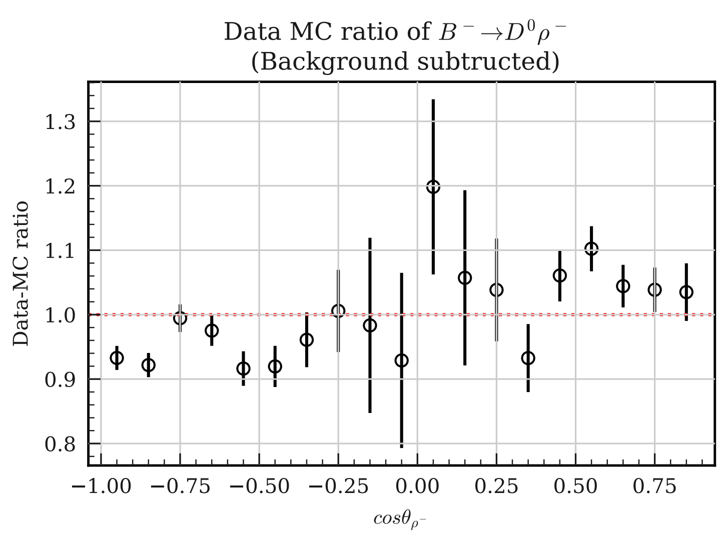


FIG. 84. data-MC ratio of $\cos\theta_\rho$ signal PDF after subtracting the background effects.

6.6. CP Fit validation using $B^0 \to D^*\pi$

6.6.1. *Mixing*

We performed mixing fit to $B^0 \to D^*\pi$ data. $B^0 \to D^*\pi$ decay is a flavor-specific decay and $B^0\bar{B}^0$ mixing can be measured by using qr and the charge of $\pi^{\pm}$. The fitting is performed to ΔE and Δt. The ΔE PDF for all components and the Δt PDF except for the signal are the same as that were used for the Δt resolution function calibration using $B^0 \to D^*\pi$ data which is shown in **??**. Equation 37 shows the Δt PDF for the signal, where q is a product of tagged B^0 flavor and the charge of the $\pi^{\pm}$ from the signal. It is the same as the $B \to \rho\rho$ signal without CPV effects but the mixing is included. The free parameters in the fitting is τ, N_{signal}, N_{qq}, N_{BB}, and Δm_d. The Δt resolution function shape parameters are fixed to the same value as Table 41.

$$\begin{aligned}
&PDF_{obs}(\Delta t, \sigma, q) = \\
&\int_{-1}^{1} d\cos\theta^* \left[\frac{3}{4}(1-\cos^2\theta^*)\right] \int_{-\infty}^{\infty} d\Delta t_{\mathrm{MC}} \int_{\frac{|\Delta t_{\mathrm{MC}}|}{1+K\cdot(sign)\Delta t_{\mathrm{MC}}\cos\theta^*}}^{\infty} d\Sigma_\tau \\
&\frac{1}{2\tau}\exp\left(\frac{-\Sigma_\tau}{\tau}\right)\mathcal{R}(\delta\Delta t, \sigma)[[1 - q\Delta\omega + q\mu(1-2\omega)] \\
&+ [q(1-2w) + \mu(1-q\Delta w)]\cos(\Delta m_d \Delta\tau)]
\end{aligned} \tag{37}$$

Table 45 shows the fit results and fitted lifetime and mixing frequency were consistent with the world average. Figure 85 shows the fit projection, and the fitted distribution agrees with the data well.

Parameter	Fit	MC expectation
τ	1.497 ± 0.027	1.519
Δm_d	0.503 ± 0.016	0.507
N_{Signal}	4499.9 ± 72.0	4987.5
$N_{q\bar{q}}$	45.8 ± 23.5	13.6
$N_{B\bar{B}}$	261.9 ± 41.1	359.7

TABLE 45. mixing fit to $B \to D^*\pi$ data

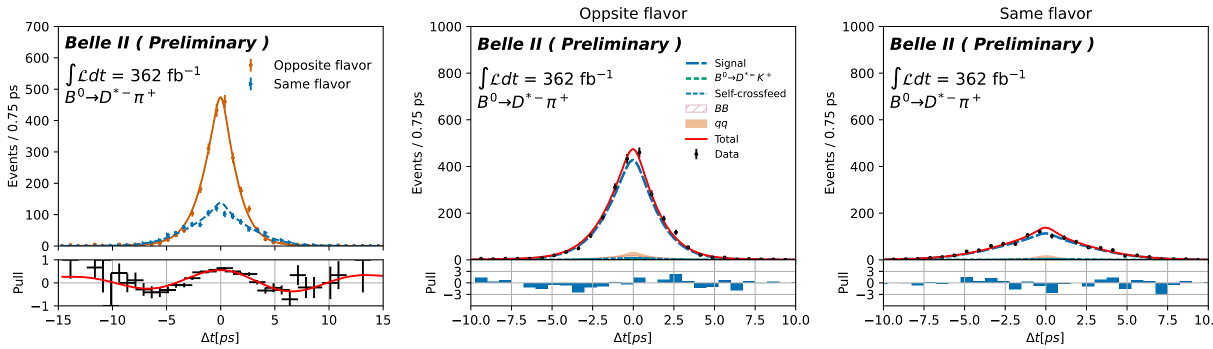


FIG. 85. Mixing fit result to $B^0 \to D^*\pi$ data. The left shows the total fit projection for the opposite and same flavor, and the middle and the right show the fit projection for each flavor.

6.6.2. CP fit

We also performed a CP fit to $B^0 \to D^*\pi$ whose CP parameter is consistent with 0 within its uncertainty to validate the Δt fit analysis procedure. The fitting procedure is almost the same as Section 6.6.1 but only 2 things are changed. Firstly, for q, we only used flavor tag information and didn't use flavor information from the signal side. Secondly, we used the same Δt PDF as $B^0 \to \rho^+\rho^-$ signal for $B^0 \to D^*\pi$. Table 46 shows the fit result, and the fitted CP parameters are consistent with 0. The fit projection is shown in Figure 86 and Figure 87.

Parameter	Fit	MC expectation
N_{Signal}	2987.8 ± 58.8	3234.3
$N_{q\bar{q}}$	33.5 ± 17.4	3.0
$N_{B\bar{B}}$	201.4 ± 33.3	218.0
$A_{\rm CP}$	0.018 ± 0.036	0.0
$S_{\rm CP}$	0.031 ± 0.053	0.0

TABLE 46. *CP fit result to $B^0 \to D^*\pi$ data.*

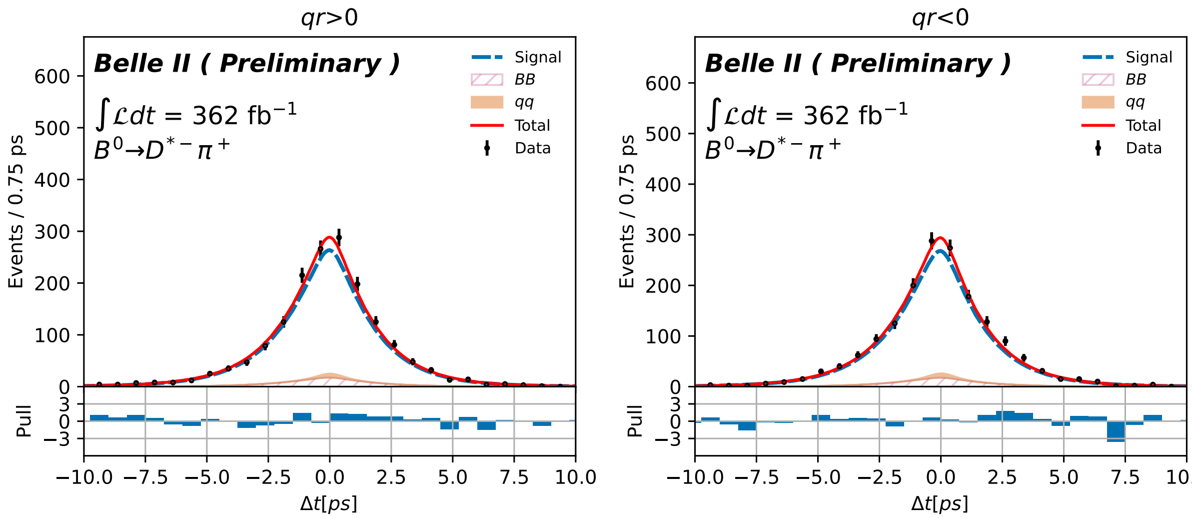


FIG. 86. Fit result to $B^0 \to D^*\pi$ data.

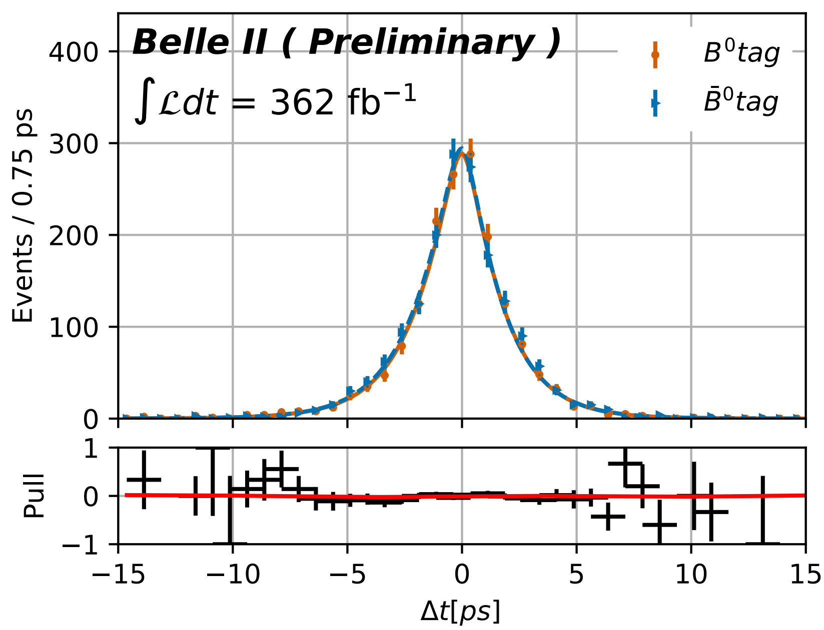


FIG. 87. Asymmetry plot of the fitting to $B^0 \to D^*\pi$ data.

6.7. *CP* Fit validation using $B^0 \to \rho^+\rho^-$

6.7.1. *Lifetime fit*

We validated Δt modeling and fitting by doing a lifetime fit to $B^0 \to \rho^+\rho^-$ samples. In the validation, the flavor tagging information was not used. all fit parameters are masked. We floated B^0 lifetime τ and only checked the lifetime. The fitted lifetime is $\tau = 1.41 \pm 0.13$ ps which is consistent with the world average. The fit projection is shown in Figure 88

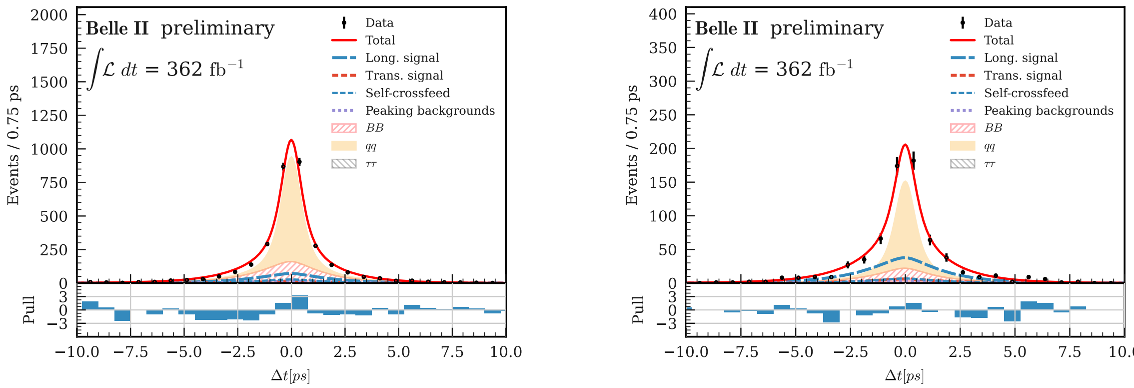


FIG. 88. The lifetime fit projection to $B^0 \to \rho^+\rho^-$ data. (Left: Total fit projection, Right: Fit projection in the signal-enhanced region.)

6.7.2. CP fit with a random tagging

We also performed a validation for the CP fit by doing a fit to $B^0 \to \rho^+\rho^-$ data with random flavor tagging while keeping $|r|$ value. Table 47 shows the result of the fitting, and the fitted CP parameters were consistent with 0. Figure 89 and Figure 90 shows the fit projection.

Parameter	Fit
S_{CP}	-0.070 ± 0.186
A_{CP}	0.079 ± 0.121

TABLE 47. CP fit to the $B^0 \to \rho^+\rho^-$ data with random flavor tagging

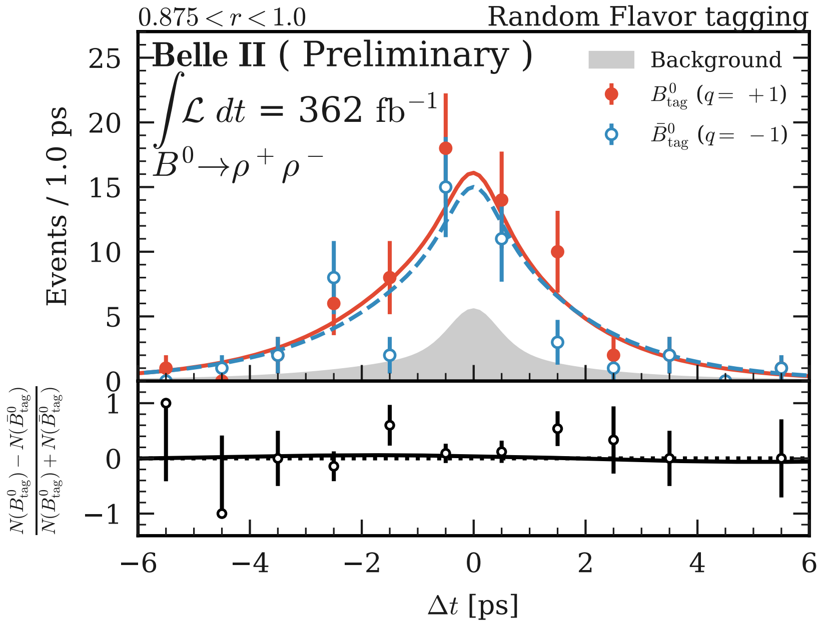


FIG. 89. CP fitting to $B^0 \to \rho^+\rho^-$ data with random flavor tagging in the signal-enhanced region.

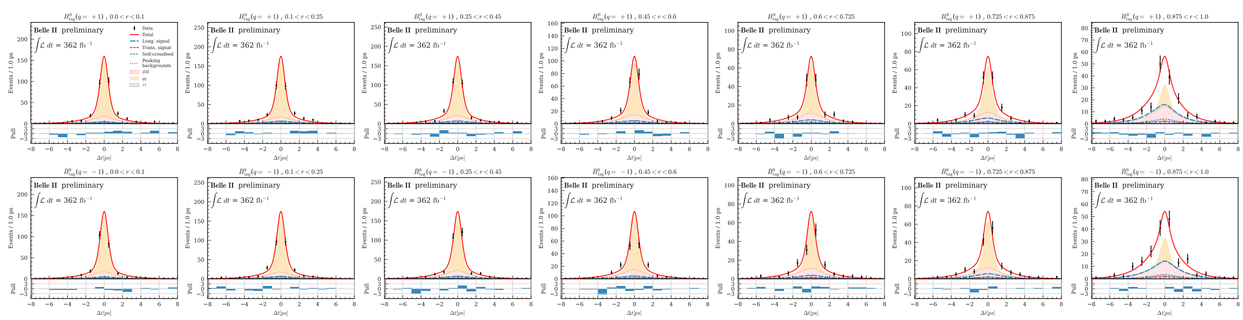


FIG. 90. Fit projection for the CP fit in the each qr bins.

7. RESULTS

We perform fit to $B^0 \to \rho^+\rho^-$ data. Figure 91 shows the fit projections to the variables for the signal-extraction fit. We obtain

$$\mathcal{B}(B^0 \to \rho^+\rho^-) = \left(2.89^{+0.23}_{-0.22}\right) \times 10^{-5}, \tag{38}$$
$$f_L = 0.921^{+0.024}_{-0.025}, \tag{39}$$

where the uncertainties are statistical only, with a statistical correlation coefficient of -0.11. Table 48 shows the yields of the signal and background yields.

We then performed a fitting to extract CP violation parameters. We obtain

$$S = -0.26 \pm 0.19, \tag{40}$$
$$C = 0.02 \pm 0.12 \tag{41}$$

where the uncertainties are statistical only, with a statistical correlation coefficient of -0.06. Figure 92 shows the fit projections in the best flavor tagging bin at $q = \pm 1$ and background-subtracted distribution.

Component	Yields
LP signal	$436.3^{+34.2}_{-33.5}$
TP signal	$65.4^{+24.3}_{-22.6}$
LP signal (SxF)	$151.0^{+11.9}_{-11.6}$
TP signal (SxF)	$5.6^{+1.9}_{-1.8}$
$q\bar{q}$	$1410.2^{+76.5}_{-75.7}$
Combinatorial $B\overline{B}$	$849.2^{+73.3}_{-72.3}$
$B^0 \to \rho^+\pi^-\pi^0$	$44.9^{+93.3}_{-91.9}$
$B^0 \to \pi^+\pi^-\pi^0\pi^0$	$-98.0^{+62.2}_{-60.2}$
$B^0 \to \pi^+\pi^-\pi^0$	$-1.2^{+18.8}_{-15.9}$
$B^0 \to a_1^0\pi^0$	$32.0^{+28.6}_{-26.5}$
"rare peaking" backgrounds	$-31.9^{+45.3}_{-42.9}$

TABLE 48. Yields of the signal and background components obtained by the signal-extraction fitting.

In the CP fitting, the event by event fraction was fixed to the fit result in the signal extraction. The CP parameter uncertainty should include the uncertainty of the statistical uncertainty of $\mathcal{B}$ or f_L in the signal extraction fit. To estimate the uncertainty, CP fit was performed while changing $\mathcal{B}$ or f_L within the uncertainty in the signal extraction fitting. Figure 93 and Table 49 show the observed bias when $\mathcal{B}$ or f_L is varied. Table 50 shows the quadratic sum of all observed shifts.

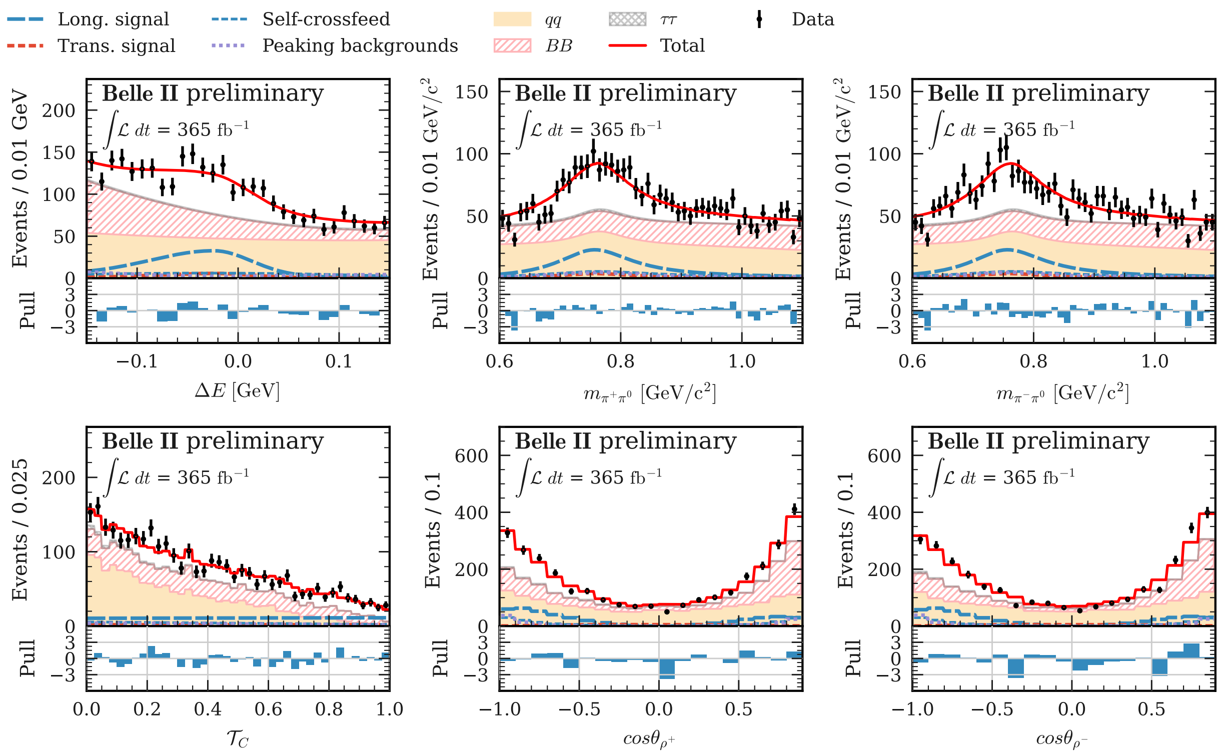


FIG. 91. Distributions of the ΔE, $\mathrm{m}_{\pi^\pm\pi^0}$, $\mathcal{T}_\mathrm{C}$, and $\cos\theta_{rho^\pm}$ for the signal-extraction fit.

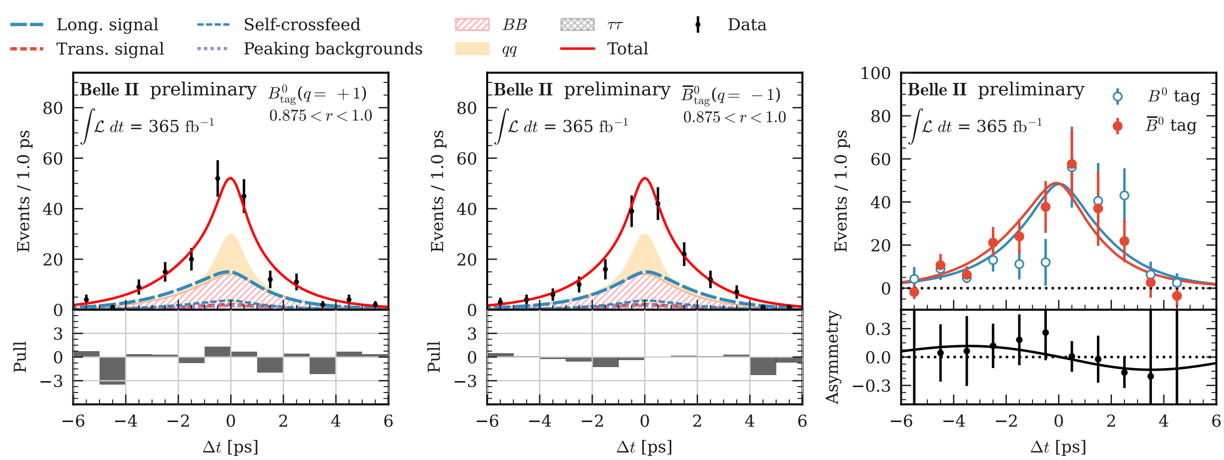


FIG. 92. Distributions of the fitting to Δt. The left and middle plots correspond to the fit projections in the best flavor tagging bin at $q = +1$ and $q = -1$. The right plot shows the Δt distributions for the longitudinal signals obtained by the $_s\mathcal{P}lot$ technique [19].

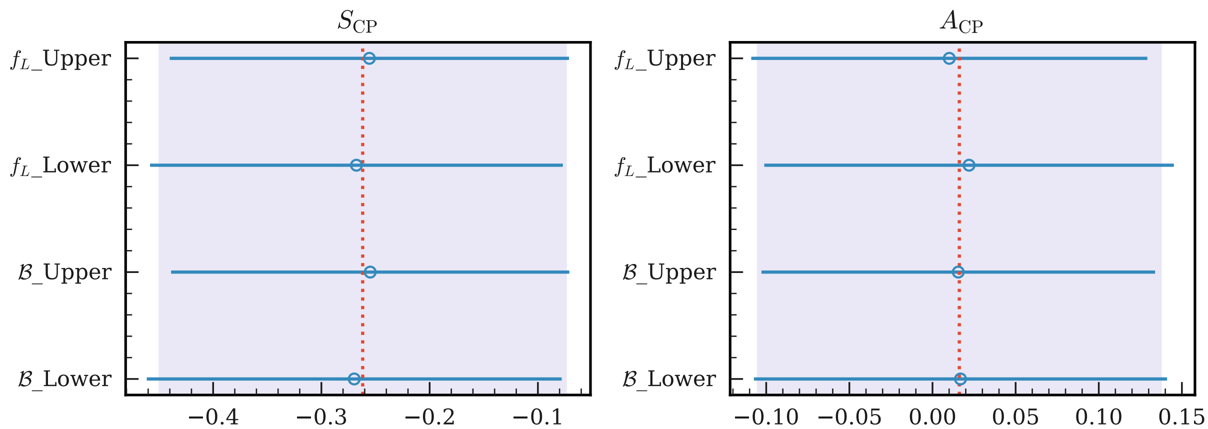


FIG. 93. Observed shifts in CP parameters by event fraction

	S_{CP}	A_{CP}
f_L+	0.0062	-0.006
f_L-	-0.0058	0.0059
$\mathcal{B}$+	0.0069	-0.0006
$\mathcal{B}$-	-0.0078	0.0007

TABLE 49. Observed shifts in CP parameters by event fraction

Parameter	Uncertainty
S_{CP}	${}^{+0.0093}_{-0.0097}$
A_{CP}	${}^{+0.0059}_{-0.0061}$

TABLE 50. Uncertainties in CP violation parameters due to event fraction

8. SYSTEMATIC UNCERTAINTIES

Table 51 lists the total statistical and systematic uncertainties. The source of each systematic uncertainty is described in this section.

	$\mathcal{B}(\%)$	f_L	S_{cp}	A_{cp}
Statistical uncertainty	$^{+7.93}_{-7.58}$	$^{+0.024}_{-0.025}$	± 18.8	± 12.1
Total systematic uncertainty	$^{+10.1}_{-9.51}$	$^{+0.017}_{-0.015}$	$^{+0.083}_{-0.078}$	$^{+0.054}_{-0.061}$
Resolution	—	—	$^{+0.034}_{-0.044}$	$^{+0.014}_{-0.019}$
Event fraction	—	—	$^{+0.009}_{-0.010}$	± 0.006
Δt PDF for $q\bar{q}$ and $B\bar{B}$	—	—	$^{+0.038}_{-0.018}$	$^{+0.001}_{-0.007}$
Physics Parameters	—	—	$^{+0.014}_{-0.016}$	± 0.003
Interference	± 1.20	± 0.005	± 0.028	± 0.017
Tag side interference	—	—	± 0.005	± 0.021
Wrong tag fraction	—	—	$^{+0.002}_{-0.003}$	± 0.005
Tracking	± 0.54	—	—	—
π^0 efficiency	± 7.67	—	—	—
PID	± 0.08	—	—	—
$\mathcal{T}_C$	± 2.87	—	—	—
Data-MC mis-modeling	$^{+3.51}_{-1.70}$	$^{+0.008}_{-0.003}$	$^{+0.006}_{-0.011}$	$^{+0.006}_{-0.015}$
$\mathcal{B}$'s of peaking backgrounds	$^{+0.94}_{-0.98}$	± 0.001	$^{+0.006}_{-0.005}$	± 0.001
Single candidate selection	± 0.55	± 0.003	± 0.013	± 0.019
SCF ratio	$^{+2.96}_{-2.45}$	$^{+0.002}_{-0.003}$	$^{+0.005}_{-0.004}$	$^{+0.000}_{-0.007}$
Signal model	$^{+1.14}_{-2.02}$	± 0.002	$^{+0.011}_{-0.014}$	$^{+0.004}_{-0.003}$
$q\bar{q}$ model	$^{+0.49}_{-0.51}$	$^{+0.001}_{-0.002}$	$^{+0.022}_{-0.010}$	± 0.002
$B\bar{B}$ model	$^{+1.00}_{-0.40}$	$^{+0.003}_{-0.001}$	± 0.009	$^{+0.005}_{-0.007}$
$\tau^+\tau^-$ model	$^{+0.17}_{-0.26}$	$^{+0.000}_{-0.001}$	± 0.001	± 0.000
Peaking model	$^{+1.36}_{-1.01}$	$^{+0.003}_{-0.005}$	$^{+0.008}_{-0.004}$	$^{+0.004}_{-0.002}$
f_{00}	$^{+1.67}_{-1.50}$	—	—	—
$N_{\Upsilon(4S)}$	± 1.45	—	—	—
MC sample size	± 0.24	± 0.002	—	—
Fit bias	± 1.03	± 0.012	± 0.020	± 0.006
Mis-Alignment	—	—	± 0.014	± 0.005
Background CP Violation	—	—	$^{+0.038}_{-0.036}$	$^{+0.037}_{-0.041}$
$\tau^+\tau^-$ background yield	$^{+0.65}_{-0.69}$	± 0.000	± 0.009	$^{+0.001}_{-0.000}$
CP Violation in TP signal	—	—	$^{+0.008}_{-0.002}$	$^{+0.004}_{-0.002}$

TABLE 51. Total statistical and systematic uncertainties.

8.1. Tracking

We estimate the uncertainty of the charged tracking efficiency using $e^+e^- \to \tau^+\tau^-$ data. We use the event where one τ decays leptonically ($\tau \to l\nu_l\nu_\tau (l = e, \mu)$ and the other decays hadronically ($\tau \to 3\pi\nu_\tau$). We identified $e^+e^- \to \tau^+\tau^-$ events from the hadronic tau decay and measured the charged tracking efficiency using the leptonic τ decays. We obtained the data-simulation ratio to be

0.05 GeV $< p <$ **0.12 GeV :** 0.947 ± 0.0196,

0.12 GeV $< p <$ **0.16 GeV :** 0.985 ± 0.0173,

0.16 GeV $< p <$ **0.20 GeV :** 0.983 ± 0.0188, and

0.20 GeV $< p$ **:** ± 0.0027.

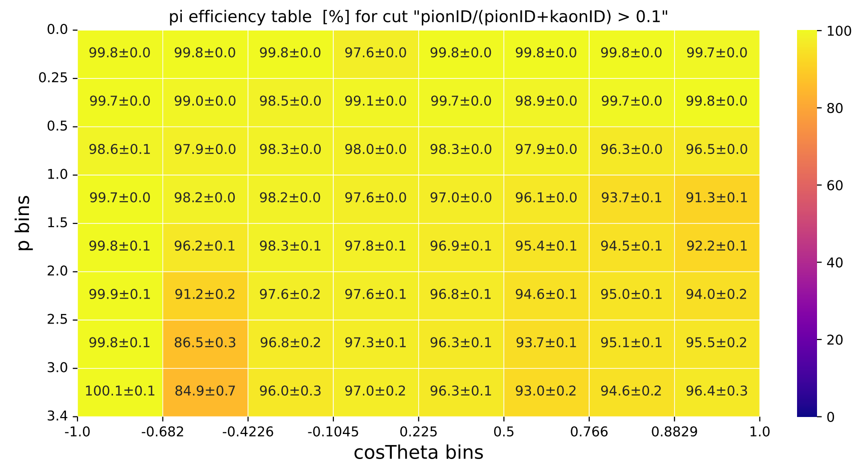
The weighted average of the uncertainty is 0.29% for each track in the final state of the target mode. We linearly add this number, $0.29\% \times 2$ tracks, to make the total uncertainty of the tracking efficiency in this $B^0 \to \rho^+\rho^-$ analysis.

8.2. PID efficiency

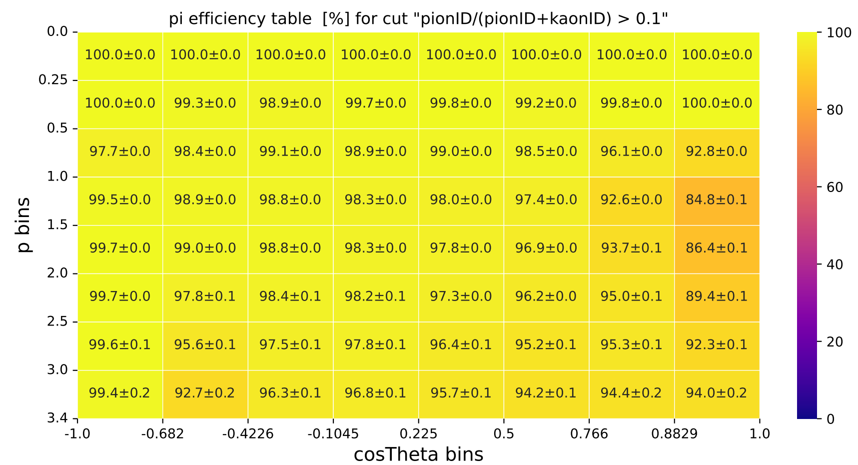
We measured the data-simulation discrepancy in the PID efficiency using the $D^{*-} \to D^0\pi^+$ data. The D^0 flavor can be identified using the charge of the slow $\pi^\pm$. We measured the Data-MC ratio in the pion ID using the π from $D^0 \to K^+\pi^-$. The Data-MC discrepancies in the charged pion identification are given in the space parameterized by $\pi^\pm$ momentum and $\cos\theta$ in the Lab. frame, as shown in Figure 94. The estimated Data/MC efficiency ratio is 0.9946 ± 0.0004. The uncertainty is calculated to be $0.0004/0.9946 = 0.0004$ per $\pi^\pm$.

8.3. π^0 efficiency

We measured the Data-MC ratio in the π^0 efficiency using the decays $D^0 \to K^-\pi^+\pi^0$, $D^0 \to K^-\pi^+$. We measured the ratio of the yield of the two decay modes and compared this with the branching fraction ratio measured by the other experiments. We also used $B^+ \to \pi^+ D^{*0}(D^{*0} \to \pi^0 D^0, D^0 \to K^-\pi^+)$ for the slow π^0s. Figure 95 shows the Data/MC ratio of the π^0 efficiency and its uncertainty. The uncertainty is dominated by the branching fraction ratio of $D^0 \to K^-\pi^+\pi^0$ and $D^0 \to K^-\pi^+$. The weighted average efficiency of the two π^0 in $B^0 \to \rho^+\rho^-$ and its error was computed while considering their momentum distributions. The estimated Data/MC efficiency ratio is 1.0015 ± 0.0401. The uncertainty is calculated to be $0.0401/1.0015 = 0.0400$ per π^0.



(a) Data



(b) MC15rd

FIG. 94. PID efficiency for $\pi^{\pm}$.The x axis and y axis are $\cos\theta$ and $\pi^{\pm}$ momentum in the Lab. frame, respectively.

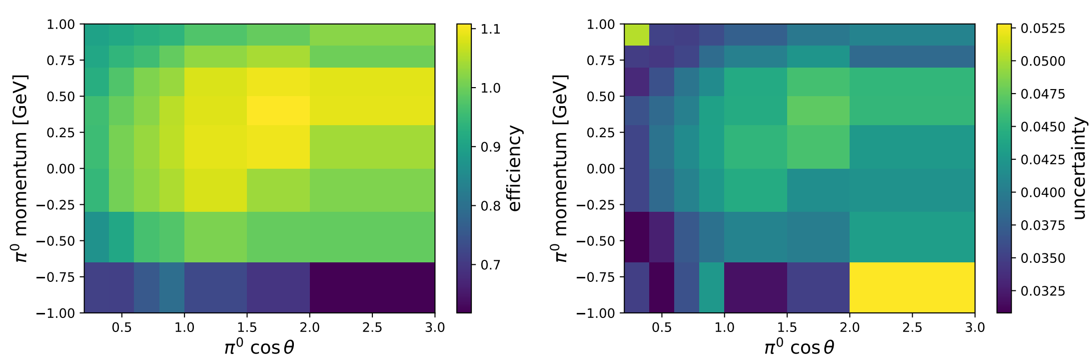


FIG. 95. (Left) Measurement Data/MC ratio of the π^0 efficiency in [0.2, 3.0] GeV and [-1.0, 1.0] in $\cos\theta$. (Right) The uncertainties in the same binning.

8.4. Continuum suppression efficiency

The Data-MC ratio of CSMVA efficiency was discussed in Section 6.3. The measured Data-MC ratio was 1.082 ± 0.031. The observed difference is corrected. The statistical uncertainty gives 2.87% systematic uncertainty to $\mathcal{B}$. Other fit parameters are not affected by the signal efficiency.

8.5. Data-MC mis-modelling for $cos\theta_\rho$

As discussed in Section 6.5.1, the Data-MC discrepancy of $\cos\theta_\rho$ was observed in the sideband region and was corrected by applying the Data-MC ratio in P_π. The systematic uncertainty arising from the Data-MC discrepancy of $\cos\theta_\rho$ was estimated by performing fittings while changing the PDF shape for qq and BB within the remaining Data-MC difference after the correction including the statistical uncertainties. In addition, we also considered the remaining Data-MC discrepancy in $cos\theta_\rho$ PDF found in $B \to D^0\rho$.

	$\mathcal{B}$	f_L	$S_{\rm CP}$	$A_{\rm CP}$
Signal	$^{+1.00}_{-0.0}$	$^{+0.008}_{-0.000}$	$^{+0.00}_{-0.01}$	$^{+0.00}_{-0.01}$
qq	$^{+0.08}_{-0.44}$	$^{+0.000}_{-0.003}$	$^{+0.00}_{-0.00}$	$^{+0.00}_{-0.00}$
BB	$^{+0.05}_{-0.22}$	$^{+0.002}_{-0.000}$	$^{+0.00}_{-0.00}$	$^{+0.00}_{-0.00}$
peaking	$^{+0.13}_{-0.04}$	$^{+0.002}_{-0.00}$	$^{+0.00}_{-0.00}$	$^{+0.00}_{-0.00}$
Total	$^{+1.02}_{-0.49}$	$^{+0.008}_{-0.003}$	$^{+0.001}_{-0.010}$	$^{+0.01}_{-0.01}$

TABLE 52. Systematic uncertainty related to $cos\theta_\rho$ PDF modeling. The systematics for the $\mathcal{B}$ is absolute.

8.6. Best candidate selection

The best candidate selection is performed using the method in Section 3.7. To consider the systematic uncertainty due to the best candidate selection, we compared the fitted results obtained when the best candidate selection was performed in a nominal way and when it was performed randomly. The random best candidate selection was performed with a different random seed for each fit. 1000 different random seeds were applied for this test in total. The systematic uncertainty was estimated to be the average difference between them.

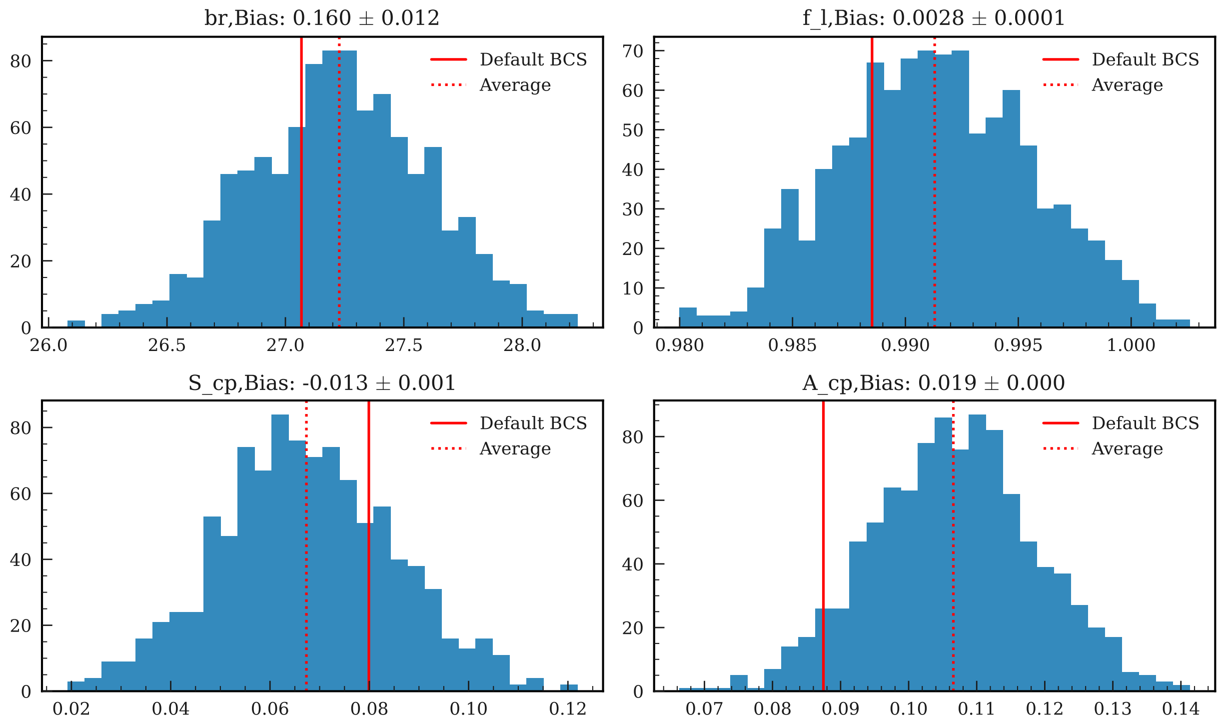


FIG. 96. Systematic uncertainty due to best candidate selection. The red solid line represents the MC-fit results for our best candidate selection.

8.7. self-crossfeed ratio

In the fitting, the ratio of the self-crossfeed events to the signal is fixed to the value calculated using MC samples. We estimated a bias on the physics parameters that the MC-based self-crossfeed ratio yields by changing the ratio up to ± 50% from the nominal value. The fitting was performed with the nominal self-crossfeed ratio. Fifty bootstrap samples were generated and fitted to dataset for each self-crossfeed ratio value. Figure 97 shows the test result. Conservatively, we assessed the systematic uncertainty corresponding to ±20% of the self-crossfeed difference which is shown in Table 53

Parameter	Systematic uncertainty
$\mathcal{B}$	$^{+0.86}_{-0.71}$
f_L	$^{+0.002}_{-0.003}$
$S_{\rm CP}$	$^{+0.005}_{-0.004}$
$A_{\rm CP}$	$^{+0.000}_{-0.007}$

TABLE 53. Systematic uncertainty due to self-crossfeed ratio

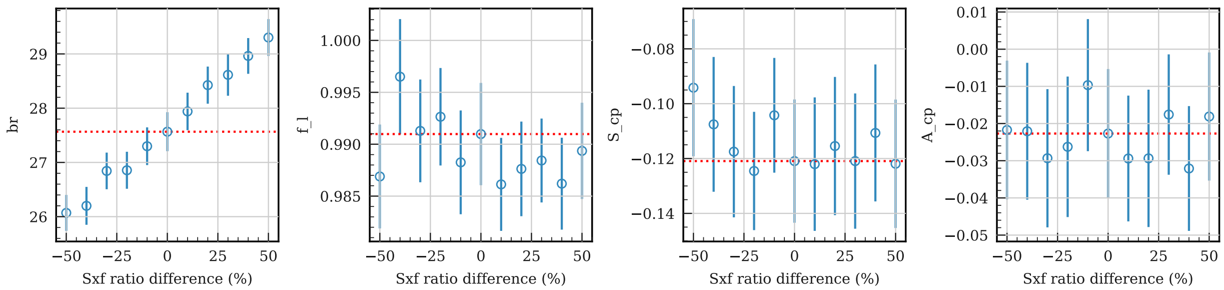


FIG. 97. The bias dependency on the self-crossfeed ratio to the signal yield

8.8. PDF modelling

To consider the systematic uncertainty due to the statistical uncertainty of the PDF shape, fittings were performed while changing the PDF shape parameters one by one within the statistical uncertainty of the fit result to the data. Table 54 shows the systematic uncertainties due to the PDF shapes. The systematic uncertainty for each component is a sum of squares of the observed bias for all the PDF shape parameters.

	$\mathcal{B}(\%)$	f_L	S_{cp}	A_{cp}
Signal model	$^{+1.14}_{-2.02}$	± 0.002	$^{+0.011}_{-0.014}$	$^{+0.004}_{-0.003}$
$q\bar{q}$ model	$^{+0.49}_{-0.51}$	$^{+0.001}_{-0.002}$	$^{+0.022}_{-0.010}$	± 0.002
$B\bar{B}$ model	$^{+1.00}_{-0.40}$	$^{+0.003}_{-0.001}$	± 0.009	$^{+0.005}_{-0.007}$
$\tau\tau$ model	$^{+0.17}_{-0.26}$	$^{+0.000}_{-0.001}$	± 0.001	± 0.000
Peaking model	$^{+1.37}_{-1.01}$	$^{+0.003}_{-0.005}$	$^{+0.008}_{-0.004}$	$^{+0.004}_{-0.002}$

TABLE 54. Systematic uncertainty due to PDF shape.

8.9. Branching fraction of peaking backgrounds

The branching fractions of the several peaking background modes are fixed in the fitting. Table 55 shows peaking background modes and their branching fractions. To consider the systematic uncertainty arising from the uncertainties of their branching fractions, fitting was performed while changing the fixed value for those branching fractions to the data. Table 56 displays the results. The systematic uncertainty is calculated from the squared sum of the observed biases. The systematic uncertainty due to the peaking backgrounds is much smaller than statistical uncertainty.

mode	$\mathcal{B}(\times 10^{-6})$
$B^0 \to \rho^0\pi^0\pi^0$	10.0 ± 10.0
$B^0 \to a_1^+\pi^-$	26.0 ± 5.0
$B^0 \to f^0\pi^0\pi^0$	10.0 ± 10.0
$B^0 \to \omega^0\pi^0$	10.0 ± 10.0
$B^0 \to \rho^- K^{*+}$	10.3 ± 2.6
$B^0 \to \rho^- K_0^{*+}$	28.0 ± 12.0

TABLE 55. Peaking background mode whose yield was fixed in the fitting

	$\mathcal{B}(\%)$	f_L	S_{cp}	A_{cp}
$\mathcal{B}$ of peaking backgrounds	$^{+0.94}_{-0.98}$	± 0.001	$^{+0.006}_{-0.005}$	± 0.001

TABLE 56. Systematic uncertainty due to the uncertainty of the $\mathcal{B}$ of peaking backgrounds

8.10. $\tau\bar{\tau}$ background yields

In the fitting, the $\tau\bar{\tau}$ background yield is fixed to MC expectation (40.2). To estimate the systematic uncertainty, we performed fit to data while changing the $N_{\tau\tau}$ within $\pm 100\%$ from the MC expectation. Table 57 summarizes the result of the estimation.

	$\mathcal{B}(\%)$	f_L	S_{cp}	A_{cp}
$\tau\tau$ background yield	$^{+0.65}_{-0.69}$	± 0.000	± 0.009	$^{+0.001}_{-0.000}$

TABLE 57. Systematic uncertainties due to $N_{\tau\tau}$

8.11. Fit bias

Fitter validation was performed in Section 4.7.1 and Section 5.7.1. The fitter introduces a slight bias in the signal branching fraction due to the correlation with the yields of the peaking backgrounds. Table 58 shows the systematic uncertainties related to the fitter bias. To consider the systematic uncertainties of it, all the observed biases were assessed. In this analysis, the bias was not corrected in the final fitting and assessed systematic errors of this bias to both positive and negative sides, conservatively.

Parameter	Mean of the bootstrap study	Input value	Bias
$\mathcal{B}(\times 10^{-6})$	27.4 $\pm$ 0.1	27.7	0.3
f_L	1.001 $\pm$ 0.001	0.99	0.012
S_{cp}	-0.118 $\pm$ 0.006	-0.14	0.02
A_{cp}	0.006 $\pm$ 0.004	0.0	0.006

TABLE 58. Systematic uncertainties due to the fitter bias

8.12. Δt resolution function

In the fitting, we fixed the resolution function shape parameters, and some of them were calibrated using data, and the others were fixed to the fit result for $B^0 \to \rho^+\rho^-$ signal MC samples. We calculated the systematics due to both, and Table 59 shows the total systematic uncertainty due to the resolution function shape. In this section we discuss the estimation for both systematic uncertainties.

Parameter	Systematic uncertainty
S_{cp}	$^{+0.034}_{-0.044}$
A_{cp}	$^{+0.017}_{-0.0014}$

TABLE 59. Systematic uncertainty due to the resolution function parameters

8.12.1. Resolution function parameters calibrated using data

The Δt resolution function shape is validated using $B^0 \to D^{*-}\pi^+$ samples. The systematic uncertainty is estimated by varying the resolution function parameters by $\pm$ 1 σ. Figure 98 and Table 60 shows the observed shift in CP violation parameters due to the resolution function parameters calibrated by $B^0 \to D^*\pi$ data. Table 61 shows the quadratic of the observed differences.

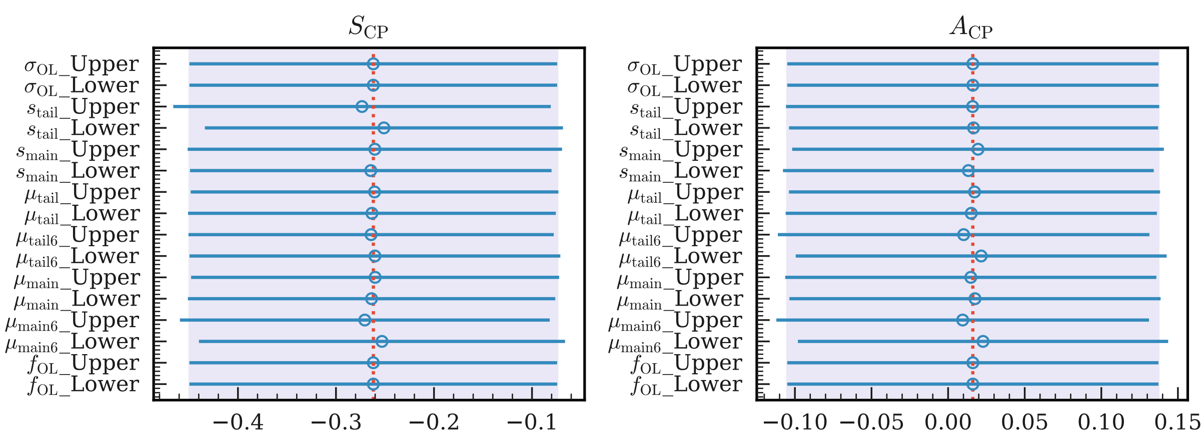


FIG. 98. Systematic uncertainty due to the statistical uncertainty of the fitting to $B^0 \to D^{*-}\pi^+$ data.

	$S_{\rm CP}$	$A_{\rm CP}$
$\sigma_{\rm OL}$+	-0.0001	0.0000
$\sigma_{\rm OL}$-	-0.0001	0.0000
$s_{\rm tail}$+	-0.0116	-0.0001
$s_{\rm tail}$-	0.0108	0.0005
$s_{\rm main}$+	0.0015	0.0034
$s_{\rm main}$-	-0.0026	-0.0029
$\mu_{\rm tail}$+	0.0012	0.0011
$\mu_{\rm tail}$-	-0.0015	-0.001
$\mu_{\rm tail6}$+	-0.0023	-0.006
$\mu_{\rm tail6}$-	0.0016	0.0055
$\mu_{\rm main}$+	0.0018	-0.0013
$\mu_{\rm main}$-	-0.0018	0.0014
$\mu_{\rm main6}$+	-0.0087	-0.0066
$\mu_{\rm main6}$-	0.0088	0.0067
$f_{\rm OL}$+	-0.0001	0.0000
$f_{\rm OL}$-	-0.0001	0.0000

TABLE 60. Observed bias due to the statistical uncertainty of the resolution function

Parameter	Systematic uncertainty
$S_{\rm CP}$	$^{+0.0142}_{-0.0151}$
$A_{\rm CP}$	$^{+0.0094}_{-0.0095}$

TABLE 61. Systematic uncertainty due to resolution function shape calibrated by $B^0 \to D^*\pi$ data.

8.12.2. Resolution function parameters fixed to MC

Some of the resolution function parameters are fixed to the fit result for $B^0 \to \rho^+\rho^-$ signal MC in the both resolution function validation and Δt fitting to $B^0 \to \rho^+\rho^-$. To estimate the systematic uncertainty of this, we performed fittings to $B^0 \to \rho^+\rho^-$ samples while floating the resolution function parameters one by one with $S_{\rm CP}$ and $A_{\rm CP}$ floated. Figure 99 shows the observed shift for each parameters, and Table 62 shows the quadratic sum of all shifts.

Parameter	Systematic uncertainty
S_{cp}	$^{+0.0304}_{-0.0415}$
A_{cp}	$^{+0.0106}_{-0.0169}$

TABLE 62. Systematic uncertainty due to the resolution function parameters which are fixed to the fit result for $B^0 \to \rho^+\rho^-$ signal MC

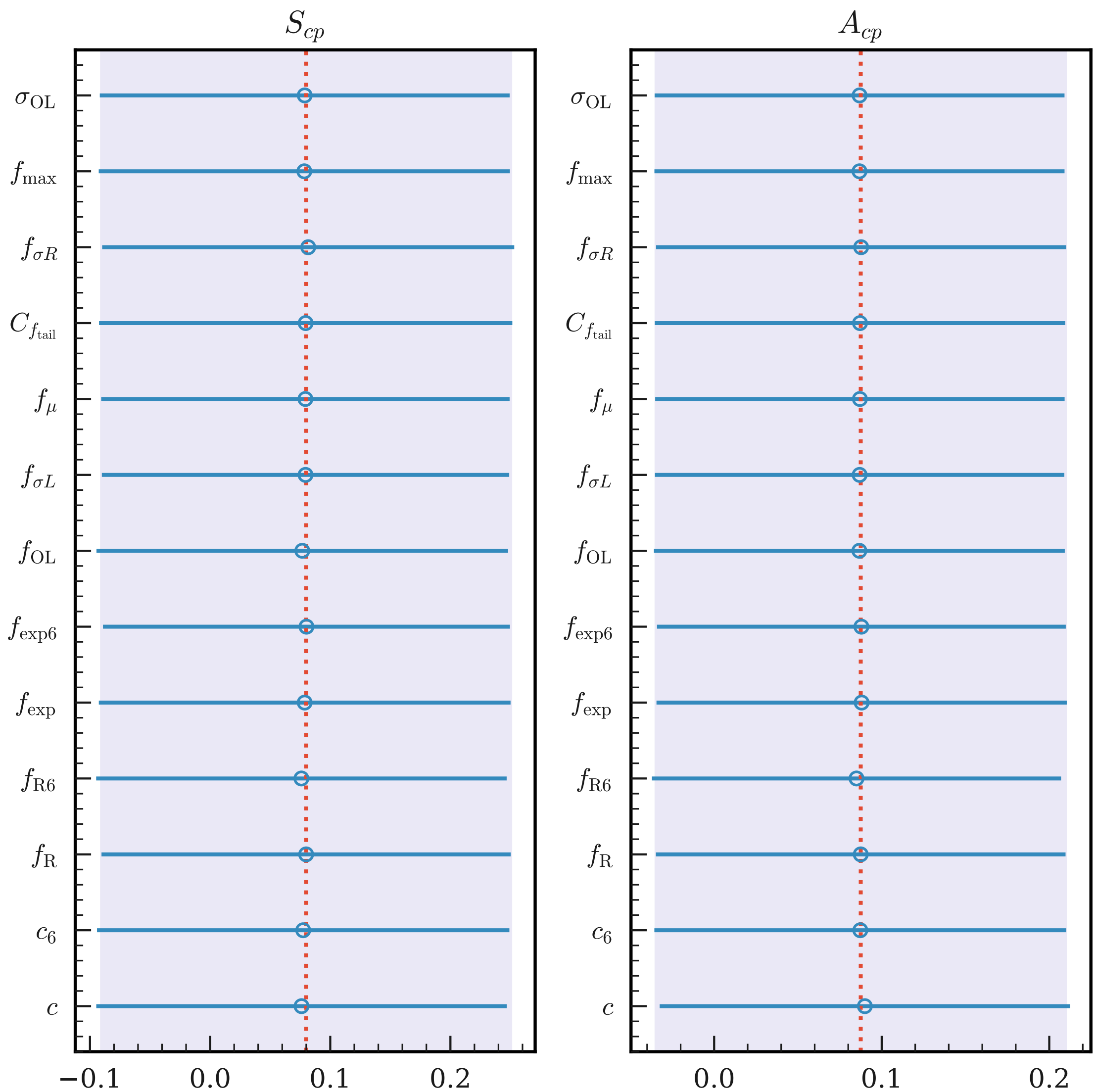


FIG. 99. Systematic uncertainty due to the resolution function parameters which are fixed to the fit result for $B^0 \to \rho^+\rho^-$ signal MC

8.13. Background Δt PDF shape

The Δt PDF shape for $q\bar{q}$ and $B\bar{B}$ is determined from the fit result to the off-resonance and sideband data. To consider the systematic uncertainty, the CP was performed while varying the background Δt PDF shape within the uncertainty obtained in the sideband fit. Figure 100 and Table 63 show the observed shift when the background Δt PDF shape parameters fluctuate. Table 63 shows the quadratic sum of all observed biases.

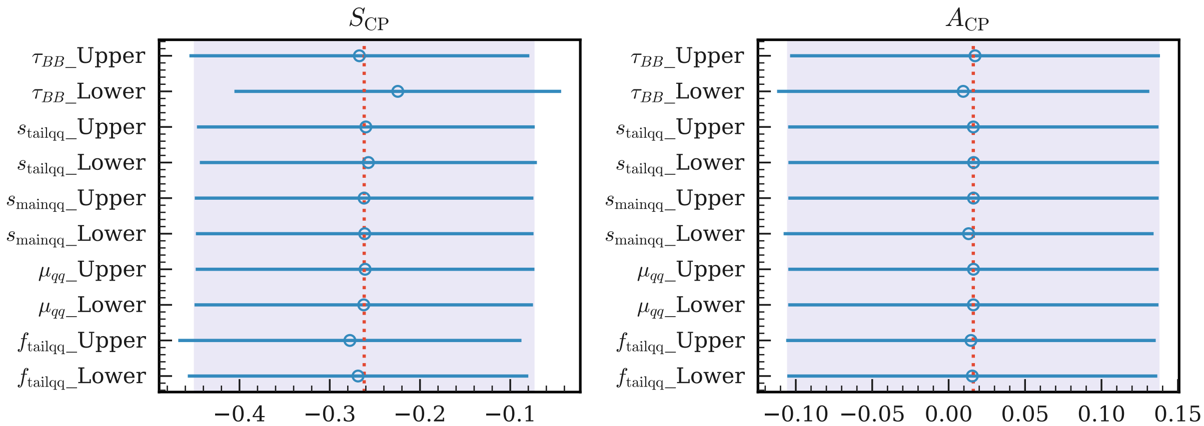


FIG. 100. Observed bias due to the uncertainty of the background Δt PDF shape.

	$S_{\rm CP}$	$A_{\rm CP}$
τ_{BB}+	-0.0053	0.0011
τ_{BB}-	0.0372	-0.0066
$s_{\rm tailqq}$+	0.0018	0.0000
$s_{\rm tailqq}$-	0.0046	0.0002
$s_{\rm mainqq}$+	-0.0001	0.0001
$s_{\rm mainqq}$-	0.0006	-0.0031
μ_{qq}+	0.0010	0.0001
μ_{qq}-	-0.0005	0.0000
$f_{\rm tailqq}$+	-0.0159	-0.0015
$f_{\rm tailqq}$-	-0.0068	-0.0007

TABLE 63. Observed bias due to the uncertainty of the background Δt PDF shape.

Parameter	Systematic uncertainty
$S_{\rm CP}$	$^{+0.0376}_{-0.0181}$
$A_{\rm CP}$	$^{+0.0012}_{-0.0075}$

TABLE 64. Systematic uncertainty due to background Δt PDF shape

8.14. Wrong tag fraction

The systematic uncertainty due to the wrong tagging fraction is calculated by performing a fitting while changing the wrong tag fraction parameters within the uncertainty. To consider the correlation of the parameters, we changed parameters one by one within the uncertainty, and at the same time, all parameters were also changed using correlation. The uncertainty and correlation used in the paper are the same as the result of [10]. Figure 101 and Table 65 shows the observed bias in CP parameters, and total systematic uncertainty related to the wrong tag fraction is in Table 66.

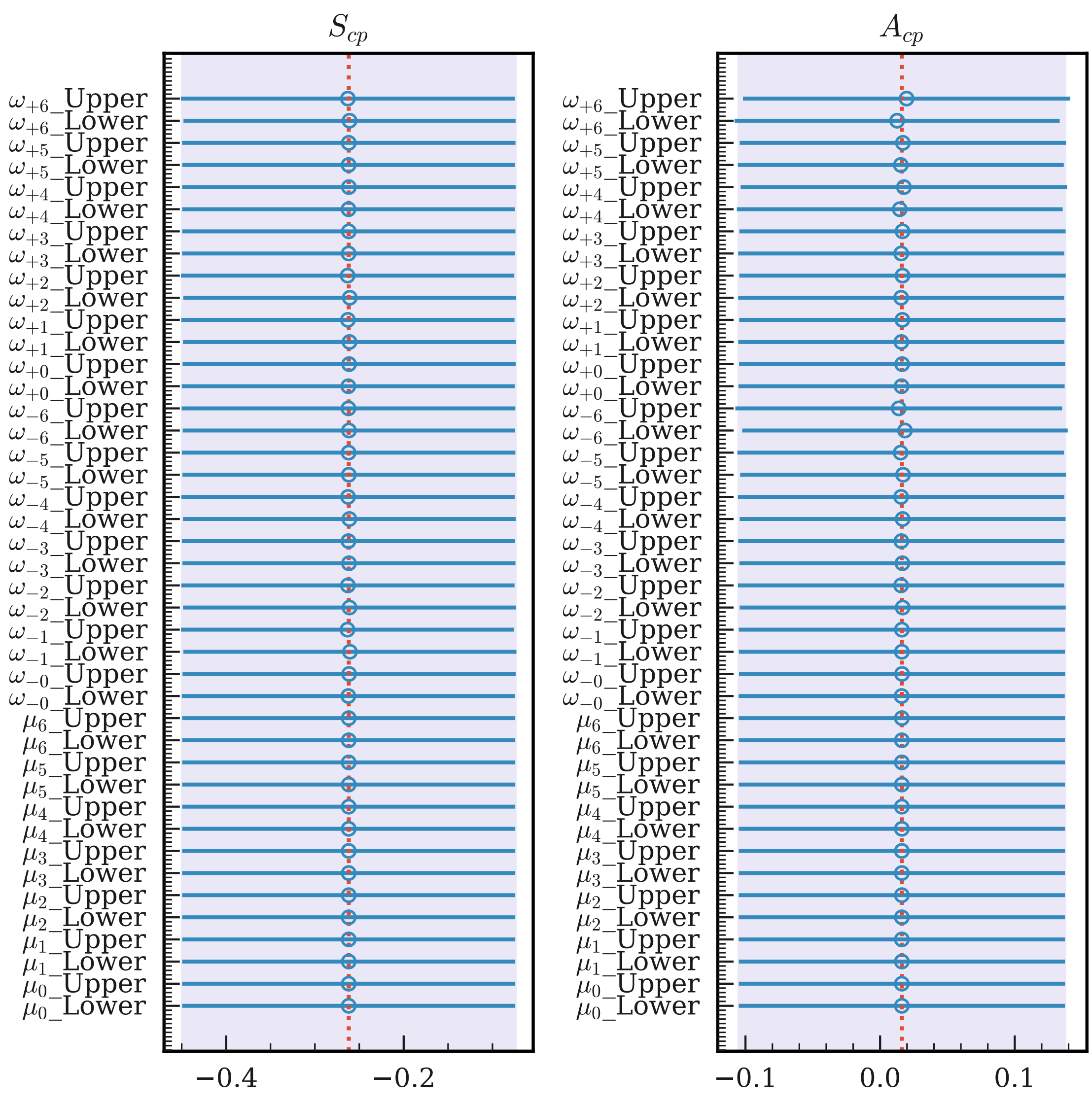


FIG. 101. Systematic uncertainty due to the statistical uncertainty of the wrong tagging fraction

	S_{CP}	A_{CP}
ω_{+6}+	-0.001	0.0035
ω_{+6}-	0.0009	-0.0035
ω_{+5}+	0.0000	0.0008
ω_{+5}-	-0.0002	-0.0007
ω_{+4}+	0.0001	0.0016
ω_{+4}-	-0.0003	-0.0015
ω_{+3}+	0.0000	0.0005
ω_{+3}-	-0.0002	-0.0005
ω_{+2}+	-0.0013	0.0005
ω_{+2}-	0.0011	-0.0005
ω_{+1}+	-0.001	0.0004
ω_{+1}-	0.0008	-0.0004
ω_{+0}+	0.0004	0.0002
ω_{+0}-	-0.0005	-0.0002
ω_{-6}+	-0.0003	-0.0023
ω_{-6}-	0.0002	0.0023
ω_{-5}+	-0.0002	-0.0009
ω_{-5}-	0.0000	0.0009
ω_{-4}+	-0.0008	-0.0005
ω_{-4}-	0.0007	0.0006
ω_{-3}+	-0.0004	-0.0004
ω_{-3}-	0.0002	0.0004
ω_{-2}+	-0.001	-0.0006
ω_{-2}-	0.0008	0.0006
ω_{-1}+	-0.0014	0.0000
ω_{-1}-	0.0013	0.0000
ω_{-0}+	0.0004	0.0002
ω_{-0}-	-0.0005	-0.0001
μ_6+	-0.0001	0.0000
μ_6-	-0.0001	0.0000
μ_5+	-0.0001	0.0000
μ_5-	-0.0001	0.0000
μ_4+	-0.0001	0.0000
μ_4-	-0.0001	0.0000
μ_3+	-0.0001	0.0000
μ_3-	-0.0001	0.0000
μ_2+	-0.0001	0.0000
μ_2-	-0.0001	0.0000
μ_1+	-0.0001	0.0000
μ_1-	-0.0001	0.0000
μ_0+	-0.0001	0.0000
μ_0-	-0.0001	0.0000

TABLE 65. Observed bias in CP parameters due to the statistical uncertainty of the wrong tagging fraction

Parameter	Systematic uncertainty
S_{cp}	$^{+0.0024}_{-0.0029}$
A_{cp}	$^{+0.0049}_{-0.0047}$

TABLE 66. Total systematic uncertainty due to wrong tag fraction

8.15. MC statistics

Signal efficiency is fixed to the value estimated using signalMC samples, and $\mathcal{B}$ and f_L are affected by the statistical uncertainty of the signal efficiency. Table 67 shows the systematic uncertainty from the uncertainty of the signal efficiency due to MC statistics.

Parameter	$\mathcal{B}(\times 10^{-6})$	f_L
Systematic uncertainty	$\pm$ 0.07	$\pm$ 0.002

TABLE 67. Systematic uncertainty due to MC statistics.

8.16. Physics parameters

In the Δt fitting, B^0 lifetime τ and B^0 mixing frequency Δm_d are fixed to the value of the PDG2023 [8]. To consider the systematic uncertainty, CP fits were performed while changing it within the uncertainty. Figure 102 and Table 68 show all observed biases, and Table 69 shows the squared sum of it.

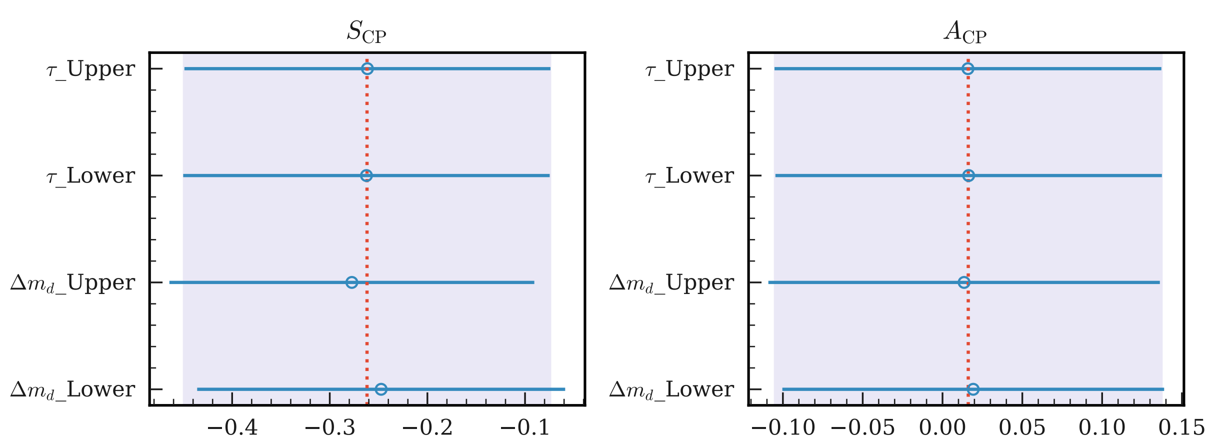


FIG. 102. Observed bias in CP parameters by physics parameters

	S_{CP}	A_{CP}
τ+	0.0004	-0.0002
τ-	-0.0006	0.0003
Δm_d+	-0.0155	-0.0026
Δm_d-	0.0145	0.0031

TABLE 68. Observed bias in CP parameters by physics parameters

Parameter	Systematic uncertainty
S_{CP}	${}^{+0.0145}_{-0.0156}$
A_{CP}	${}^{+0.0031}_{-0.0026}$

TABLE 69. Systematic uncertainty due to physics parameters

8.17. Number of the $B^0\bar{B}^0$ pairs

The measured the number of the $B\bar{B}$ pair is $(387.1 \pm 5.6) \times 10^6$. $\mathcal{B}$ is determined using it, and the systematic uncertainty of it is 0.4×10^{-6}.

8.18. Branching ratio of $\Upsilon(4S) \to B^0\bar{B}^0$

f^{+-}/f^{00} which is the ratio of Branching ratio of $\Upsilon(4S) \to B^+B^-$ and Branching ratio of $\Upsilon(4S) \to B^0\bar{B}^0$, is measured to be $f^{+-}/f^{00} = 1.065 \pm 0.052$. [20] It corresponds to $\pm 2.6\%$ uncertainty on the number of $B^0\bar{B}^0$ pair, and it means measured $\mathcal{B}$ has a $\pm 2.6\%$ of the systematic uncertainty.

8.19. Interference

Decay modes whose final state particles are the same as $B^0 \to \rho^+\rho^-$,$\pi\pi\pi^0\pi^0$, can interfere to $B^0 \to \rho^+\rho^-$ signals. In this analysis, $B^0 \to \rho\pi\pi^0$, $B^0 \to a_1\pi$, and $B^0 \to \pi\pi\pi^0\pi^0$ can interfere to the signals. We generated toy MC samples considering possible interference and perform fitting to it, and estimate the systematic uncertainty. To consider the interference effect, amplitudes for these mode was considered. Equation 42 shows the amplitude used to consider the interference. For the signal, $\mathcal{A}_{m_{\pi}^{+}\pi^0}$ can be describes using relativistic Breit-Wigner function, shown in Equation 43. $\mathcal{A}_{\cos\theta_\rho}$ can be modelled using a template histogram of signal MC samples. For $B^0 \to \rho\pi\pi^0$ and $B^0 \to a_1\pi$, $\mathcal{A}_{m_{\pi}^{+}\pi^0}$ is modelled using a sum of relativistic Breit-Wigner function and a first-order Chebyshev polynomial function, and $\mathcal{A}_{\cos\theta_\rho}$ is modelled using a template histograms of MC samples. Finally, for $B^0 \to \pi\pi\pi^0\pi^0$ samples, it doesn't have ρ in the intermediate state, and $\mathcal{A}_{m_{\pi}^{+}\pi^0}$ is modelled using a first-order Chebyshev polynomial function and $\mathcal{A}_{\cos\theta_\rho}$ is modelled using a template histograms of MC samples. All PDF shape parameters are determined from the fitting to each signal MC samples.

$$\mathcal{A} = \mathcal{A}_{m_{\pi}^{+}\pi^0}\mathcal{A}_{m_{\pi^-\pi^0}}\mathcal{A}_{\cos\theta_{\rho^+}}\mathcal{A}_{\cos\theta_{\rho^-}} \tag{42}$$

$$A(m) = \frac{p_{\rho^\pm \to \pi^\pm \pi^0}}{m^2 - m_0^2 + i m_0 \Gamma(m)} \frac{B(p_{B \to \rho^+ \rho^-})}{B(p'_{B \to \rho^+ \rho^-})} \frac{B(p_{\rho^\pm \to \pi^\pm \pi^0})}{B(p'_{\rho^\pm \to \pi^\pm \pi^0})} \tag{43}$$

To consider the interference of the signal and backgrounds, total amplitude is calculated using Equation 44, where $\mathcal{A}_{\text{total}}$ is total amplitude, $\mathcal{A}_{\rho\rho}$ and $\mathcal{A}_{\text{peaking}}$ are amplitudes for $B^0 \to \rho^+\rho^-$ and peaking background, and c_{peaking} is a fraction of the contribution of peaking backgrounds, and ϕ is the phase of the peaking backgrounds to the signal.

$$\mathcal{A}_{\text{total}} = \mathcal{A}_{\rho\rho} + c_{\text{peaking}} \exp(i\phi) \mathcal{A}_{\text{peaking}} \tag{44}$$

Figure 103 shows the distributions of the samples with and without interference between $B^0 \to \rho^+\rho^-$ and $B^0 \to \rho\pi\pi^0$. The shape and yields vary depending on the phase, as expected.

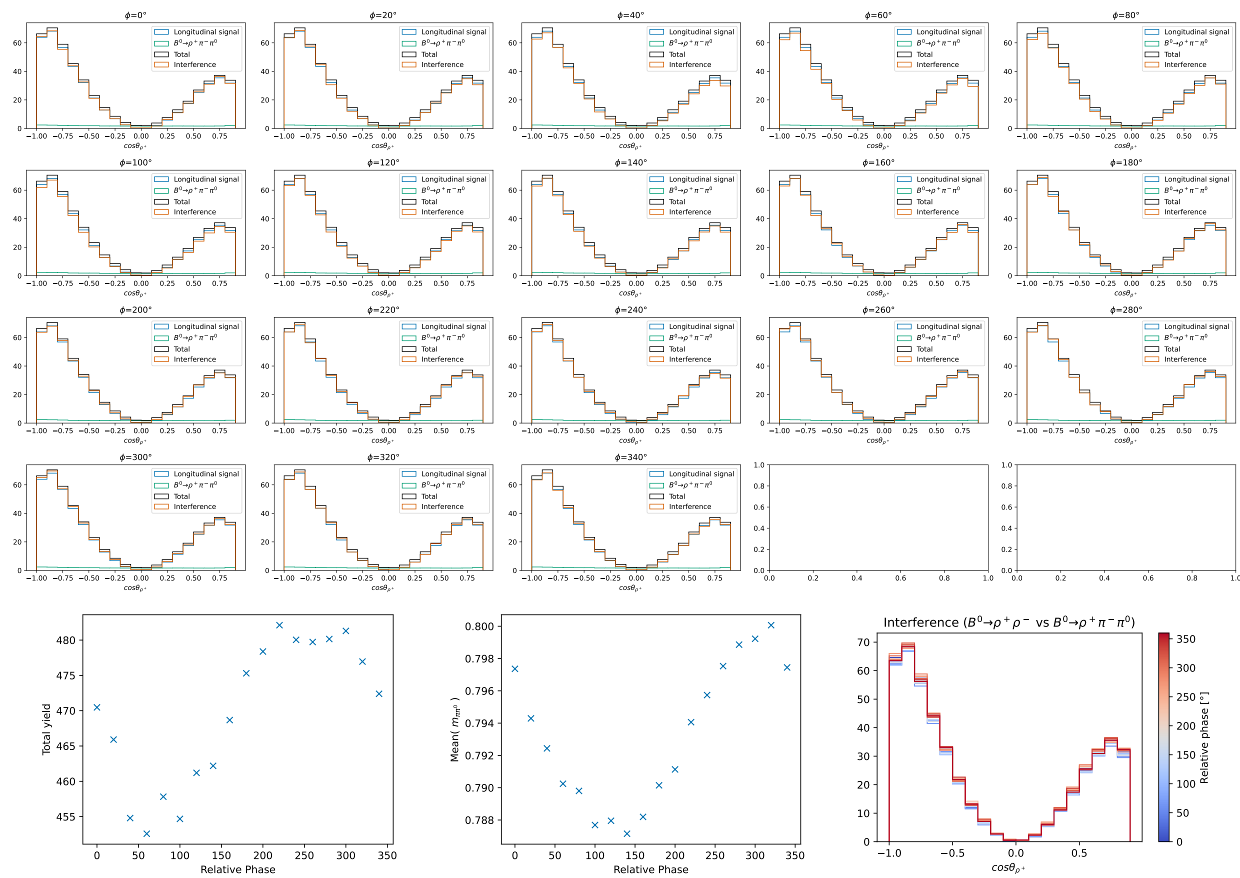


FIG. 103. Effects of interference between $B^0 \to \rho^+\rho^-$ and $B^0 \to \rho\pi\pi^0$ samples

Figure 104 shows the bias on the fit parameters due to interference. In this study, 500 toy samples with interference were generated for each relative phase and performed fitting and tested the bias. The RMS of the Observed bias in each mode is used for the systematic uncertainties. Table 70 shows the systematic uncertainty due to interference.

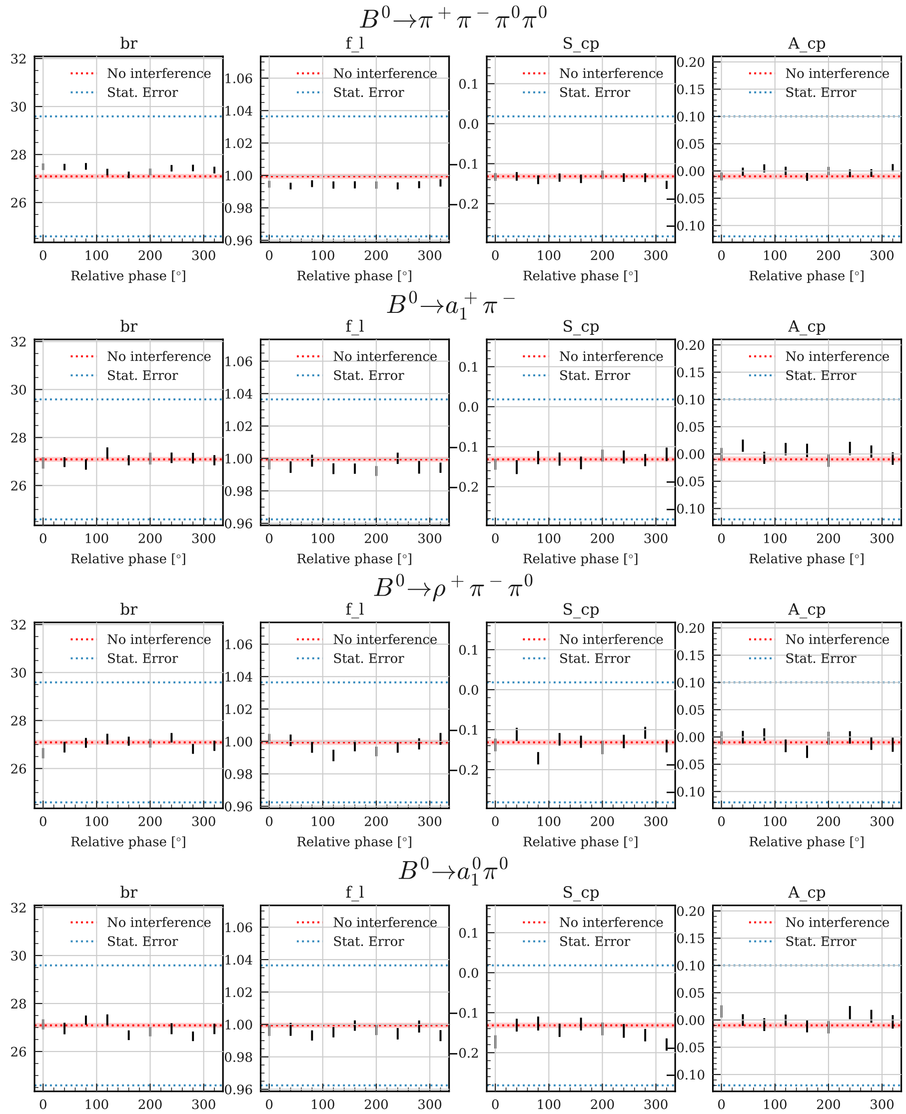


FIG. 104. Possible bias by the interference between the longitudinal signal and peaking background modes.

	$\mathcal{B}$	f_L	$S_{\rm CP}$	$A_{\rm CP}$
$B^0 \to \pi^+\pi^-\pi^0\pi^0$	±0.115	±0.001	±0.007	±0.005
$B^0 \to a_1^+\pi^-$	±0.133	±0.002	±0.010	±0.009
$B^0 \to \rho^+\pi^-\pi^0$	±0.193	±0.003	±0.018	±0.010
$B^0 \to a_1^0\pi^0$	±0.232	±0.002	±0.018	±0.010
total	±0.349	±0.005	±0.028	±0.017

TABLE 70. Systematic uncertainty due to interference

8.20. Tag side interference

The interference between CKM-favored $b \to c\bar{u}d$ and doubly-CKM-suppressed $\bar{b} \to \bar{c}u\bar{d}$ in the tag side can affect the time-dependent CPV measurement in the signal side. In the fitting, it is not considered, and the systematic uncertainty due to neglecting interference is estimated by generating TOY MC samples with and without tag-side interference effects. We formulate the time-dependence of the B^0 decay considering the tag-side interference effect as

$$\mathcal{P}_{B^0}(\Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau}\left[R_{B^0} + A_{B^0}\cos(\Delta m_d \Delta t) + S_{B^0}\sin(\Delta m_d \Delta t)\right], \tag{45}$$

$$\mathcal{P}_{\bar{B}^0}(\Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau}\left[R_{\bar{B}^0} + A_{\bar{B}^0}\cos(\Delta m_d \Delta t) + S_{\bar{B}^0}\sin(\Delta m_d \Delta t)\right]. \tag{46}$$

The parameter R, A, and S corresponds to CP-violating parameters considering the tag-side interference effects, which are defined as

$$R_{B^0,\bar{B}^0} = \frac{1+|\lambda_{CP}^2|}{2} - 2r'\mathrm{Re}[\lambda_{CP}]\cos(2\phi_1\phi_3 \mp \delta'), \tag{47}$$

$$A_{B^0,\bar{B}^0} = \pm\frac{|\lambda_{CP}^2|-1}{2} + 2r'\mathrm{Im}[\lambda_{CP}]\sin(2\phi_1\phi_3 \mp \delta'), \tag{48}$$

$$S_{B^0,\bar{B}^0} = \pm\,\mathrm{Im}[\lambda_{CP}] + r'(1-|\lambda_{CP}^2|)\sin(2\phi_1\phi_3 \mp \delta'). \tag{49}$$

The parameter $\lambda_{CP} = \frac{q}{p}\frac{\bar{\mathcal{A}}}{\mathcal{A}}$ where $\mathcal{A}$ is the amplitude of the signal decay, r' is the ratio of the amplitude of CKM-favored and doubly-CKM-suppressed decays, and δ' is the strong phase difference between CKM-favored and doubly-CKM-suppressed decays. We assume $r' = 0.018$ and $\delta' = -102°$, and generate TOY MC samples with and without tag-side interference and check the difference in the fitted CP-violation parameters. Figure 105 shows the result, and systematic uncertainty is summarised in Table 71

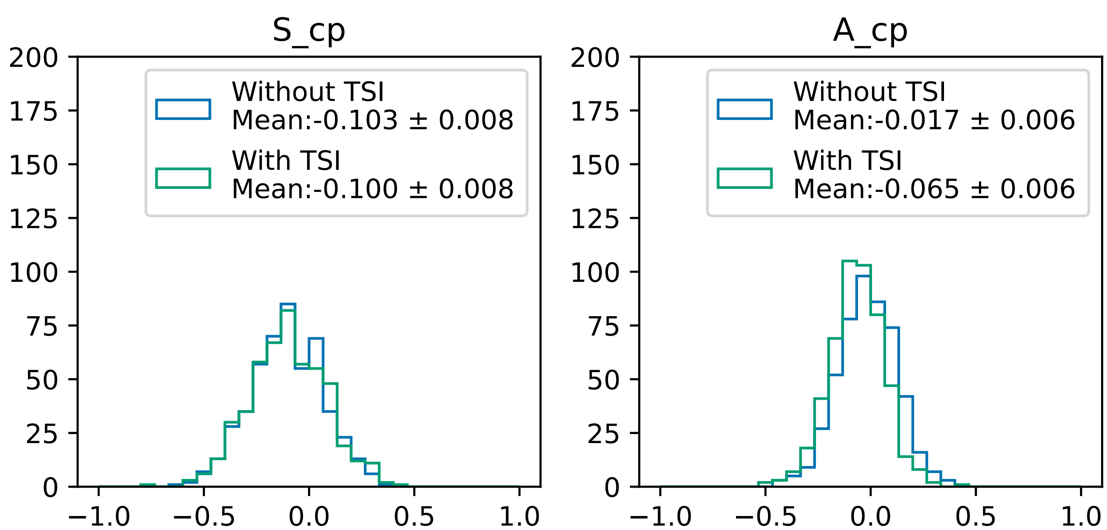


FIG. 105. Bias due to the tag-side interference

Parameter	S_{CP}	A_{CP}
Bias	0.003	-0.048

TABLE 71. Systematic uncertainty due to tag side interference

8.21. Mis-alignment

To estimate the systematic uncertainty due to the misalignment, we performed compared 4 scenarios with the nominal alignment. Figure 106 shows the results. All samples are generated with the same seeds, so we can ignore the statistical uncertainties. The systematic uncertainty is calculated by taking the maximum deviations from the nominal one.

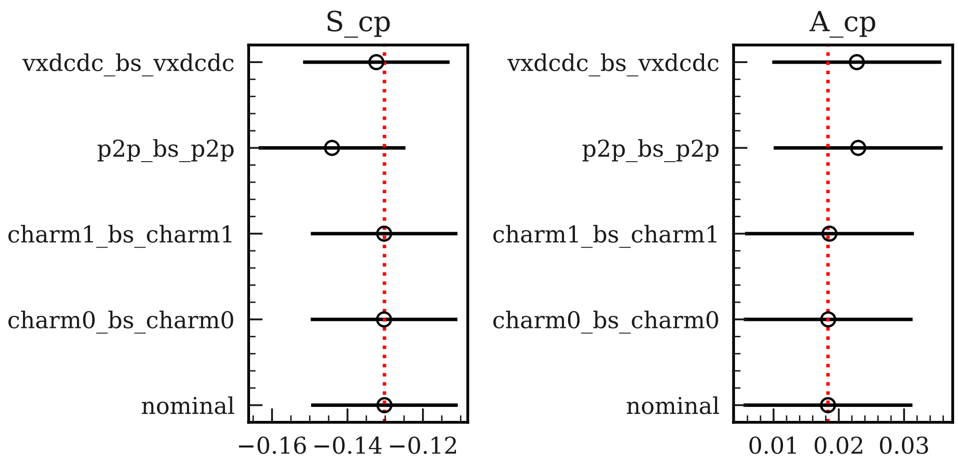


FIG. 106. Systematic uncertainty due to mis-alignment. vxdcdc is the impact of the second iteration of full global alignment including CDC wires, p2p is the difference between prompt alignment and reprocessing which is more advanced, charm1 is the ambiguity in the alignment from the Telescope weak mode, and charm0 is the statistical day-by-day variations.

8.22. Background CPV

Some peaking background modes can have CPV, and it can be the cause of the bias in S_{cp} and A_{CP} of $B^0 \to \rho^+\rho^-$. Figure 72 shows the possible CPV in the peaking background modes. We cited the result if the CPV parameters are measured, and if the CPV measurement has not been done, we assessed $\pm 50\%$ of CPV. To estimate the systematic uncertainties, we generated TOY samples changing the CPV parameters for the peaking background modes one by one, then performed bootstrap and TOY study. In this study, the signal and background modes except for the peaking background decay are the bootstrap samples, and the peaking background modes with different CPV parameters are TOY samples. Figure 107 shows the results, and all peaking decay modes affect much smaller than expected statistical uncertainty on CPV parameters of $B^0 \to \rho^+\rho^-$. The systematic uncertainty is calculated by taking a squared sum of all observed bias, and it is shown in Table 73.

Decay Mode	S_{CP}	A_{CP}
$B^0 \to \rho^+\pi^-\pi^0$	0.0 ± 0.5	0.0 ± 0.5
$B^0 \to \pi^+\pi^-\pi^0\pi^0$	0.0 ± 0.5	0.0 ± 0.5
$B^0 \to a_1^{\pm}\rho^{\mp}$	0.0 ± 0.5	0.0 ± 0.5
$B^0 \to \rho^- K^{*+}$	0.0 ± 0.0	0.21 ± 0.15
$B^0 \to \rho^- K_0^{*+}$	0.0 ± 0.0	0.0 ± 0.5
$B^0 \to \pi^{\pm}\pi^{\mp}\pi^0$	0.0 ± 0.5	0.0 ± 0.5
$B^0 \to a_1^+\pi^-$	-0.2 ± 0.4	0.05 ± 0.11
$B^0 \to a_1^0\pi^0$	0.0 ± 0.5	0.0 ± 0.5
$B^0 \to K_s\pi^0\pi^0$	0.89 ± 0.3	0.21 ± 0.2
$B^0 \to \rho^{\pm}\pi^{\mp}$	0.05 ± 0.07	0.03 ± 0.07
$B^{\pm} \to \rho^{\pm}\pi^0$	0.0 ± 0.0	0.02 ± 0.11
$B^{\pm} \to a_1^0\pi^{\pm}$	0.0 ± 0.0	0.0 ± 0.5
$B^{\pm} \to \rho^{\pm}\rho^0$	0.0 ± 0.0	0.05 ± 0.05
$B^{\pm} \to a_1^{\pm}\pi^0$	0.0 ± 0.0	0.0 ± 0.5
$B^{\pm} \to \rho^{\pm}a_1^0$	0.0 ± 0.0	0.0 ± 0.5

TABLE 72. Possible CPV in the peaking background modes.

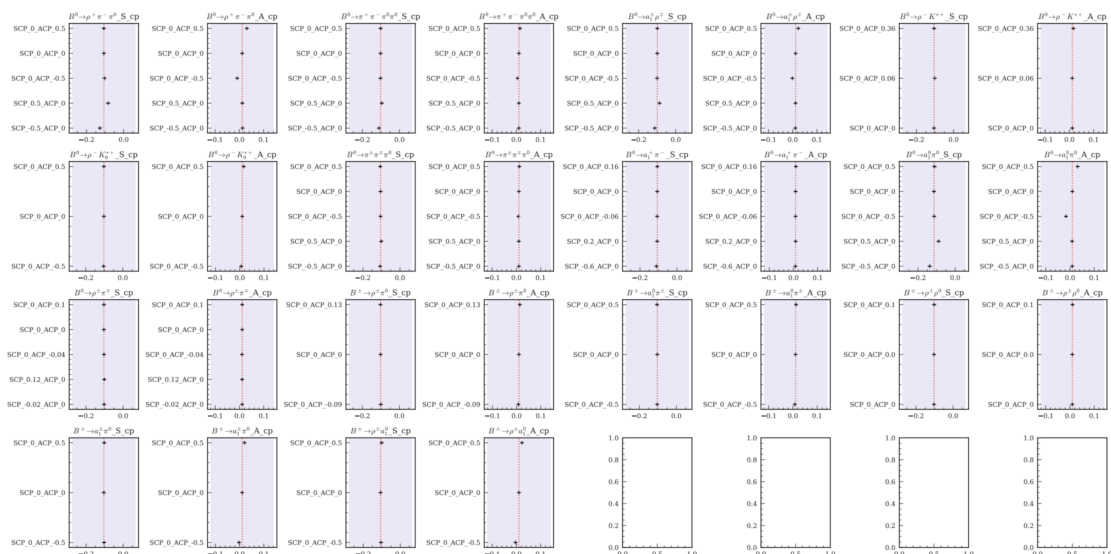


FIG. 107. Observed bias in TOY study when the CPV parameters for the peaking background were changed.The filled region is the expected statistical uncertainty of $B^0 \to \rho^+\rho^-$, and the black point shows the observed CPV parameters in the TOY study.

Parameter	Systematic uncertainty
S_{CP}	$^{+0.038}_{-0.036}$
A_{CP}	$^{+0.037}_{-0.041}$

TABLE 73. Systematic uncertainties due to the CPV in background modes.

8.23. CP violation in the transverse signals

We considered the possible CP violation in the transversely polarized $B^0 \to \rho^+\rho^-$ signals. The direct and indirect CP violating effect is assumed to be canceled by the equal contributions from the CP even and CP odd states in transversely polarized signals. We estimated a possible shift in the measured CP violation because of the shift by performing fittings to TOY MC samples assuming $\pm 50\%$ of the direct and indirect CP violation in the transverse signals. Figure 108 shows the result of the TOY studies, and we observed a much smaller shift compared with the statistical uncertainties.

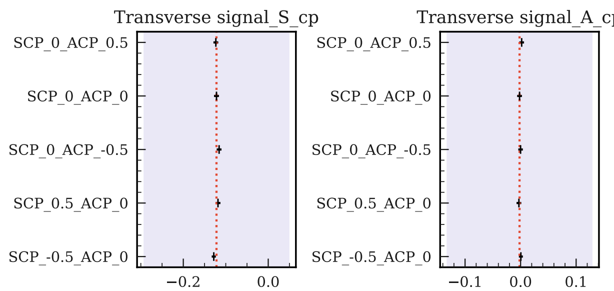


FIG. 108. Observed shift in the TOY study varying CP violating parameters of the transverse signals by $\pm 50\%$. The purple filled region corresponds to the statistical uncertainties for S_{CP} and A_{CP}

8.24. Correlation of the systematic uncertainty

The correlation of the systematic uncertainties for 4 observables is estimated by measuring the correlation coefficient of each systematic uncertainty source and combining them while considering the correlation. Firstly, the correlation coefficient for each systematic uncertainty source, for each pair of the parameters is calculated. For example, we varied the PDF shape parameters individually and measured the shift and the correlation in the four observables to estimate the correlation of the systematic uncertainties from PDF shape parameters. The same procedure is applied for all combinations of observables and all systematic uncertainties which are shown in Table 74

To combine each systematic uncertainty considering correlation, we generated a random number with the same distribution as the uncertainty.

The probability distribution for the bias due to the systematic uncertainty is

$$f(\vec{x}) = \frac{exp(-1/2\vec{x}^T \Sigma^{-1}\vec{x})}{\sqrt{(2\pi)^2|\Sigma|}}$$

where $\vec{x} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, v_1, v_2 are systematic uncertainties of the observables, Σ is covariance matrix which can be determined using the size of the each systematic uncertainties and the correlation coefficient shown in Table 74. We generated $\vec{x}$ randomly using the probability distribution for each systematic uncertainty source and summed them up. Figure 109 and Figure 110 show the correlation of the systematic uncertainty and the sum of the statistic and

	$\mathcal{B}$ - f_L	S_{CP} - A_{CP}	$\mathcal{B}$ - S_{CP}	$\mathcal{B}$ - A_{CP}	f_L-S_{cp}	f_L-A_{CP}
Resolution	0.00	0.56	0.00	0.00	0.00	0.00
Δt PDF for $q\bar{q}$ and $B\bar{B}$	0.00	-0.73	0.00	0.00	0.00	0.00
Physics Parameters	0.00	0.99	0.00	0.00	0.00	0.00
Interference	-0.67	-0.18	0.18	0.01	-0.20	0.06
Tag side interference	0.00	-1.00	0.00	0.00	0.00	0.00
Wrong tag fraction	0.00	0.48	0.00	0.00	0.00	0.00
Tracking	0.00	0.00	0.00	0.00	0.00	0.00
π^0 eff.	0.00	0.00	0.00	0.00	0.00	0.00
PID	0.00	0.00	0.00	0.00	0.00	0.00
$\mathcal{T}_C$	0.00	0.00	0.00	0.00	0.00	0.00
Data-MC mis-modeling	-0.18	-0.11	-0.30	-0.09	0.11	-0.25
$\mathcal{B}$ of peaking backgrounds	0.42	0.49	0.85	0.67	0.23	0.29
Best candidate selection	-0.40	0.04	-0.16	0.17	0.05	0.10
SxF ratio	-1.00	-1.00	1.00	-1.00	-1.00	1.00
Signal model	-0.96	0.94	0.99	0.94	-0.95	-0.92
$q\bar{q}$ model	0.16	-0.89	0.14	-0.09	-0.00	-0.02
$B\bar{B}$ model	0.27	0.39	0.10	0.03	0.07	0.08
$\tau\tau$ model	-0.90	0.47	0.68	0.05	-0.50	0.03
Peaking model	0.35	0.06	0.48	-0.20	0.25	-0.20
$f+-/f_{00}$	0.00	0.00	0.00	0.00	0.00	0.00
N_{BB}	0.00	0.00	0.00	0.00	0.00	0.00
MC stat.	0.00	0.00	0.00	0.00	0.00	0.00
Fit bias	-0.46	-0.01	-0.00	-0.03	0.00	0.03
Mis-Alignment	0.00	-1.00	0.00	0.00	0.00	0.00
Background CP Violation	0.00	0.01	0.00	0.00	0.00	0.00
$\tau\tau$ background yield	-1.00	1.00	1.00	1.00	-1.00	-1.00

TABLE 74. Correlation of the systematic uncertainties

systematic uncertainties, respectively. Table 75 shows the correlation of the four observables.

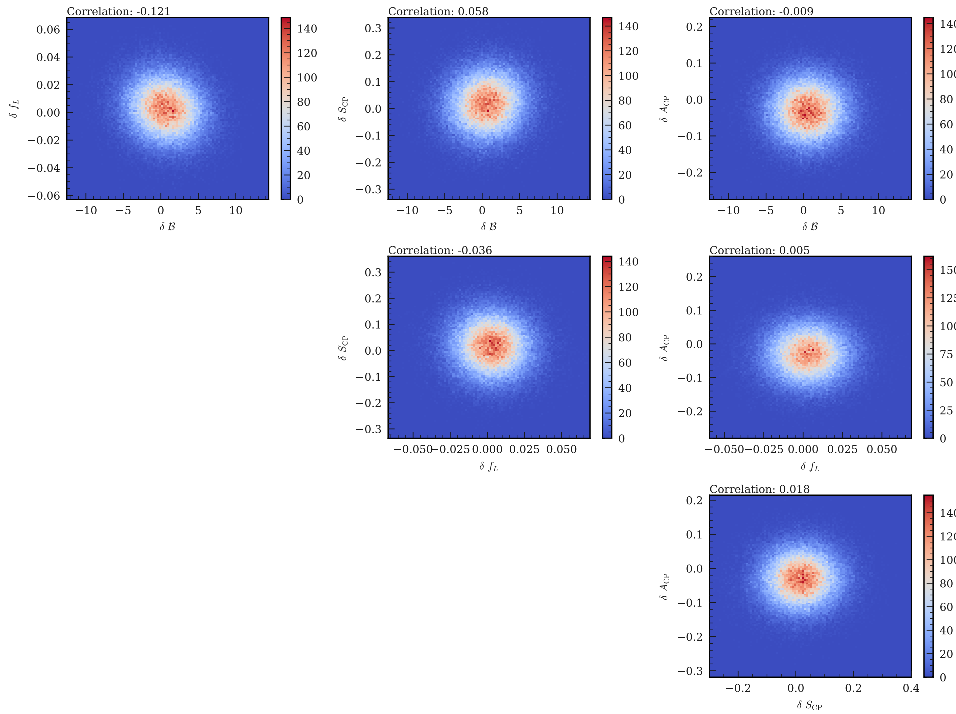


FIG. 109. Correlation of the systematic uncertainties

TABLE 75. The correlation matrix of four observables.

Stat.	f_L	S_{CP}	A_{CP}
$\mathcal{B}$	-0.11	0.04	-0.01
f_L		0.03	-0.05
S			0.06
Syst.	f_L	S_{CP}	A_{CP}
$\mathcal{B}$	-0.12	0.06	-0.01
f_L		-0.03	-0.00
S			0.02
Stat. + Syst.	f_L	S_{CP}	A_{CP}
$\mathcal{B}$	-0.11	0.04	0.00
f_L		0.02	-0.04
S			0.06

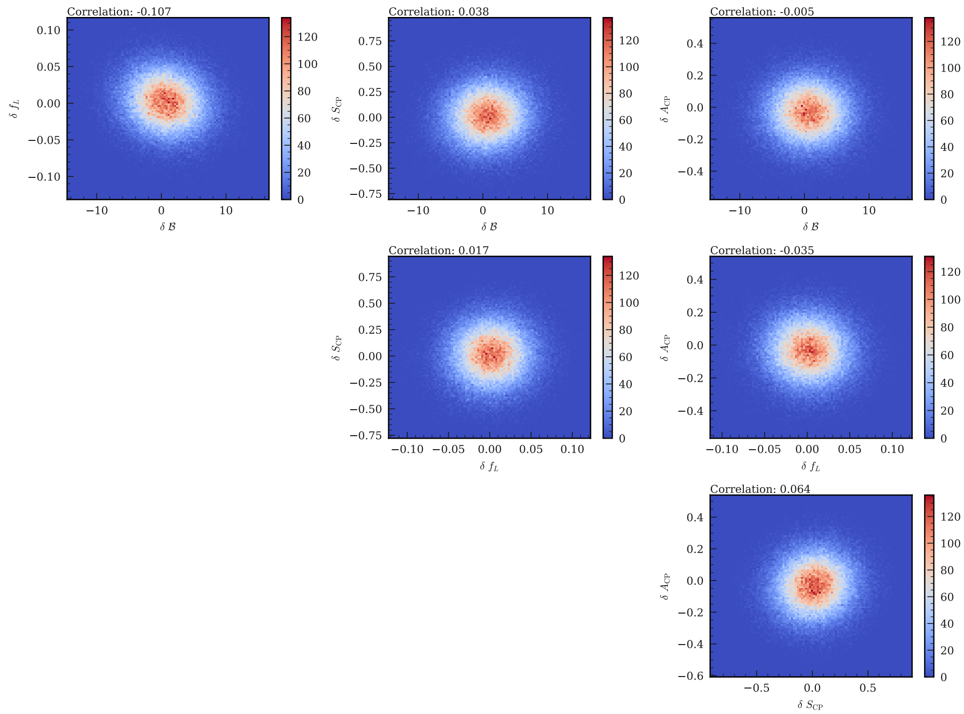


FIG. 110. Correlation of the sum of the statistical and systematic uncertainties

9. DISCUSSION

9.1. Improvement from the previous experiments

Table 76 shows the comparison with the previous experiments. Looking at $\mathcal{B}$ and f_L, our sensitivity was better than that of BaBar experiments which used a similar number of BB pairs.The sensitivity was slightly worse than that of the Belle experiment. There are several reasons, but the largest one is the number of the floating peaking background. There are many peaking background modes whose branching fractions are measured. They were fixed with some assumptions in the past experiments because extracting their yields at the same time is technically difficult and it makes the uncertainty larger. However, in our case, we improved the background model and fit procedure and extracted all possible peaking background yields from the fit. It makes the uncertainty a little larger, however, it is a more conservative way. Despite the larger uncertainty of the conservative peaking background modeling, we improved the statistical precision by significant improvements in the event selection. We reduced the background largely using a machine learning algorithm that didn't exist at the time of the previous experiments. It reduced the number of background events in the fit region by a factor of ten or more compared to the previous experiments. It also reduced the mis-reconstructed signal fraction to 35%. It was 64% in the Belle experiments, and 51% in the BaBar experiments. We improved our analysis to not depend on the assumption in estimating the effects from peaking backgrounds or mis-reconstructed yields. At the same time, the statistical precision is improved thanks to the advanced event selection using machine learning.

The impact of the conservative treatment on the peaking background yields is not so large in the CP violation measurement, so the precision is improved by the advanced event selection. The statistical precision scaled by the statistics of the CP violating parameters was improved by 10% compared to the Belle experiment.

	$\mathcal{B}$	f_L	$S_{\rm CP}$	$A_{\rm CP}$	$N_{BB}(\times 10^6)$
Belle II	$29.1^{+2.3+3.1}_{-2.2-3.0}$	$0.921^{+0.024+0.017}_{-0.025-0.015}$	$-0.26 \pm 0.19 \pm 0.08$	$0.02 \pm 0.12^{+0.05}_{-0.06}$	388
Belle	$28.3 \pm 1.5 \pm 1.5$	$0.988 \pm 0.012 \pm 0.006$	$-0.13 \pm 0.15 \pm 0.005$	$0.0 \pm 0.1 \pm 0.06$	772
BaBar	$25.5 \pm 2.1^{+3.6}_{-3.9}$	$0.992 \pm 0.024^{+0.026}_{-0.013}$	$-0.17 \pm 0.20^{+0.05}_{-0.06}$	$-0.01 \pm 0.15 \pm 0.06$	383

TABLE 76. Comparison of the results with the previous experiments.

9.2. ϕ_2 extraction

9.2.1. Strategy to extract ϕ_2

Firstly, we need to develop and validate a method to extract ϕ_2. In the development, we use the world average values for the inputs and confirm that the calculated ϕ_2 is consistent with the existing results, for example, the PDG value, the Belle results, and the BABAR results. Here, GammaCombo is applied for mathematical treatments, and its calculation results will undergo a consistency check. In addition, we also develop our original tool for the ϕ_2 extraction for understanding its procedure well and cross-checking. After the cross-checks of the calculation results, we measure ϕ_2 using the new $\mathcal{B}$, f_L, $S_{\rm cp}$, $A_{\rm cp}$ for

$B^0 \to \rho^+\rho^-$. In this step, the inputs of the ϕ_2 calculator except for $B^0 \to \rho^+\rho^-$ are set to the world average.

9.2.2. *Isospin Analysis*

The $B \to \rho\rho$ decay has contributions from both tree and penguin amplitudes. Due to the penguin contribution in $B^0 \to \rho^+\rho^-$ decay, the time-dependent CP-violating parameter $S_{\rm CP}$ is polluted, and the direct CP-violating parameter $A_{\rm CP}$ can appear. It is called "Penguin pollution". To remove the effect of the penguin pollution, we extract ϕ_2 using an isospin analysis [21]. A $\rho\rho$ system can have a state with isospin $I = 2$ or $I = 0$. The tree decay can lead to either $I = 2$ or $I = 0$ states. On the other hand, the gluonic penguin decay only allows decay to the $I = 0$ state. Thus, the $I = 2$ state has only a tree contribution, while the $I = 0$ state has both tree and penguin contributions. The $\rho^\pm\rho^0$ state has only $I = 2$, so it has contributions only from the tree decay. The other two decays ($\rho^+\rho^-, \rho^0\rho^0$) can result in either $I = 2$ or $I = 0$ states, leading to contributions from both tree and penguin amplitudes. The relationships of amplitudes among $B \to \rho^+\rho^-$, $B \to \rho^0\rho^0$, and $B \to \rho^\pm\rho^0$, are given by

$$\begin{aligned} \mathcal{A}_{+-} + \mathcal{A}_{00} &= \mathcal{A}_{+0} \\ \bar{\mathcal{A}}_{+-} + \bar{\mathcal{A}}_{00} &= \bar{\mathcal{A}}_{+0}, \end{aligned} \tag{50}$$

where $\mathcal{A}_{ij}$ is the amplitude of $B \to \rho^i\rho^j$, and $i, j = +, -, 0$. These relationships can be shown in triangles, as illustrated in Figure 111. $\Delta\phi_2$ can be extracted by using the angle between the two isospin triangles.

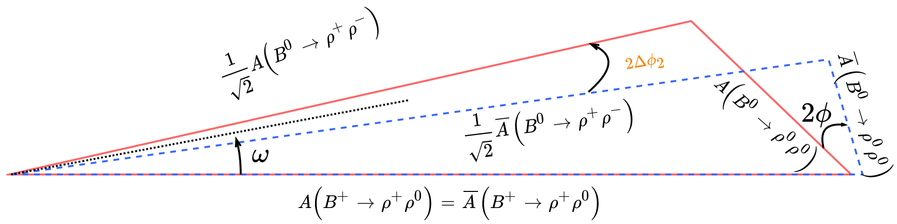


FIG. 111. Isospin triangle of $B \to \rho\rho$.

9.2.3. Procedure to extract ϕ_2

Likelihood

To extract ϕ_2, we use the $\mathcal{B}$, f_L, $S_{\rm CP}$, and $A_{\rm CP}$ of $B \to \rho^+\rho^-$ and $B \to \rho^0\rho^0$, and $\mathcal{B}$ and f_L, of $B^\pm \to \rho^\pm\rho^0$. Using the amplitudes, the observables can be written as

$$
\begin{aligned}
\mathcal{B}(\rho^+\rho^-) =& \tau_{B^0}\frac{\mathcal{A}_{+-}^2 + \bar{\mathcal{A}}_{+-}^2}{2} \\
\mathcal{B}(\rho^+\rho^0) =& \tau_{B^\pm}\mathcal{A}_{+0}^2 \\
\mathcal{B}(\rho^0\rho^0) =& \tau_{B^0}\frac{\mathcal{A}_{00}^2 + \bar{\mathcal{A}}_{00}^2}{2} \\
\mathcal{S}_{\rm cp}(\rho^+\rho^-) =& \sqrt{1 - \mathcal{A}_{\rm CP}(\rho^+\rho^-)^2}\sin(2\phi_2 + 2\Delta\phi_2) \\
\mathcal{S}_{\rm cp}(\rho^0\rho^0) =& \sqrt{1 - \mathcal{A}_{\rm CP}(\rho^0\rho^0)^2}\sin(2\phi_2 + 2\phi)
\end{aligned}
\tag{51}
$$

where $\mathcal{B}(\rho^i\rho^j)$ is the branching fraction of the longitudinaly polarized component of $B \to \rho^i\rho^j$, τ_B is the lifetime of $B^\pm$ or B^0, ϕ_2 is the angle in the unitarity triangle, and $\Delta\phi_2$ is the shift in ϕ_2 due to the penguin pollution. The parametrizations in the fitting are given by

$$
\begin{aligned}
\mathcal{A}_\pm =& \sqrt{a_{+-}(1 - \mathcal{A}_{\rm CP}(\rho^+\rho^-))} \\
\bar{\mathcal{A}}_\pm =& \sqrt{a_{+-}(1 + \mathcal{A}_{\rm CP}(\rho^+\rho^-))} \\
\mathcal{A}_{+0} =& \sqrt{a_{+0}} \\
\bar{\mathcal{A}}_{+0} =& \sqrt{a_{+0}} \\
\mathcal{A}_{00} =& \sqrt{\mathcal{A}_{+-}^2/2 + \mathcal{A}_{+0}^2 - \sqrt{2}\mathcal{A}_{+-}\mathcal{A}_{+0}\cos(\omega - \Delta\phi_2)} \\
\bar{\mathcal{A}}_{00} =& \sqrt{\bar{\mathcal{A}}_{+-}^2/2 + \bar{\mathcal{A}}_{+0}^2 - \sqrt{2}\bar{\mathcal{A}}_{+-}\bar{\mathcal{A}}_{+0}\cos(\omega + \Delta\phi_2)}
\end{aligned}
\tag{52}
$$

where a_{+-}, a_{+0} are the square root of the length of the $\mathcal{A}_{+-}$ and $\mathcal{A}_{+0}$, respectively, and ω and $\Delta\phi_2$ are the angles shown in Figure 111. We calculated the observables using Equation 51 and Equation 52 using a_{+-}, a_{+0}, ω, ϕ_2, $\Delta\phi_2$, $\mathcal{A}_{\rm CP}(\rho^+\rho^-)$, $\mathcal{A}_{\rm CP}(\rho^0\rho^0)$ as free parameters. Equation 53 shows the likelihood in the fitting, where $\Delta\vec{x} \equiv \vec{x}_{\rm obs} - \vec{x}_{\rm pred}$.
$\vec{x}_{\rm obs} = (\mathcal{B}(\rho^+\rho^-), \mathcal{B}(\rho^+\rho^0), \mathcal{B}(\rho^0\rho^0), \mathcal{S}_{\rm CP}(\rho^+\rho^+), \mathcal{S}_{\rm CP}(\rho^0\rho^0), \mathcal{A}_{\rm CP}(\rho^+\rho^+), \mathcal{A}_{\rm CP}(\rho^0\rho^0))$ using the measured one, and $\vec{x}_{\rm pred}$ is one from calculation.

$$\chi^2 \equiv -2\log\mathcal{L} = \Delta\vec{x}\Sigma^{-1}\Delta\vec{x}^{-1} \tag{53}$$

In the fitting, we minimized the likelihood using Minuit and obtained the ϕ_2 value and the uncertainty. In addition, we fixed ϕ_2 and minimized the likelihood by other floating parameters while changing ϕ_2 by 0.1°. Then, we calculated the confidence interval by integrating degree-of-freedom 1 chi-square distribution from $\Delta\chi^2$ to inf for each ϕ_2, where $\Delta\chi^2 \equiv \chi^2 - \min(\chi^2)$.

Validation using Belle result

To validate the procedure of ϕ_2 extraction shown in Section 9.2.3, we performed ϕ_2 extraction using the ϕ_2 measurement performed in Belle [22]. Table 77 shows the input which were used the ϕ_2 extraction, and the extracted ϕ_2 was $\phi_2 = (93.7 \pm 10.6)^\circ$.

Parameter	Value
$\mathcal{B}(B^0 \to \rho^+\rho^-)$	$(28.3 \pm 1.5 \pm 1.5) \times 10^{-6}$
$f_L(B^0 \to \rho^+\rho^-)$	$0.988 \pm 0.012 \pm 0.023$
$\mathcal{B}(B^\pm \to \rho^\pm\rho^0)$	$(31.7 \pm 8.8) \times 10^{-6}$
$f_L(B^\pm \to \rho^\pm\rho^0)$	0.95 ± 0.11
$\mathcal{B}(B^0 \to \rho^0\rho^0) \times f_L(B^0 \to \rho^0\rho^0)$	$(0.21 \pm 0.34) \times 10^{-6}$
$\mathcal{S}_{\rm CP}(B^0 \to \rho^+\rho^-)$	$-0.13 \pm 0.15 \pm 0.05$
$\mathcal{A}_{\rm CP}(B^0 \to \rho^+\rho^-)$	$0.00 \pm 0.10 \pm 0.06$
$\tau_{B^\pm}/\tau_{B^0}$	1.076

TABLE 77. Inputs for the ϕ_2 extraction performed at Belle

We performed the crosscheck by the procedure developed within the framework of GammaCombo [23] and by the self-developed tools. Figure 112 and Figure 113 show the ϕ_2 extraction results using GammaCombo and self-developed tool, respectively. Both results show a good agreement with Belle's result. In the ϕ_2 extraction using the self-developed tool, the uncertainty estimated by Hesse shows a good agreement, and the uncertainty was reduced a little by using Minos error estimation. It it due to the approximation in the Hesse, and in principle, Minos error estimation is more appropriate because it does a likelihood scan in reality. This study demonstrated the reliability of the ϕ_2 extraction, and we will use Minos error estimation for a better error estimation.

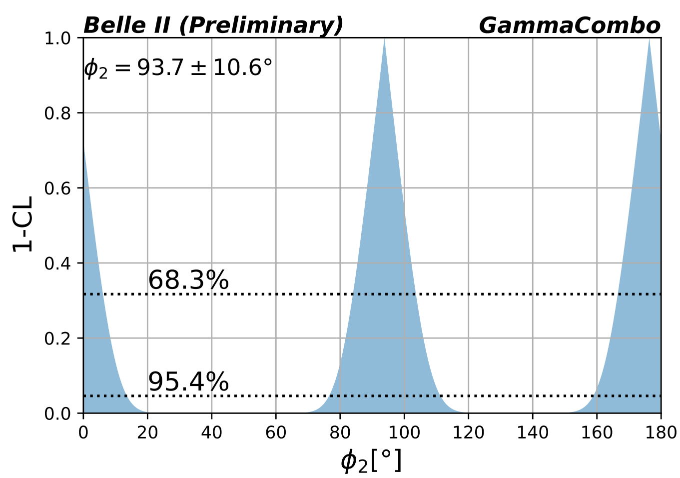


FIG. 112. ϕ_2 extraction using the same inputs as belle with GammaCombo

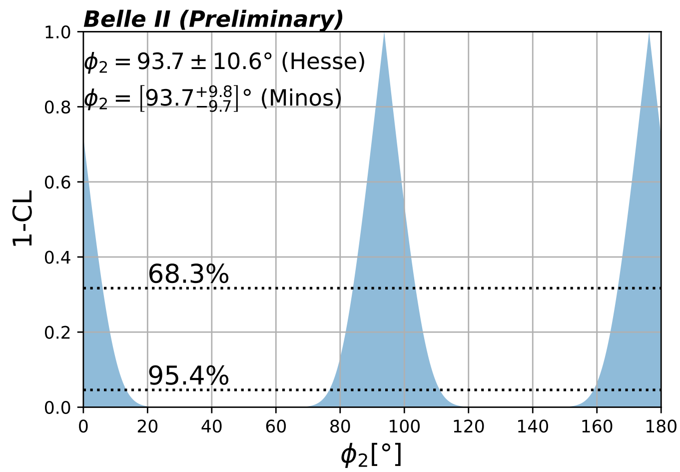


FIG. 113. ϕ_2 extraction using the same inputs as belle with self-developed tools

9.2.4. ϕ_2 extraction

The ϕ_2 extraction performed at Belle was limited by the precision of the branching fraction and longitudinal polarization measurement of $B^{\pm} \to \rho^{\pm}\rho^0$, and also time-dependent CP violation measurement of $B^0 \to \rho^0\rho^0$ was missing. To maximize the precision of the ϕ_2 measurement and demonstrate the impact of the new result from the Belle II, we plan to extract ϕ_2 using a combination of Belle II and world average.

ϕ_2 extraction using world average

Before going to the combination of the Belle II and world average, we calculated ϕ_2 using the world average of $B \to \rho\rho$ measurement. Figure 78 shows the inputs. Figure 114 shows the result of the ϕ_2 extraction using teh world average, and ϕ_2 was determined to be $\phi_2 = (91.9^{+4.7}_{-4.9})^{\circ}$.

Parameter	Value
$\mathcal{B}(B^0 \to \rho^+\rho^-)$	$(27.76 \pm 1.84) \times 10^{-6}$
$f_L(B^0 \to \rho^+\rho^-)$	0.99 ± 0.02
$\mathcal{B}(B^{\pm} \to \rho^{\pm}\rho^0)$	$(24.9 \pm 1.9) \times 10^{-6}$
$f_L(B^{\pm} \to \rho^{\pm}\rho^0)$	0.950 ± 0.016
$\mathcal{B}(B^0 \to \rho^0\rho^0)$	$(0.93 \pm 0.14) \times 10^{-6}$
$f_L(B^0 \to \rho^0\rho^0)$	0.71 ± 0.06
$\mathcal{S}_{\mathrm{CP}}(B^0 \to \rho^+\rho^-)$	-0.15 ± 0.13
$\mathcal{A}_{\mathrm{CP}}(B^0 \to \rho^+\rho^-)$	0.00 ± 0.09
$\mathcal{S}_{\mathrm{CP}}(B^0 \to \rho^0\rho^0)$	0.3 ± 0.7
$\mathcal{A}_{\mathrm{CP}}(B^0 \to \rho^0\rho^0)$	0.2 ± 0.9
$\tau_{B^{\pm}}/\tau_{B^0}$	1.076 ± 0.004

TABLE 78. Inputs for the ϕ_2 extraction using world average

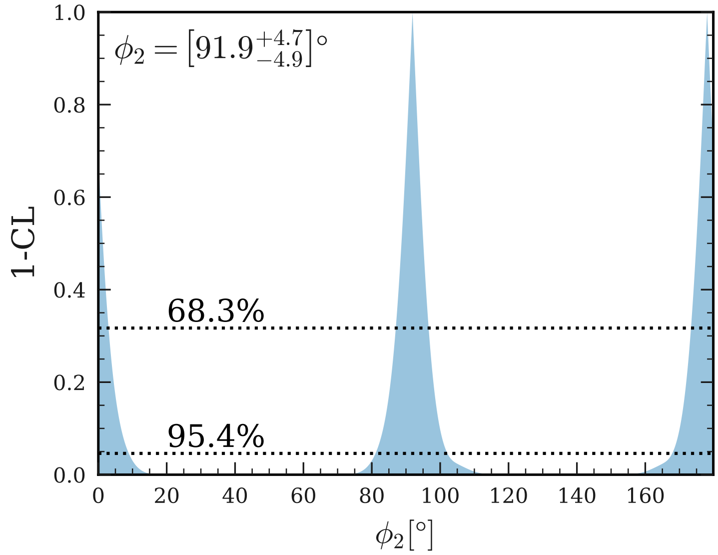


FIG. 114. ϕ_2 extraction using the world average of $B \to \rho\rho$

Branching fraction correction using updated f_{+-}/f_{00}

It was assumed that the branching fractions of $\Upsilon(4S) \to B^0\bar{B}^0, B^+B^-$ are the same in the previous experiments. In 2023, the Belle experiment measured [24]

$$f_{+-}/f_{00} = \frac{\Gamma(\Upsilon(4S) \to B^+B^-)}{\Gamma(\Upsilon(4S) \to B^0\bar{B}^0)} = 1.065 \pm 0.012(Stat.) \pm 0.051(Syst.).$$

Considering the result, the branching fraction of $B^+ \to \rho^+\rho^0$ increases and $B^0 \to \rho^+\rho^-, \rho^0\rho^0$ decreases, which leads to a larger $\Delta\phi_2$. We corrected the branching fractions considering the value and uncertainty of f_{+-}/f_{00} as shown in Table 80.

Figure 115 shows the ϕ_2 extraction using the corrected world average, the extracted ϕ_2 is

$$\phi_2 = 91.5^{+4.9}_{-5.4}.$$

The precision of the ϕ_2 extraction got worse by the correction, because of the larger $\Delta\phi_2$ as expected. The effect was ignored in the previous experiments and it is necessary to include them in the ϕ_2 extraction.

Parameter	Value
$\mathcal{B}(B^0 \to \rho^+\rho^-)$	$(28.6 \pm 1.9) \times 10^{-6}$
$f_L(B^0 \to \rho^+\rho^-)$	0.99 ± 0.02
$\mathcal{B}(B^\pm \to \rho^\pm\rho^0)$	$(23.3 \pm 1.8) \times 10^{-6}$
$f_L(B^\pm \to \rho^\pm\rho^0)$	0.950 ± 0.016
$\mathcal{B}(B^0 \to \rho^0\rho^0)$	$(0.98 \pm 0.16) \times 10^{-6}$
$f_L(B^0 \to \rho^0\rho^0)$	0.71 ± 0.06
$\mathcal{S}_{\rm CP}(B^0 \to \rho^+\rho^-)$	-0.15 ± 0.13
$\mathcal{A}_{\rm CP}(B^0 \to \rho^+\rho^-)$	0.00 ± 0.09
$\mathcal{S}_{\rm CP}(B^0 \to \rho^0\rho^0)$	0.3 ± 0.7
$\mathcal{A}_{\rm CP}(B^0 \to \rho^0\rho^0)$	0.2 ± 0.9
$\tau_{B^\pm}/\tau_{B^0}$	1.076 ± 0.004

TABLE 79. Corrected inputs for the ϕ_2 extraction using the world average by f_{+-}/f_{00} results by the Belle experiment. The uncertainty of the f_{+-}/f_{00} is included in the uncertainties of the branching fractions.

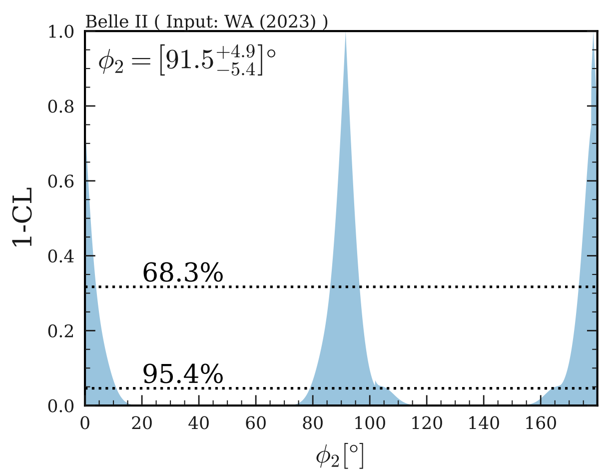


FIG. 115. ϕ_2 extraction using the corrected world average of $B \to \rho\rho$ by $f+-/f_{00}$ result by the Belle experiment

ϕ_2 extraction using the combination of the world average and Belle II

We combined our results

$$\begin{aligned}
\mathcal{B}(B^0 \to \rho^+\rho^-) &= \left(2.94^{+0.23+0.31}_{-0.22-0.30}\right) \times 10^{-5}, \\
f_L &= 0.921^{+0.024+0.017}_{-0.025-0.015}, \\
S &= -0.26 \pm 0.19 \pm 0.08, \\
A &= 0.02 \pm 0.12^{+0.05}_{-0.06}
\end{aligned}$$

and the world average corrected by the f_{+-}/f_{00} to extract ϕ_2. Figure 116 shows the result of the ϕ_2 extraction using the combination of the world average and Belle II. We obtain $\phi_2 = [92.6^{+4.5}_{-4.8}]^\circ$ which has about 10% better precision compared with the ϕ_2 extraction using the corrected world average ($\phi_2 = [91.5^{+4.9}_{-5.4}]^\circ$)

Parameter	Value
$\mathcal{B}(B^0 \to \rho^+\rho^-)$	$(28.0 \pm 1.8) \times 10^{-6}$
$f_L(B^0 \to \rho^+\rho^-)$	0.97 ± 0.017
$\mathcal{B}(B^\pm \to \rho^\pm\rho^0)$	$(23.3 \pm 1.8) \times 10^{-6}$
$f_L(B^\pm \to \rho^\pm\rho^0)$	0.950 ± 0.016
$\mathcal{B}(B^0 \to \rho^0\rho^0)$	$(0.98 \pm 0.16) \times 10^{-6}$
$f_L(B^0 \to \rho^0\rho^0)$	0.71 ± 0.06
$\mathcal{S}_{\rm CP}(B^0 \to \rho^+\rho^-)$	-0.17 ± 0.11
$\mathcal{A}_{\rm CP}(B^0 \to \rho^+\rho^-)$	-0.01 ± 0.07
$\mathcal{S}_{\rm CP}(B^0 \to \rho^0\rho^0)$	0.3 ± 0.7
$\mathcal{A}_{\rm CP}(B^0 \to \rho^0\rho^0)$	0.2 ± 0.9
$\tau_{B^\pm}/\tau_{B^0}$	1.076 ± 0.004

TABLE 80. Inputs for the ϕ_2 extraction using the combination of the corrected world average by f_{+-}/f_{00} and $B^0 \to \rho^+\rho^-$ results from this analysis.

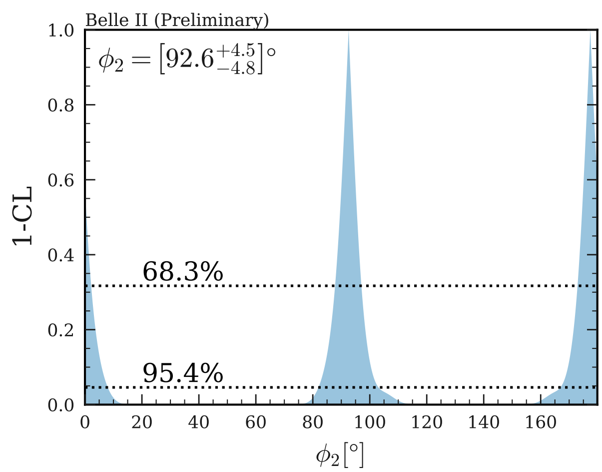


FIG. 116. ϕ_2 extraction using the combination of the world average and Belle II

9.3. Future prospects

We estimate the future sensitivity of the $B^0 \to \rho^+\rho^-$ analysis and its impact on the ϕ_2 precision. In the $\mathcal{B}$ and f_L, its statistical uncertainties are dominated by the π^0 efficiency that is due to the branching fraction uncertainty of the control mode. This is not expected to be improved in the future Belle II experiment because it is a control mode in the experiment. In the $S_{\rm CP}$ and $A_{\rm CP}$, the dominant systematic uncertainties are due to the background Δt PDF and Δt resolution function parameters. They are limited by statistical uncertainties and can be improved with statistics. Another large systematic uncertainty comes from the CP violation in the background B decays. Some of them can be improved by statistics. We assume no systematic improvement on $\mathcal{B} \times f_L$, and the systematic uncertainty on $S_{\rm CP}$ and $A_{\rm CP}$ can be improved by statistics considering the above. The estimated sensitivity of the $B^0 \to \rho^+\rho-$ observables are listed in Table 81.

0.365 ab^{-1}		
Paramaeter	value	uncertainty
$\mathcal{B} \times f_L$	27.720	1.738
$S_{\rm cp}$	-0.174	0.101
$A_{\rm cp}$	-0.007	0.071
1.0 ab^{-1}		
Paramaeter	value	uncertainty
$\mathcal{B} \times f_L$	27.720	1.669
$S_{\rm cp}$	-0.174	0.077
$A_{\rm cp}$	-0.007	0.055
10.0 ab^{-1}		
Paramaeter	value	uncertainty
$\mathcal{B} \times f_L$	27.720	1.614
$S_{\rm cp}$	-0.174	0.030
$A_{\rm cp}$	-0.007	0.021
50.0 ab^{-1}		
Paramaeter	value	uncertainty
$\mathcal{B} \times f_L$	27.720	1.608
$S_{\rm cp}$	-0.174	0.014
$A_{\rm cp}$	-0.007	0.010

TABLE 81. Expected sensitivity of the $B^0 \to \rho^+\rho^-$ observables.

We estimate the precision of ϕ_2 under this assumption, the results are shown in Figure 117. We expect to reach about $\pm 3.3[^\circ]$ of the precision by the future $B^0 \to \rho^+\rho^-$ analysis at the Belle II experiment. The sensitivity of the $B^0 \to \rho^+\rho^-$ observables was estimated by Ref. [25]. Our estimation is more powerful than the expectation in this reference because of the improvement in the statistical and systematic precision. The uncertainty in the CP violating parameter in our analysis ± 0.19 for S_{CP} and ± 0.12 for A_{CP}. They become ± 0.016 and ± 0.010 scaling by $\frac{1}{\sqrt{50\ \rm ab^{-1}/365\ fb^{-1}}}$, which corresponds to factor 2 improvement from the expectation. We can also expect improvements in the systematic uncertainty because we reduced the fraction of the mis-reconstructed signals from the previous measurements as discussed in Sec. 9.1. There are similar improvements in the fraction of the contamination from the peaking background. The major systematic uncertainty is due to mis-reconstructed signals, CP violation in the B meson decays, and interference with other decays. These can be reduced by improving the purity of the signals, which was improved about factor two from previous measurements in our analysis.

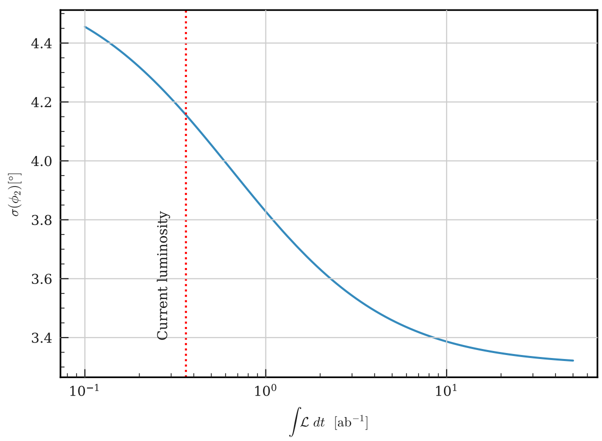


FIG. 117. Expected precision of ϕ_2 by the future improvement on $B^0 \to \rho^+\rho^-$ observables by the Belle II experiment.

Parameter	uncertainty
$\mathcal{B} \times f_L$	$\pm 0.4 \pm 2.6$
S_{CP}	$\pm 0.03 \pm 0.03$
A_{CP}	$\pm 0.02 \pm 0.03$

TABLE 82. The sensitivity of the $B^0 \to \rho^+\rho^-$ observables with 50 ab^{-1} data estimated by Ref. [25].

We also estimate the sensitivity on the ϕ_2 from $B \to \rho\rho$ measurements with the expected improvement in other measurements. We use the expected precision of other $B \to \rho\rho$ observables with 50 ab^{-1} data which was estimated by Ref. [25], shown in Table 83. We perform the isospin analysis using the expected sensitivity of $B^0 \to \rho^+\rho^-$ and values in Table 83, and estimate the precision of ϕ_2 to be $\pm 0.8°$.

Parameter	Uncertainty
$\mathcal{B}(\rho^+\rho^0)$	± 1.7
$\mathcal{B}(\rho^0\rho^0)$	± 0.13
$S_{CP}(\rho^0\rho^0)$	± 0.10
$A_{CP}(\rho^0\rho^0)$	± 0.09

TABLE 83. Expected precision of the $B \to \rho\rho$ measurements with 50 ab^{-1} data estimated by Ref. [25].

The precision of ϕ_2 using $B \to \rho\rho$,$B \to \pi\pi$, and $B \to \rho\pi$ decays are estimated to be $\pm 1.5°$, $\pm 3°$, and $\pm 1.5°$, and the precision of ϕ_2 combining all measurements is estimated to be $\pm 1°$**??**. Our analysis method improves the ϕ_2 precision using $B \to \rho\rho$ to be $\pm 0.8°$, and the precision combining all measurements to be $\pm 0.7°$.

9.4. Constraint on the BSM physics

The updated ϕ_2 is a useful observable to give a constraint on the BSM effects in the BSM model. The constraint on the BSM is performed by the CKM fitter group [26]. Using the same procedure, we estimate the impact of our analysis and possible improvement in the future.

We only consider the BSM effect in the loop process. Also, we only consider the BSM contribution in the B_d system. The variables used in the constraint are ϕ_1, ϕ_2, ϕ_3, $|V_{ud}|$, $|V_{us}|$, $|V_{cd}|$, $|V_{ub}|$, and Δm_d. ϕ_3, $|V_{ud}|$, $|V_{us}|$, $|V_{cd}|$, $|V_{ub}|$ are measured by the branching fraction or direct CP violation parameter in the tree process. They won't be affected by the BSM with our assumption. In contrast, ϕ_1, ϕ_2, Δm_d can be affected by the BSM contribution because they are measured by the $B^0\bar{B}^0$ mixing. We write the BSM contribution in B^0-$\bar{B}^0$ mixing as

$$\mathcal{A} = (1 - h_d \exp(2i\sigma_d))\mathcal{A}_{\mathrm{SM}} \equiv r_d^2 \exp(2i\theta_d)\mathcal{A}_{\mathrm{SM}}, \tag{54}$$

where $\mathcal{A}$ is the amplitude of the B^0-$\bar{B}^0$ mixing, $\mathcal{A}$ is the B^0-$\bar{B}^0$ mixing amplitude in the SM, h_d is the magnitude of the BSM to the SM, and σ_d is the relative weak phase of the BSM to the SM. Also, we use r_d and θ_d for an alternative parametrization of the equation for simplicity.

Using the parametrization, the observables can be written as in Table 84.

parameter	
ϕ_1	$\phi_{1\mathrm{SM}} + \theta_d$
ϕ_2	$\phi_{2\mathrm{SM}} - \theta_d$
ϕ_3	$\phi_{3\mathrm{SM}}$
$\|V_{ud}\|$	1-1/2λ^2
$\|V_{us}\|$	λ
$\|V_{cb}\|$	$A\lambda^2$
$\|V_{ub}\|$	$\|A\lambda^3(\rho - i\eta)\|$
Δm_d	$(0.502\ \mathrm{ps}^{-1})\left(\frac{\|V_{td}\|}{0.0086}\right)^2 \frac{f_{B_d}^2 \mathcal{B}_{B_d}}{(0.17\ \mathrm{GeV})^2} r_d^2$

TABLE 84. Parameterization of the observables considering BSM effects in theB^0-$\bar{B}^0$ mixing. λ, A, ρ, and η are the CKM parameters in Wolfenstein parameterization, and $\phi_{i\mathrm{SM}}$ is the ϕ_i calculated from λ, A, ρ, and η.

Using the parametrization above, we calculate the observables and calculate $\Delta\chi^2$ as

$$\Delta\chi^2 = \vec{x}\mathit{\Sigma}^{1}\vec{x}, \tag{55}$$

where $\mathit{\Sigma}$ is a covariance matrix of the observables and $\vec{x}$ is the difference of $\vec{x}_{\mathrm{obs}}$ and $\vec{x}_{\mathrm{theo}}$. $\vec{x}$ corresponds to observed values, and $\vec{x}_{\mathrm{theo}}$ corresponds to calculated observables using Table 84. In the SM, the parameter Δm_d which is the mass difference of the B^0 mesons is

$$\Delta m_d = \frac{G_F}{6\pi^2} f_{B_d}^2 \mathcal{B}_{B_d} \eta_B m_b m_W^2 S\left(\frac{m_t^2}{m_W^2}\right) |V_{tb}V_{td}^*|^2, \tag{56}$$

where G_F is Fermi constant, η_B is QCD correction factor, m_i are masses of i quarks, $S(x)$ is Inami-lim function, f_{B_q} is the decay constant, and $\mathcal{B}$ is the bag factor. We approximate the Δm_d as [**?**]

$$\Delta m_d = (0.502\ \mathrm{ps}^{-1})\left(\frac{|V_{td}|}{0.0086}\right)^2 \frac{f_{B_d}^2 \mathcal{B}_{B_d}}{(0.17\ \mathrm{GeV})^2}. \tag{57}$$

We then multiply r_d^2 which is possible BSM contribution to the parameter Δm_d in this assumption.

We calculate $\Delta\chi^2$ using λ, A, ρ, η, h_d, and σ_d as fit parameters and calculate the confidence interval in the two-dimensional plane of h_d and σ_d. We constrain the BSM effect using

parameter	
ϕ_1	$[22.84^{+0.45}_{-0.44}]^\circ$
ϕ_2	$[91.5^{+4.5}_{-5.4}]^\circ$
ϕ_3	$[67.1^{+3.3}_{-3.5}]^\circ$
$\lvert V_{ud}\rvert$	0.97373 ± 0.00053
$\lvert V_{us}\rvert$	0.21635 ± 0.00038
$\lvert V_{cb}\rvert$	0.0422 ± 0.00044
$\lvert V_{ub}\rvert$	0.00382 ± 0.0002
Δm_d	0.5065 ± 0.0033
$\mathcal{B}_d$	0.856 ± 0.026
f_{B_d}	0.189 ± 0.003

TABLE 85. Observables used for the fit. Observables except for ϕ_2 are cited from [3], and ϕ_2 is the value calculated by us using the world average of the observables related to $B \to \rho\rho$ decays.

the world average and the combination of the world average and this analysis. Figure 118 shows the constraint on the BSM effect in two scenarios. The precision of the ϕ_2 is improved by 10 %, but the excluded area is not changed so much because of the shift in the central value of the ϕ_2. The new constraint gives 5% stronger constraint around $\sigma_d = 90^\circ$ and slightly increases the excess around $\sigma_d = 180^\circ$ by 8%.

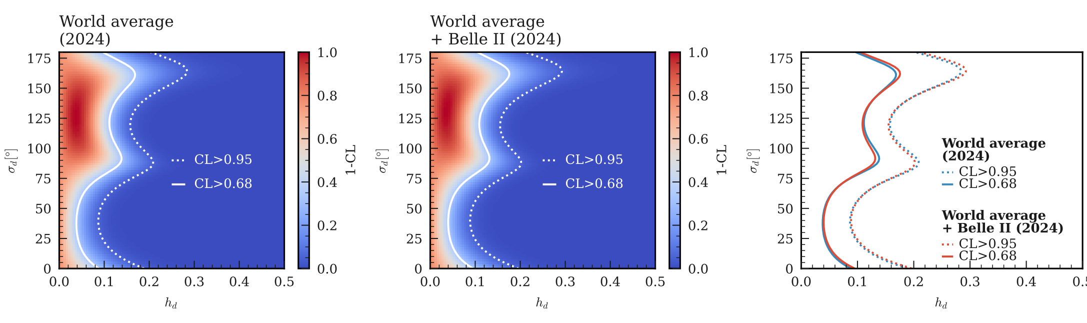


FIG. 118. Constraint to the BSM effects in B^0-$\bar{B}^0$ mixing. The left one is the current world average without Belle II, the middle one is the combined result of the world average and Belle II, and the right one is the comparison. The solid line is the 68% confidence interval, and the dotted line corresponds to the 95% confidence interval.

We also estimate the expected improvement sensitivity of the BSM constraint with 50 ab^{-1} data. We cite the expected sensitivities and values assuming no BSM contributions, which is summarized by Ref. [4], shown in Table 86. The ϕ_2 value is expected to be improved from $\pm 1^\circ$ to $\pm 0.7^\circ$ as discussed in Sec.9.3, so we estimate the sensitivity in both cases. Figure 119 shows the result of the BSM constraint in two cases. The BSM constraint is expected to be improved by 3% by the 30% improvement in the ϕ_2 precision by this analysis.

Parameter	value
ϕ_1	$(21.38 \pm 0.31)^\circ$
ϕ_2	$(91.5 \pm 1.0)^\circ$
ϕ_3	$(67.1 \pm 1.0)^\circ$
$\lvert V_{ud}\rvert$	0.974 ± 0.00022
$\lvert V_{us}\rvert$	0.225 ± 0.0006
$\lvert V_{cb}\rvert$	0.042 ± 0.0004
$\lvert V_{ub}\rvert$	$(3.56 \pm 0.08) \times 10^{-3}$
Δm_d	0.507 ± 0.004
$\mathcal{B}_{B_d}$	0.756 ± 0.00753
f_{B_d}	0.193 ± 0.00124

TABLE 86. Values and uncertainties which are expected with 50 ab^{-1} data summarized by Ref. [4].

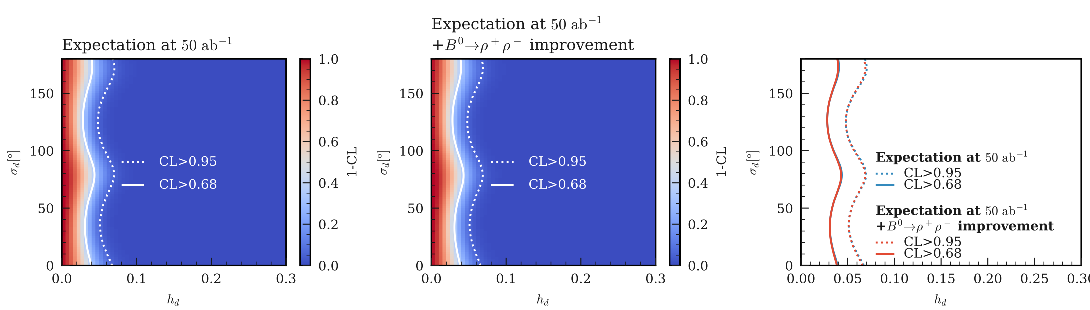


FIG. 119. Constraint to the BSM effects in B^0-$\bar{B}^0$ mixing using the expected sensitivity with 50 ab^{-1} data.

10. CONCLUSION

We measure the branching fraction, longitudinal polarization fraction, and time-dependent CP violation parameter $S_{\rm CP}$ and $A_{\rm CP}$ of $B^0 \to \rho^+\rho^-$ decay at the Belle II experiment. We used 365 fb^{-1} data that includes $(388 \pm 6) \times 10^6$ $B\bar{B}$ pairs, and we obtain

$$\mathcal{B}(B^0 \to \rho^+\rho^-) = \left(2.90^{+0.23+0.31}_{-0.22-0.29}\right) \times 10^{-5}, \tag{58}$$

$$f_L = 0.921^{+0.024+0.017}_{-0.025-0.015}, \tag{59}$$

$$S_{\rm CP} = -0.26 \pm 0.19 \pm 0.08, \tag{60}$$

$$A_{\rm CP} = -0.02 \pm 0.12^{+0.05}_{-0.06}. \tag{61}$$

Our result agrees with previous measurements. The precision of the analysis is improved by the newly implemented machine learning techniques for continuum and fake-photon suppression, and simultaneous cut-based selection optimization by the differential evolution technique. We then extracted ϕ_2, which is one of the parameters of the CKM unitarity triangle using the previous result for $B \to \rho\rho$ and our results. We obtain $\phi_2 = (91.5^{+4.5}_{-5.4})^\circ$ without our result, and $\phi_2 = (92.6^{+4.5}_{-4.8})^\circ$ with our result. The new result improves the precision of the global fit to $B \to \rho\rho$ for ϕ_2 extraction by 10%. We also discuss the impact of our analysis on the BSM constraint in B^0-$\bar{B}^0$ mixing. We perform fittings using the existing world averages and our results, and obtain 5% stronger constraint at $\sigma_d = 90^\circ$ which is the relative phase difference between SM and BSM. We also discuss the improvement of the future sensitivity in ϕ_2 measurements. Our analysis procedure will improve the precision of $B^0 \to \rho^+\rho^-$ observable by factor two, which corresponds to factor two improvement in ϕ_2 extraction in $B \to \rho\rho$ decay modes and 30% improvement in ϕ_2 extraction using $B \to \pi\pi, \rho\pi$, and $\rho\rho$ decay modes.

ACKNOWLEDGEMENT

I greatly appreciate Prof. Toru Iijima. He supervised me from when I started the research for high-energy physics. He showed me how interesting particle physics is and also helped me develop the basics of research. In addition, he gave me a lot of opportunities to do research and interact with other researchers. My research cannot be done without him.

I appreciate my co-analysts, Dr. Akimasa Ishikawa and Dr. Yu Nakazawa. I was helped a lot by their deep knowledge of particle physics and also helped from technical aspects a lot. Their contributions helped me a lot to overcome a lot of analysis challenges.

I greatly appreciate Dr. Kenji Inami. He supervised me in the TOP counter development and operations. I learned a lot from his deep knowledge of photodetectors which is essential for me to be an experimentation.

Dr. Kodai Matsuoka supervised me for the photodetector development. He was also an expert in detector development, and I learned experimental techniques a lot from deep knowledge. He continued to collaborate with me after leaving Nagoya University, and I greatly appreciate him.

I greatly appreciate Dr. Diego Tonelli, Dr. Mirco Dorigo, Dr. Thibaud Humair, Dr. Michele Veronesi, and Prof. Takeo Higuchi. They are the conveners of the analysis working groups, and their deep experience and knowledge helped to develop the difficult analysis.

I greatly appreciate Prof. Alan Schwartz, Dr. Riccardo Manfredi, Dr. Sagar Hazra, Prof. Yoshihide Sakai, and Prof. Thomas Browder. They reviewed our analysis and helped to improve the quality of the analysis. They also helped a lot in writing the papers.

In addition, I greatly appreciate Dr. Martin Bessner, Dr. Hulya Atmacan, and Dr. Yinghui Guan. They were the main members of the TOP counter operation. Thanks to their help, I contributed to the Belle II operation from the operation of the TOP counter.

I greatly appreciate Mr. Genta Muroyama, Dr. Noritsugu Tsuzuki, Dr. Kazuki Kojima, Ms. Akane Maeda, Mr. Tomoya Nakano, Mr. Yuya Adachi, Mr. Ryotaro Komori, and Mr. Koichi Ueda. They worked for photodetector development for me. Thanks to their collaboration, I could proceed with my detector developments.

Appendices

Appendix 1: Continuum suppression

Continuum impression MVA (CSMVA), that is MVA was used to suppress $e^+e^- \to q\bar{q}$ backgrounds ($q = u, d, s, c$). It was discussed in 3.4, and more detailed description is given in this section.
The MVA used in this analysis is TabNet [12], which is a kind of neural-network-based machine learning algorithm. In principle, neural networks perform better than other machine learning techniques because of an excellent expression supported by a large degree of freedom. However, in many cases, a decision-tree-based algorithm like FastBDT works better, especially in cases where the data is tabular-type, not picture or time-series. It is because of the lack of the amount of data. Neural networks require huge amounts of data to make a model larger and optimize its weight parameters. In contrast, the degree of freedom of decision-tree-based algorithms is much smaller than neural network models and can easily be trained. If the dataset is tabular, the required expression of the model is relatively smaller. The performance of it tends to depend on how the model was trained well for the datasets, not the model's expression. However, if a neural network model is trained well for tabular-type datasets, it can potentially perform better than a decision tree by learning complex correlations of the input features. To obtain a better performance of neural network models, specialized structure and limit the degree of the freedom of models. TabNet has an efficient structure, especially for tabular-type datasets, and it can be trained with relatively smaller datasets. It is why the TabNet was selected for the continuum suppression for this analysis.

The hyperparameter of the TabNet was tuned using Optuna [27], which is an algorithm for hyperparameter tuning for machine learning models using Bayesian optimization. Table 87 shows the hyperparameters that were tuned using Optuna. All were tuned simultaneously, and the minimizing target was set to the log loss of the validation data. Figure 120 shows the history of the hyperparameter tuning. The best parameter was taken and used in the other cut-based selection optimization or the later analysis.

Parameter	Description
n_d	Width of the prediction layer
n_a	Width of the attention embedding for each mask (Set same value as n_d)
n_{steps}	Number of steps in the architecture
γ	A coefficient for feature reusage in the masks.
λ_{sparse}	Sparsity loss coefficient
Mask function	A function for selecting features
Batch size	Number of samples per batch
Virtual batch size	Batch size for virtual batch normalization

TABLE 87. Hyperparameters for TabNet that were tuned

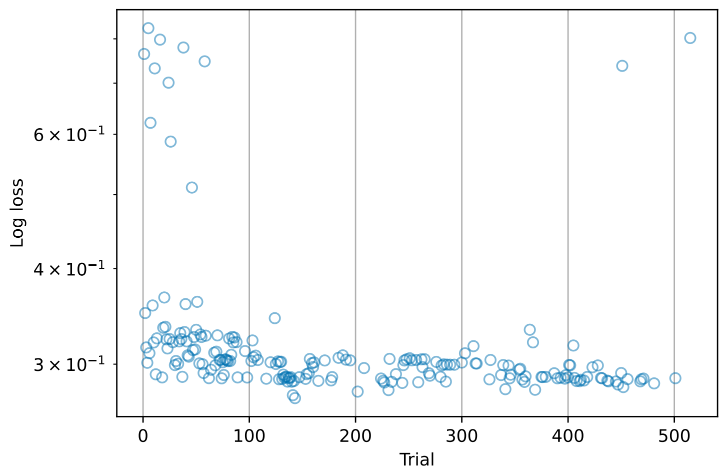


FIG. 120. Hyperparameter tuning for TabNet

Other machine-learning algorithms were also trained to confirm that TabNet shows better performance. In this comparison, FastBDT, LightGBM, and a conventional neural network, which is a multi-layer perceptron with three layers, were tested. Hyperparameters for them were tuned in the same way as TabNet. Figure 121 shows the result, and TabNet performs best. The performance improvement is 10% more background reduction in the fit region. It was tested using the dataset before the optimization of the cut-based selection, and then TabNet was tuned further while tuning the cut-based selection.

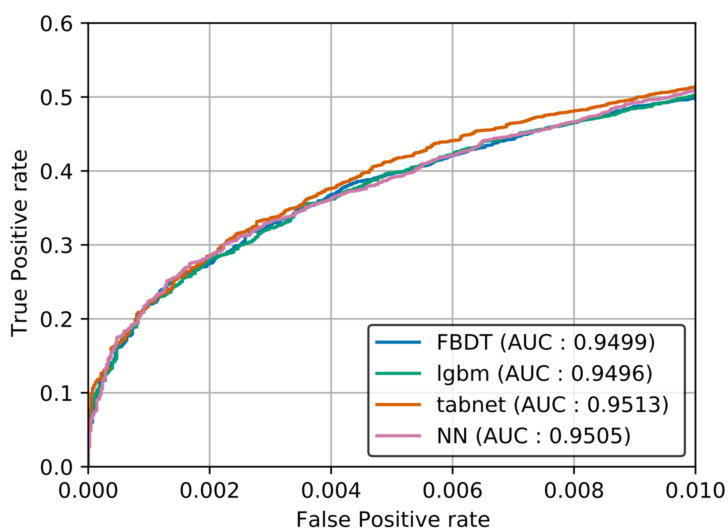


FIG. 121. Comparison of the performance of machine learning models

Figure 122 shows the output of the TabNet-based CSMVA model. Signals and backgrounds are classified well, and the performance of the training and validation dataset is consistent within the uncertainty, which means the model is not overfitted to the training samples.

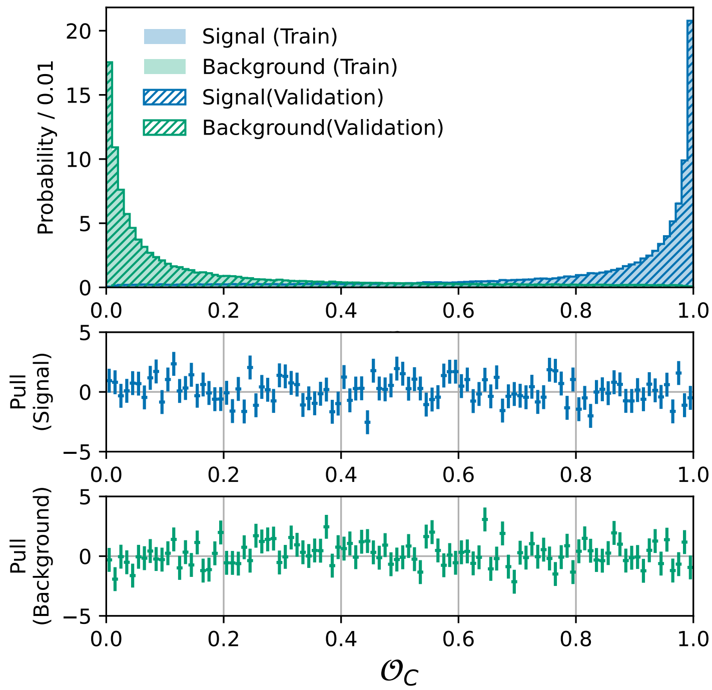


FIG. 122. Output of the TabNet-based CSMVA model

Appendix 2: Modelling

1 Signal model

Figure 123 and Figure 124 show the modelling for longitudinal signal and transverse signal distributions. The ΔE distributions are modelled using a double bifurcated Gaussian function. The $\mathcal{O}_c$, that is a continuum suppression output transformed by the integral probability of the transverse signal distribution, is modelled using a first-order Chebyshev polynomial function. The di-pion resonances are well described by a Breit-Wigner formula. In this analysis, the amplitude of the relativistic Breit-Wigner formula is

$$A(m) = \frac{p_{\rho^\pm\to\pi^\pm\pi^0}}{m^2 - m_0^2 + im_0\Gamma(m)} \frac{B(p_{B\to\rho^+\rho^-})}{B(p'_{B\to\rho^+\rho^-})} \frac{B(p_{\rho^\pm\to\pi^\pm\pi^0})}{B(p'_{\rho^\pm\to\pi^\pm\pi^0})}$$

where m_0 is the resonance mass, and $p^{(\prime)}$ is the momentum of the daughter particles of the B meson. Here, we consider the case where B decays into two dipion systems, $\pi^+\pi^0$ and $\pi^-\pi^0$. In the rest frame of the B meson, p represents the momentum of the system reconstructed as $\pi^\pm\pi^0$, and p' represents the momentum of those systems when they are truly $\rho^\pm$. $\Gamma(m)$ is

$$\Gamma(m) = \left(\frac{p}{p'}\right)^3 \left(\frac{m_0}{m}\right) \Gamma_0 \left[\frac{B(p)}{B(p')}\right]^2.$$

Γ_0 is the nominal width, and $B(p)$ is the Blatt-Weisskopf form factor:

$$B(p) = \frac{1}{\sqrt{1 + (Rp)^2}}.$$

R is $3\,\mathrm{GeV}^{-1}$. For higher energy γ, ECL can miss the deposited energy in clusters; a peak of the ΔE distribution negatively shifts. This can affect the distributions of ρ (di-pion) mass and $\cos\theta_{\rho^\pm}$. In particular, because of the stronger correlation between ΔE and $\cos\theta_{\rho^\pm}$, we slice the ΔE distributions into ten regions ([-0.15, -0.1, -0.07, -0.05, -0.04, -0.03, -0.02, -0.01, 0.01, 0.02, 0.15]GeV) and fit the $\cos\theta_{\rho^\pm}$ distributions for each region, as shown in Figure 125.

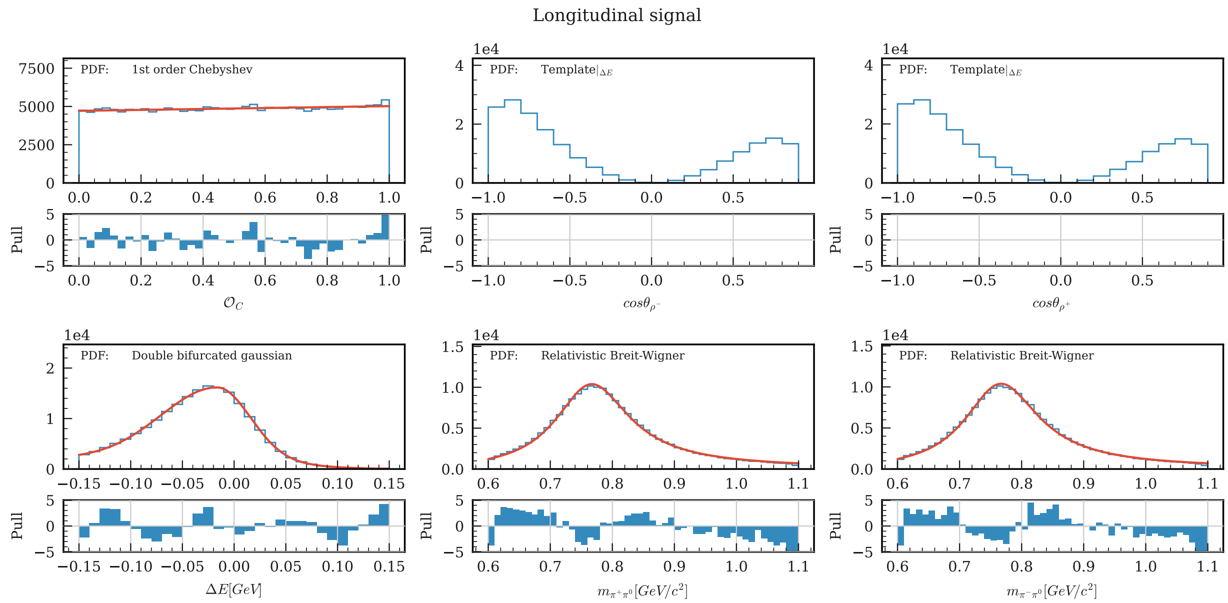


FIG. 123. Modeling fits for Longitudinal signal.

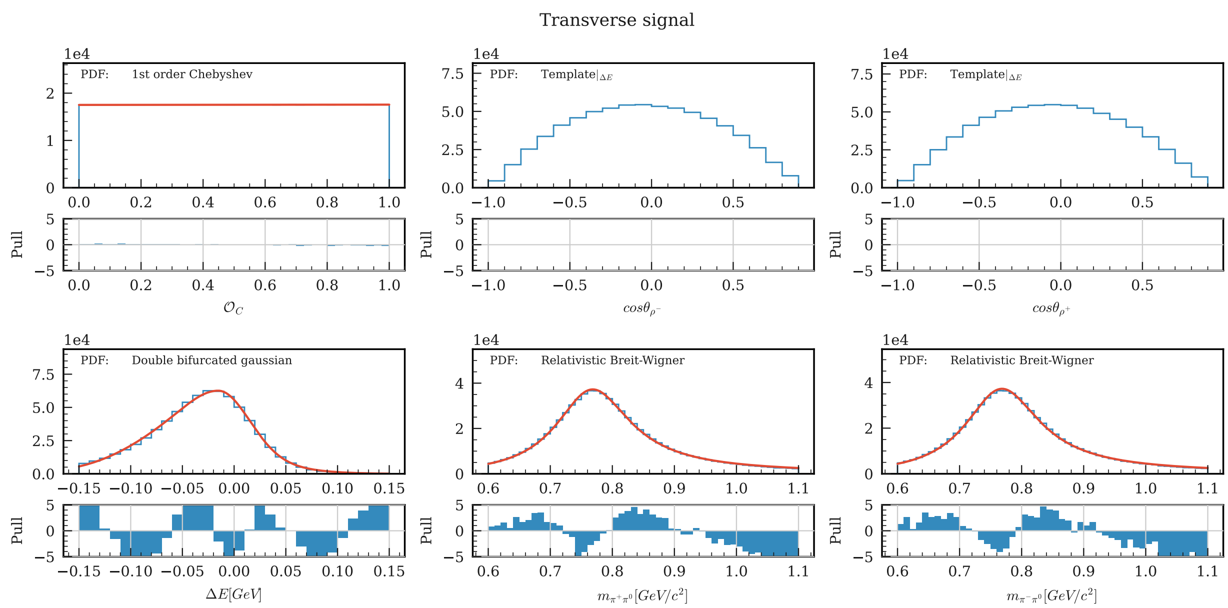


FIG. 124. Modeling fits for Transverse signal.

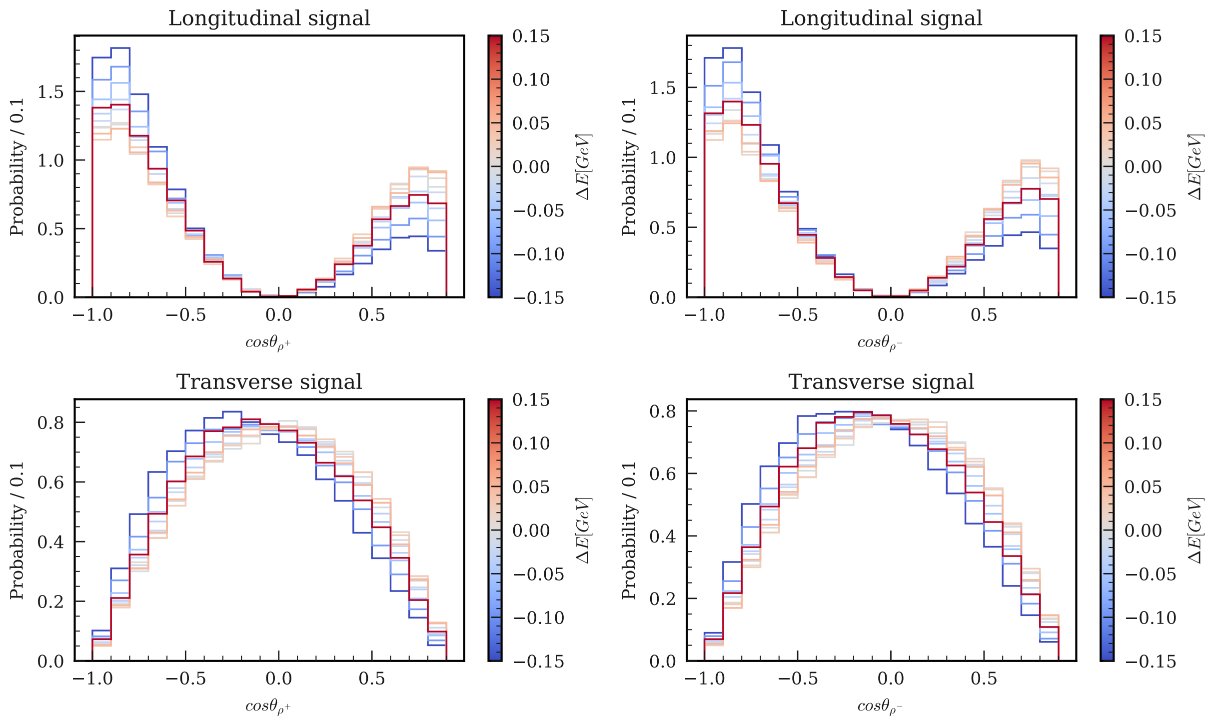


FIG. 125. Correlation between ΔE and $cos\theta_{\rho^\pm}$ for longitudinal signal transverse signals

2 Self-crossfeed signal

Some signal events are not reconstructed correctly and are called "self-crossfeed". In this analysis, they are categorized into three contributions: $2\pi^{\pm}$ of longitudinal, and $< 2\pi$ of longitudinal, and self-crossfeed of the transverse signals. $2\pi^{\pm}$ self-crossfeed is the events that have two correctly reconstructed $\pi^{\pm}$ from the longitudinal signals, but at least more than one π^0 is not reconstructed correctly due to the wrong combination of γ from $\rho^{\pm}$ or mis-reconstructed gamma by beam backgrounds or hadronic clusters. Its kinematic distribution differs from correctly reconstructed signals, but the vertex position is exactly the same as the signal events. Thus, its Δt distribution is the same as the signals, so it needs to be separated from the other backgrounds. $< 2\pi$ self-crossfeed has a contribution from correctly reconstructed $\pi^{\pm}$ or γ from the longitudinal signals, and it has information on the signals, but the distribution of the fitting variable is smeared, so it is modelled separately. The ratio of the Self-crossfeed events with no correctly reconstructed $\pi^{\pm}$ is quite small. It is less than 10% of 1π self-crossfeed, so 1π self-crossfeed and 0π self-crossfeed are summed up and treated as the $< 2\pi$ self-crossfeed. The transverse signals don't have $\mathcal{CP}$ violating effects, and the expected number of signals is very small. So, the self-crossfeed events of the transverse signals don't need to be categorized into several categories. In this analysis, mis-reconstructed events of the transverse signals are summed up and treated as the self-crossfeed of the transverse signals.

Figure 126, Figure 127, and Figure 128 shows the modelling for $2\pi^{\pm}$ and $< 2\pi\pm$, and transverse self-crossfeed events. Modelling for the ΔE and $\mathcal{O}_C$ distributions are the same function as the signal components. For $\mathrm{m}_{\pi^{\pm}\pi^0}$ distributions, the sum of a first-order Chebyshev polynomial and Breit-Wigner function is used. In self-crossfeed events, $cos\theta_{\rho^{\pm}}$ correlates to each other. The correlations are shown in Figure 129 and Figure 130. In self-crossfeed events, if one $\rho^{\pm}$ is reconstructed correctly, the other is mis-reconstructed. Then, if one shows signal-like $cos\theta_{\rho^{\pm}}$, others show more background-like distributions. That is the cause of this correlation. To consider this, a 2D histogram is used for $cos\theta_{\rho^{\pm}}$ PDF of self-crossfeed events.

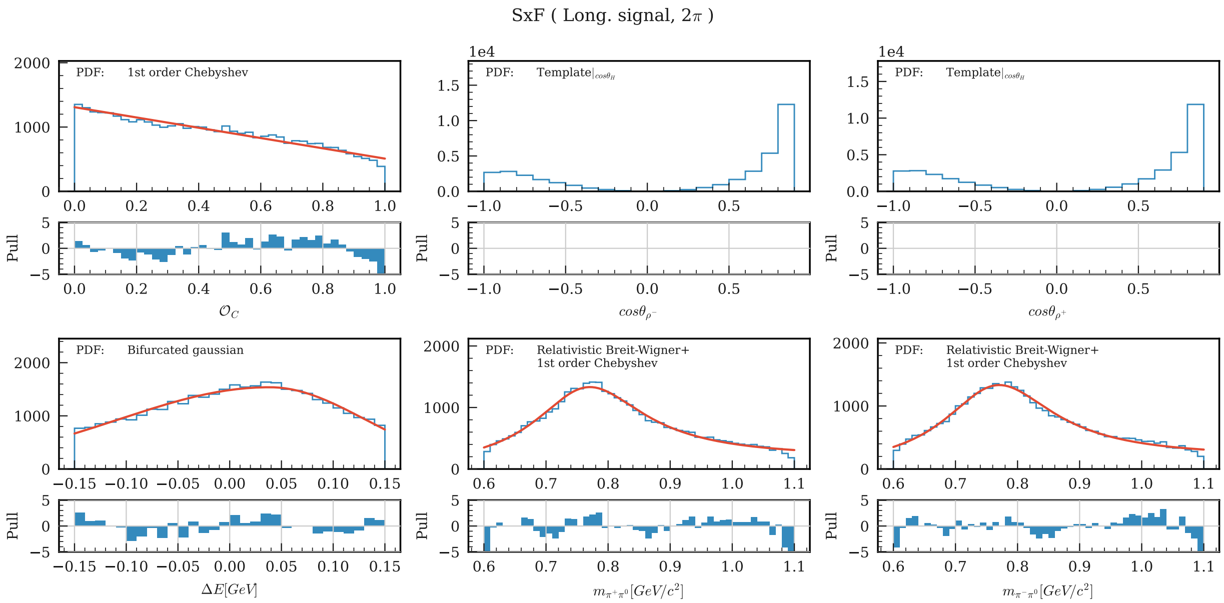


FIG. 126. Modeling fits for SxF (Long. signal, 2π).

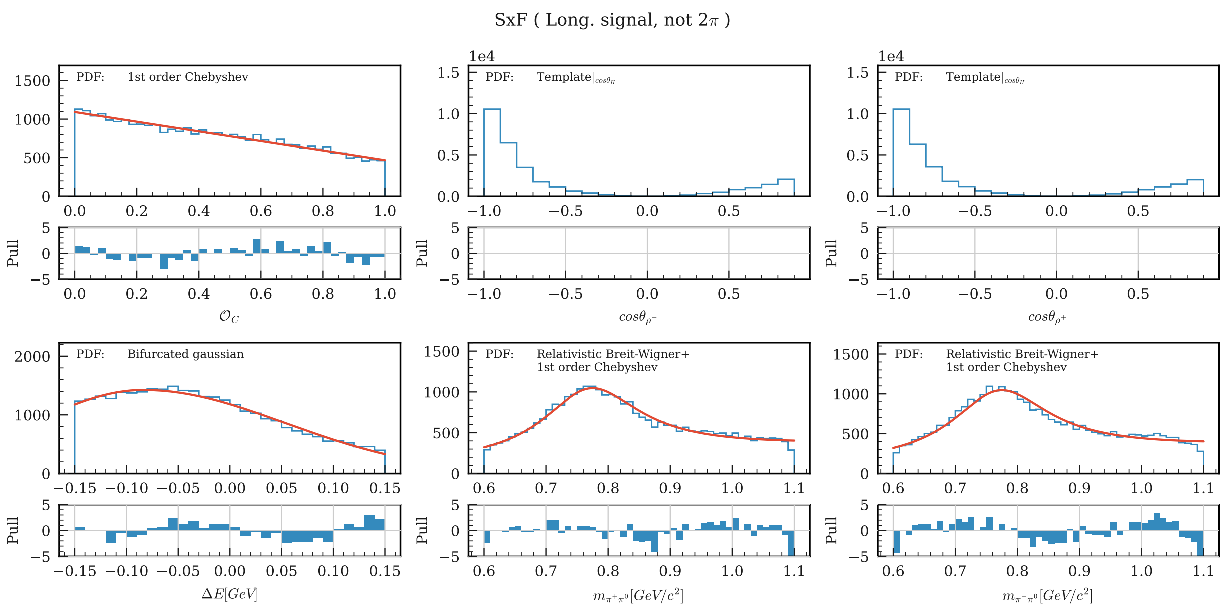


FIG. 127. Modeling fits for SxF (Long. signal, not 2π).

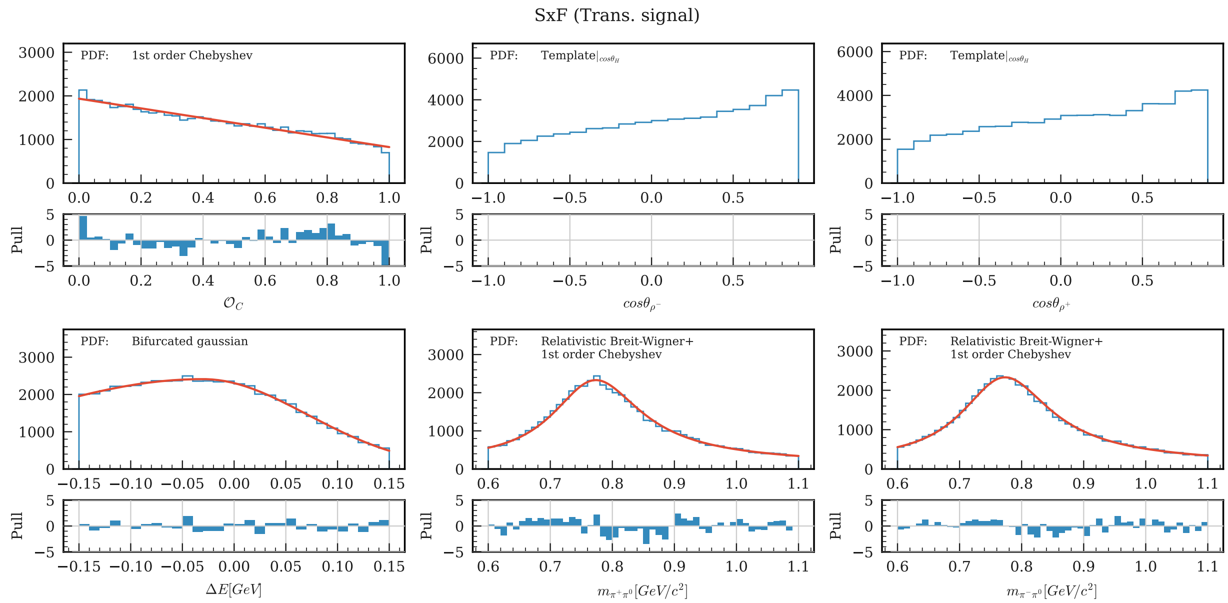


FIG. 128. Modeling fits for SxF (Trans. signal).

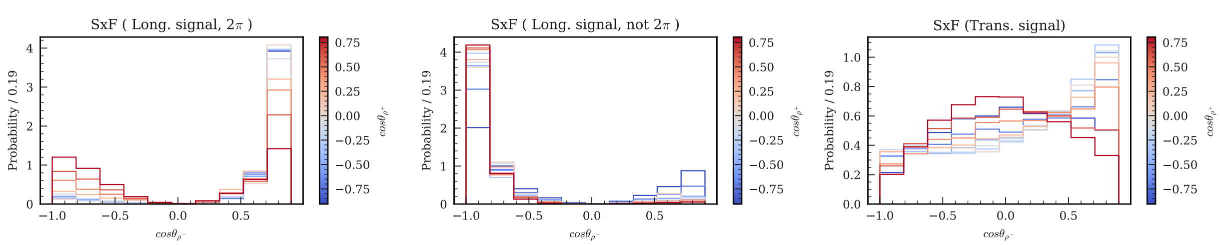


FIG. 129. Correlation between $cos\theta_{\rho^+}$ and $cos\theta_{\rho^-}$ in the self-crossfeed events

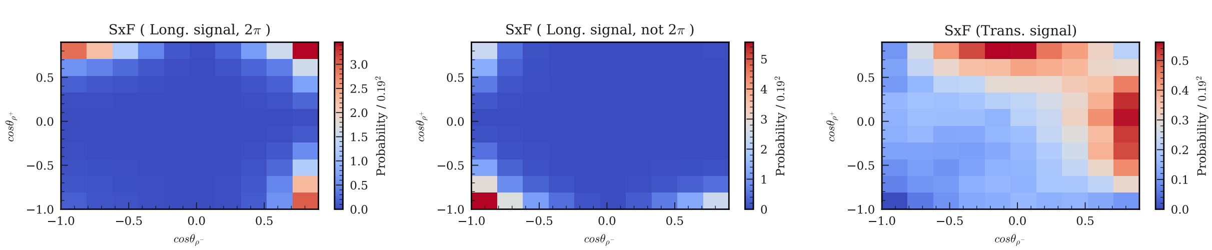


FIG. 130. PDF for $cos\theta_{\rho^+}$ and $cos\theta_{\rho^-}$ in the self-crossfeed events

3 $q\bar{q}$ and $B\bar{B}$

Figure 131 shows the modelling for the $q\bar{q}$ backgrounds. $\mathcal{O}_C$ and $cos\theta_H$ are modelled by template histograms with correlations and ΔE and $m_{\pi^{\pm}\pi^0}$ are modelled using a second order Chybyshev polynomial and the sum of the Relativistic Breit-Wigner and 1st order Chebyshev polynomial function.

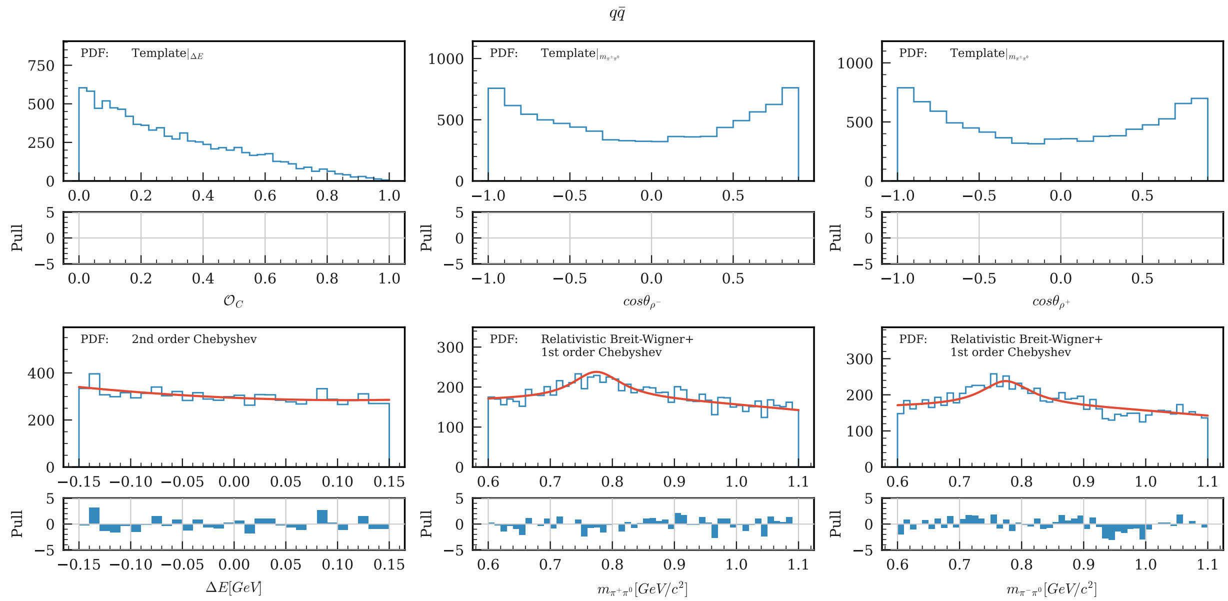


FIG. 131. Modeling fits for $q\bar{q}$.

Figure 132 shows the modelling for the $B\bar{B}$ backgrounds. $B\bar{B}$ has several decays that make a peak in the variables for the fitting, and they are modelled separately as peaking backgrounds, which is described in 5. In this analysis, $B\bar{B}$ backgrounds have only non-peaking contributions. $\mathcal{O}_C$ and $cos\theta_H$ are modelled by template histograms with correlations, and ΔE and $m_{\pi^{\pm}\pi^0}$ are modelled using a second order Chybyshev polynomial function. Several correlations is considered in the modelling for $B\bar{B}$ and $q\bar{q}$ backgrounds. The first one it the correlation between $cos\theta_{\rho^{\pm}}$ and $m_{\pi^{\pm}\pi^0}$. To consider this, we sliced the data with $m_{\pi^{\pm}\pi^0}$ distributions and then made histograms for each $m_{\pi^{\pm}\pi^0}$ slice. The same correlation was observed between $\mathcal{O}_C$ and ΔE, and the correlation was considered in the same ways.

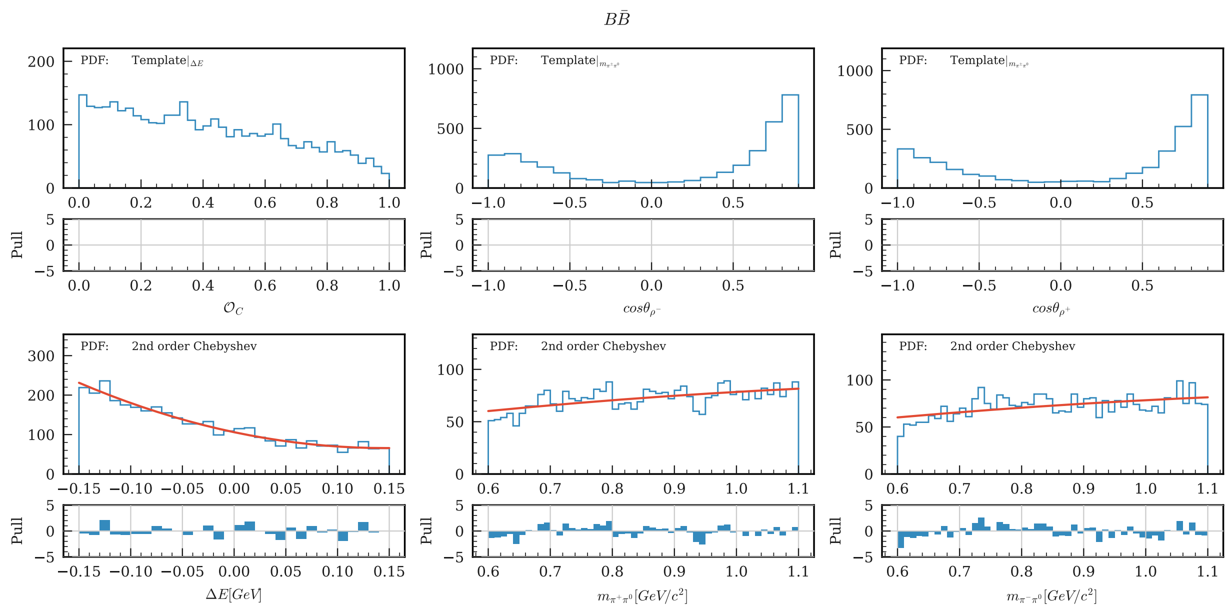


FIG. 132. Modeling fits for $B\bar{B}$.

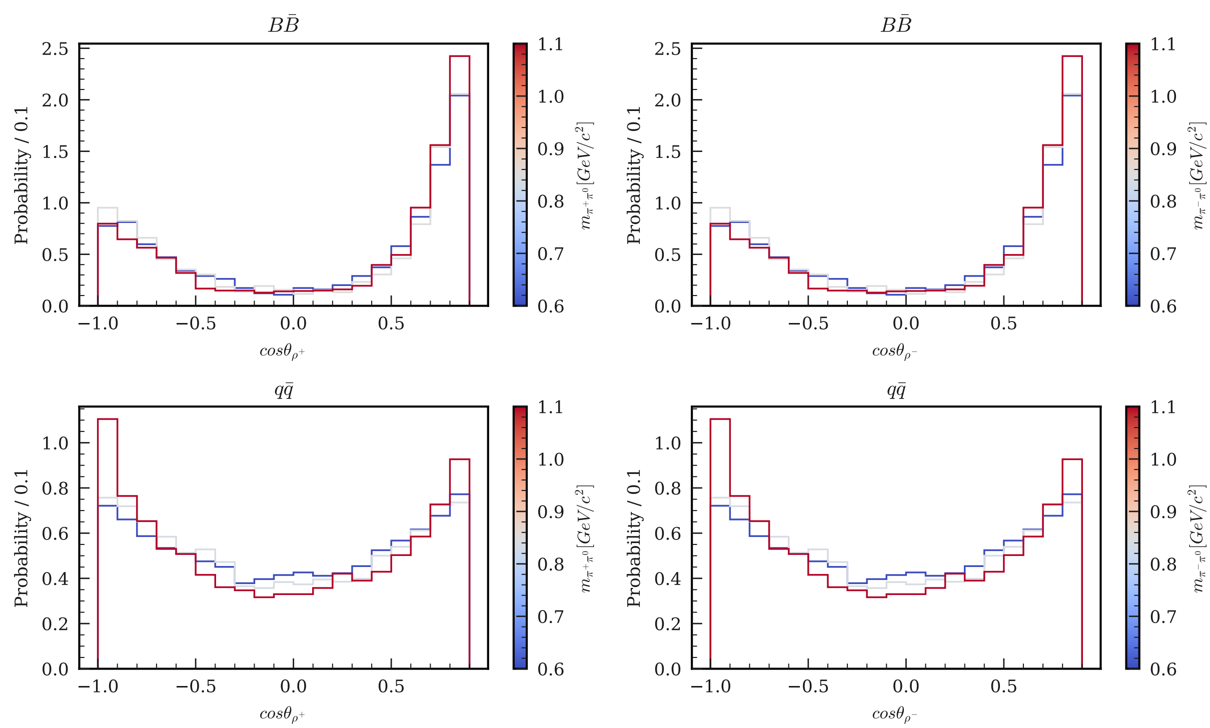


FIG. 133. $cos\theta_{\rho^\pm}$ modelling for $B\bar{B}$ and $q\bar{q}$ backgrounds.

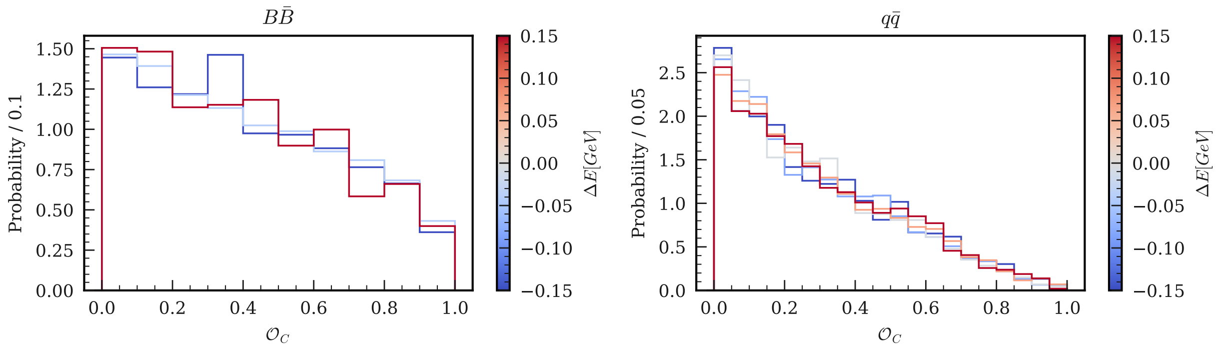


FIG. 134. $\mathcal{O}_C$ modelling for $B\bar{B}$ and $q\bar{q}$ backgrounds.

4 $\tau\tau$

Figure 135 shows the PDF modeling for $\tau\tau$ backgrounds. The CSMVA outputs are modeled by an exponential function, ΔE is modeled by a Chebyshev polynomial function (2nd order), $m_{\pi\pi^0}$ is modeled by a sum of a Chebyshev polynomial (1st order) and relativistic Breit-Wigner function, and $cos\theta_\rho$ is modeled by a template histogram.

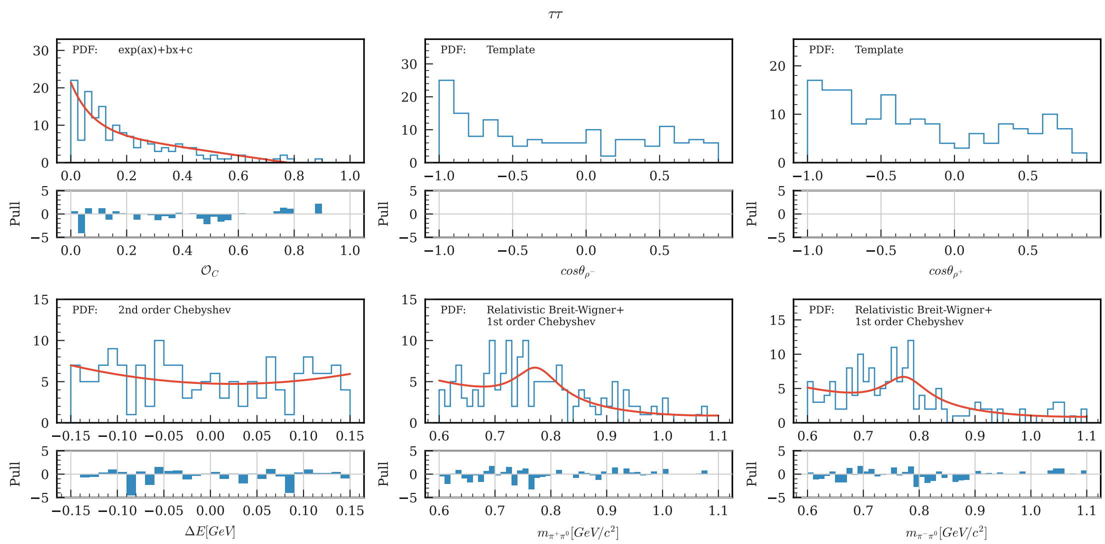


FIG. 135. Modeling fits for $\tau\tau$ using MC15rd samples

5 Peaking backgrounds

Figure 136 to Figure 152 shows the modelling for peaking backgrounds. In general, ΔE is modelled using double bifurcated Gaussian, $\mathcal{O}_C$ is modelled Chyebyshev polynomial, $m_{\pi^\pm\pi^0}$ is modelled linear function, and $\cos\theta_{\rho^\pm}$ is modelled using a template histogram, and in case the degree of freedom of the model is too much, it is reduced a little. For example, double bifurcated Gaussian to single bifurcated Gaussian.

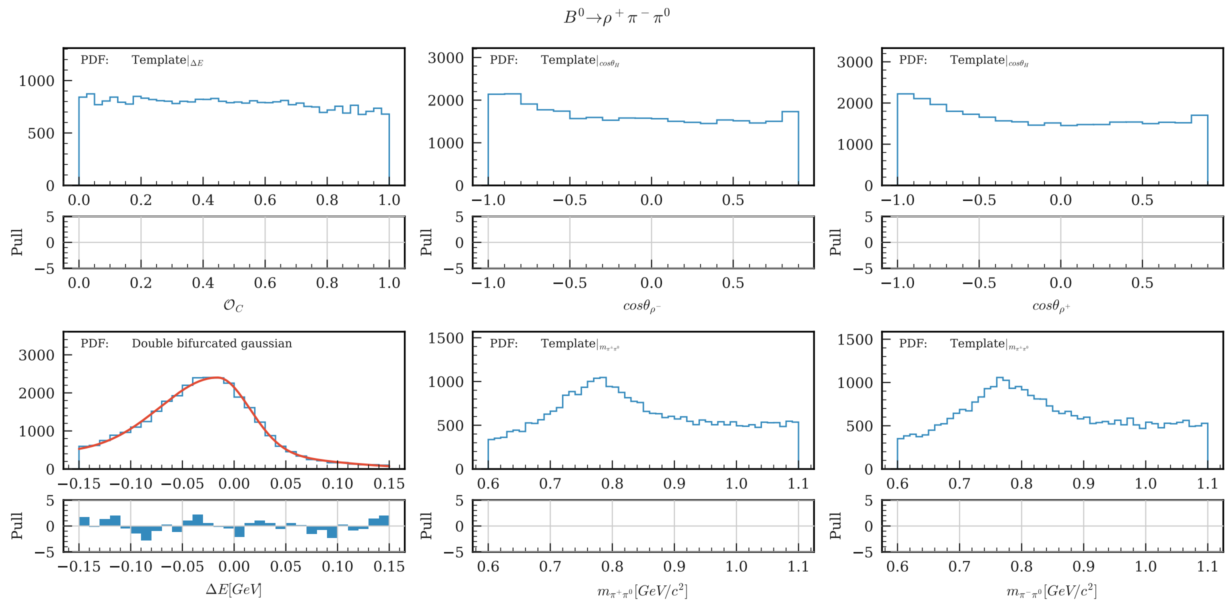


FIG. 136. Modeling fits for $B^0 \to \rho^+\pi^-\pi^0$.

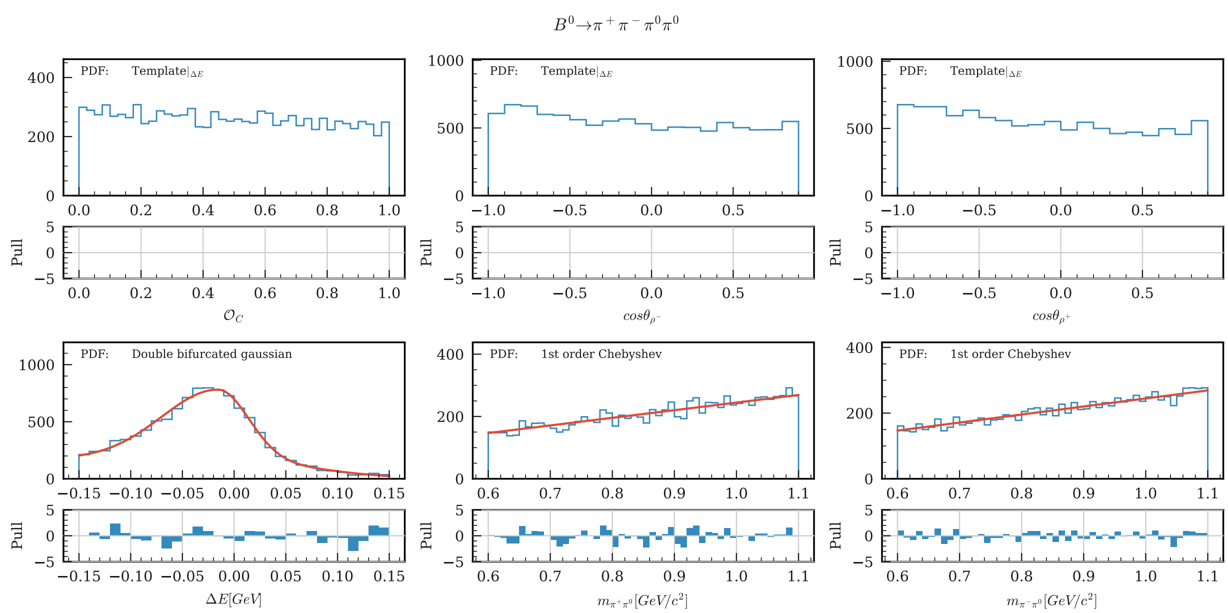


FIG. 137. Modeling fits for $B^0 \to \pi^+\pi^-\pi^0\pi^0$.

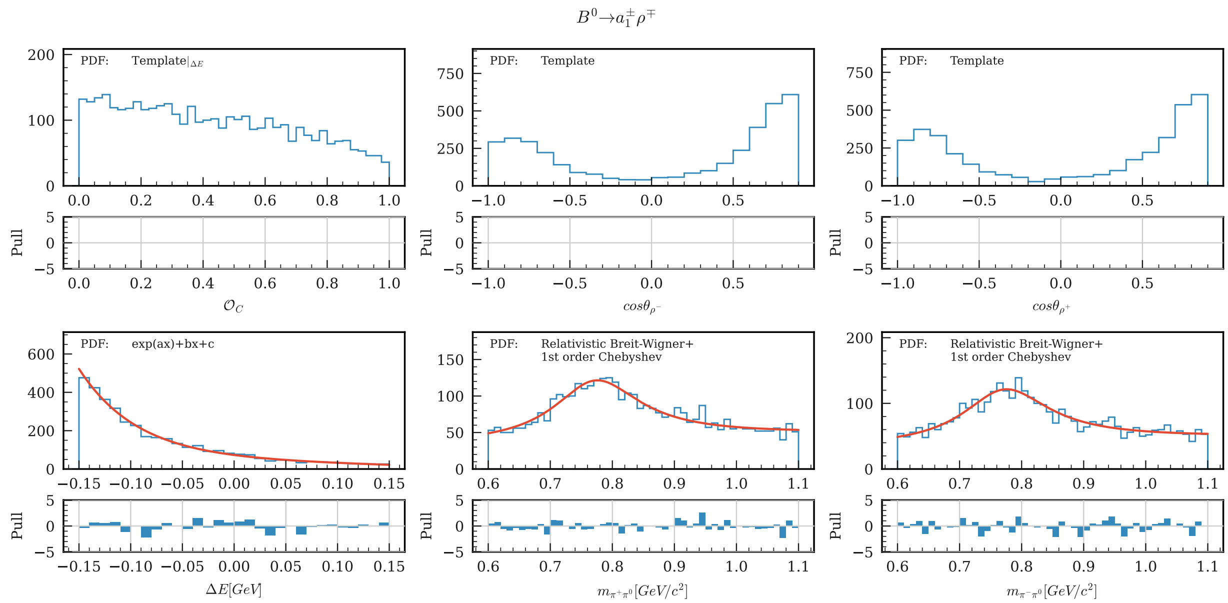


FIG. 138. Modeling fits for $B^0 \to a_1^\pm \rho^\mp$.

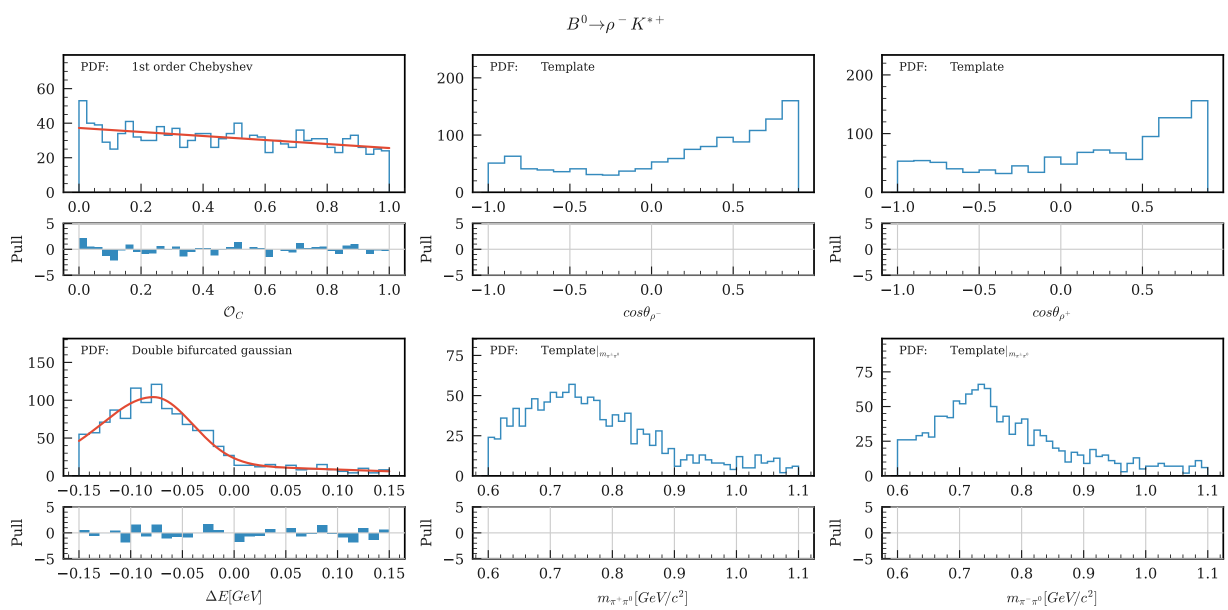


FIG. 139. Modeling fits for $B^0 \to \rho^- K^{*+}$.

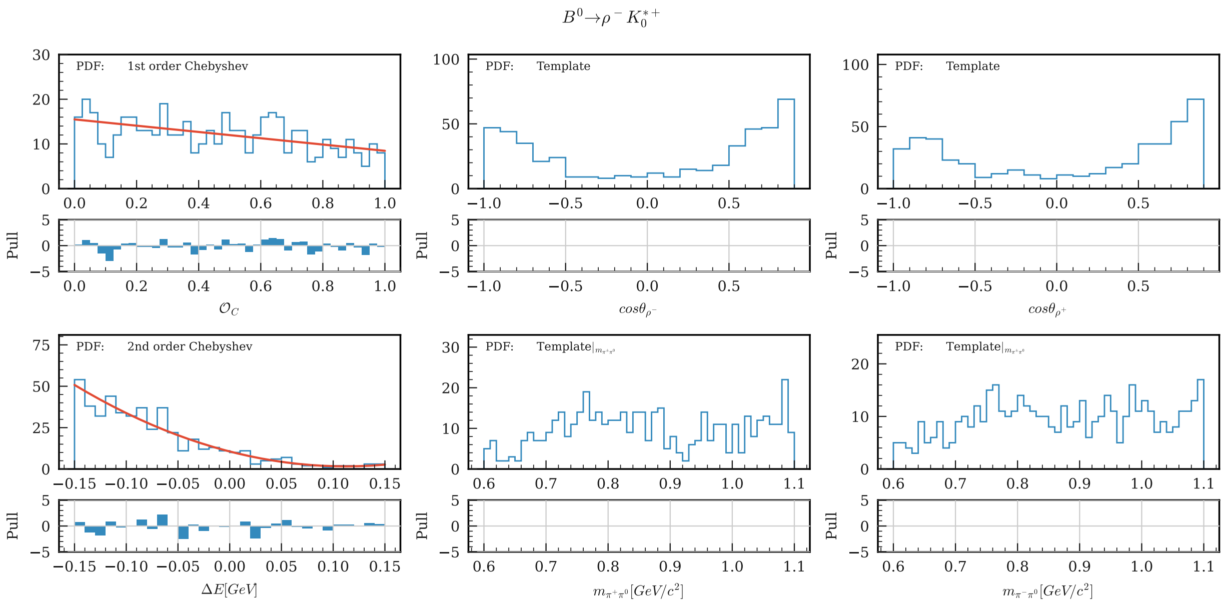


FIG. 140. Modeling fits for $B^0 \to \rho^- K_0^{*+}$.

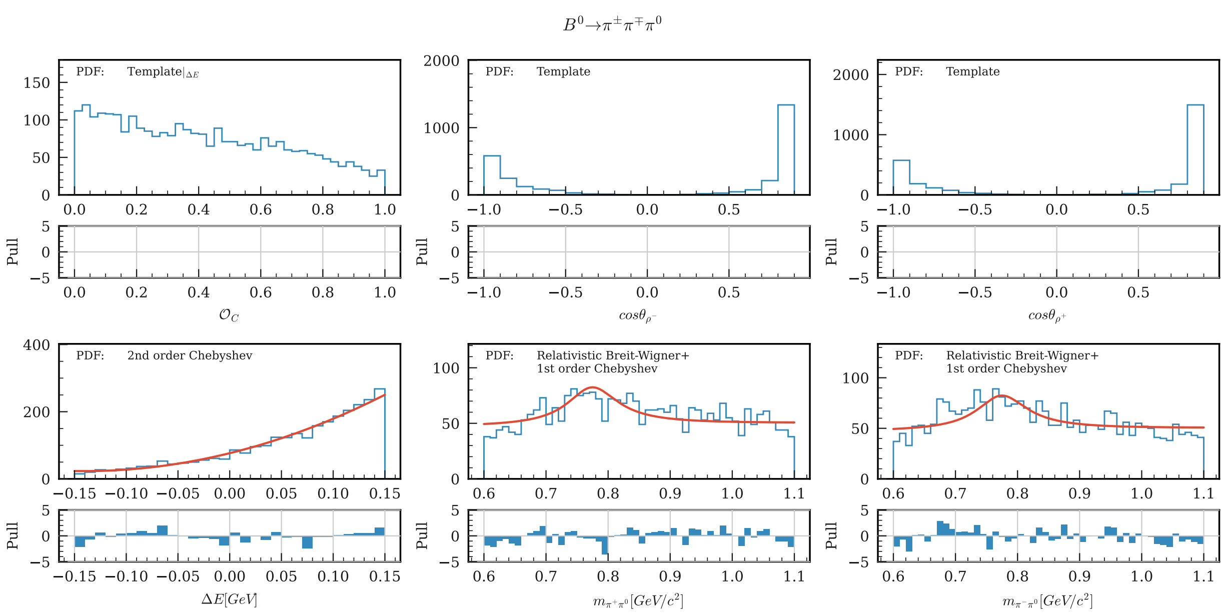


FIG. 141. Modeling fits for $B^0 \to \pi^\pm \pi^\mp \pi^0$.

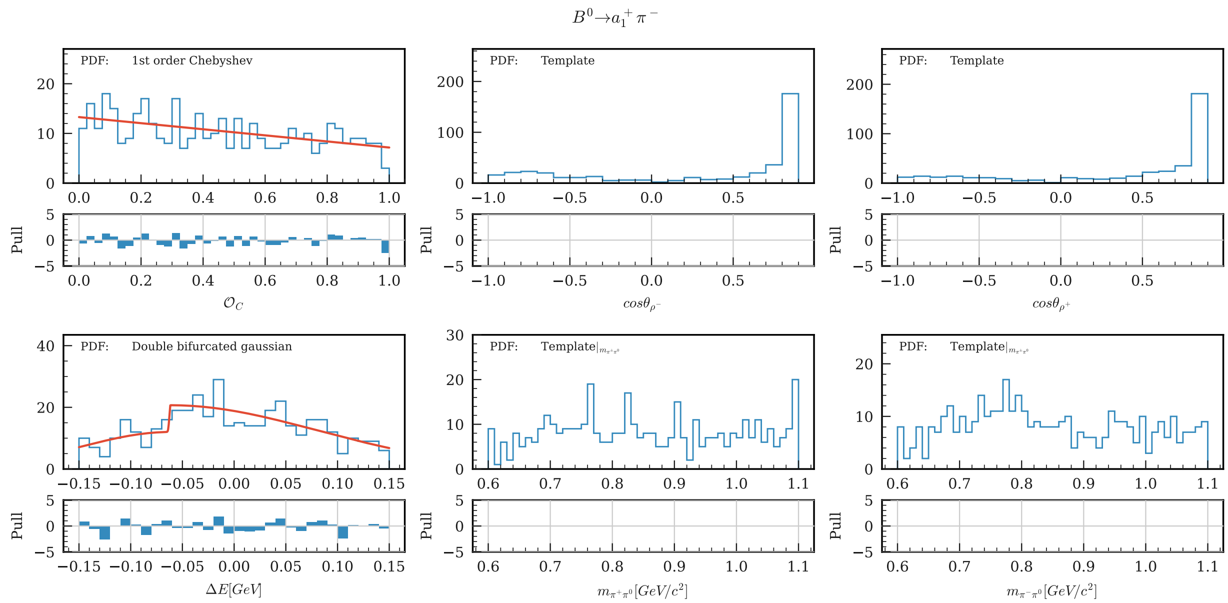


FIG. 142. Modeling fits for $B^0 \to a_1^+ \pi^-$.

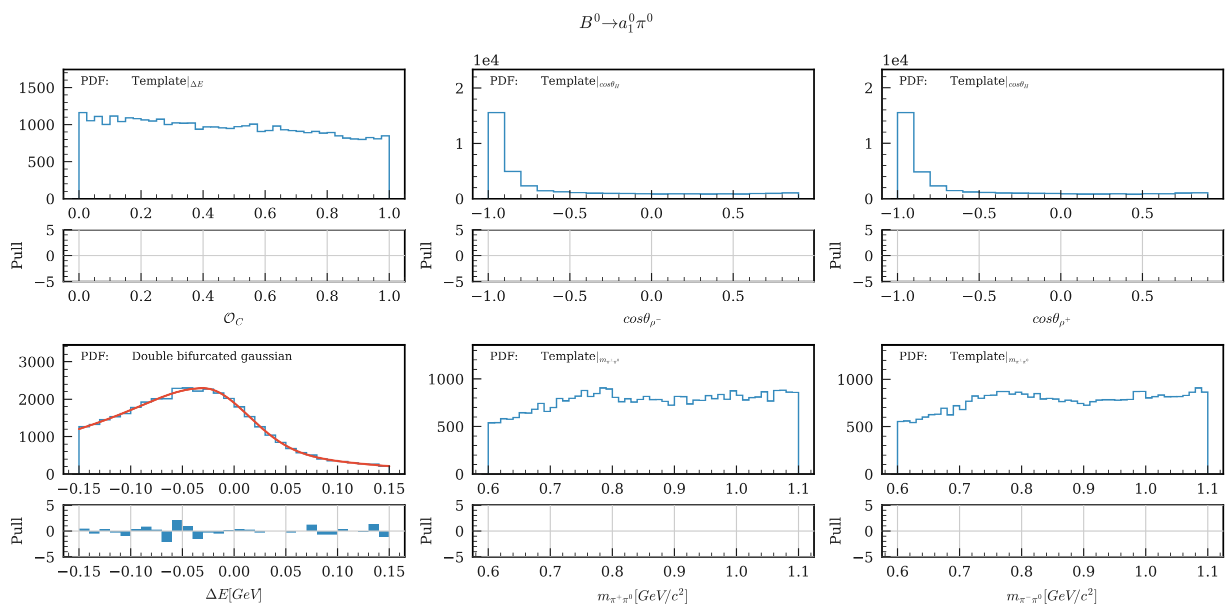


FIG. 143. Modeling fits for $B^0 \to a_1^0 \pi^0$.

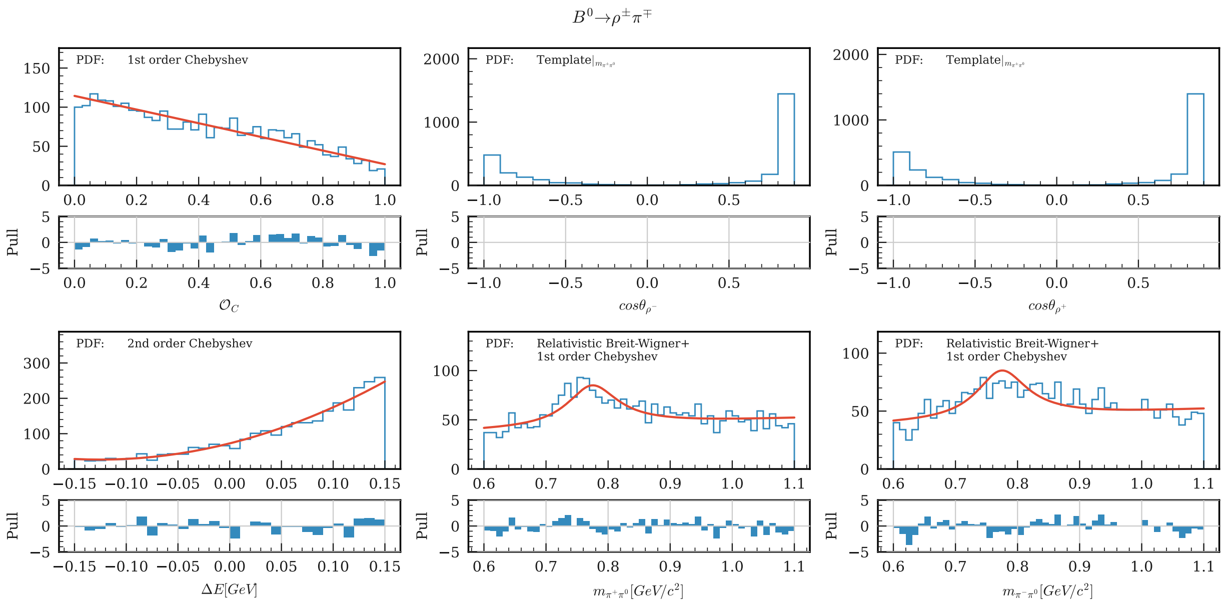


FIG. 144. Modeling fits for $B^0 \to \rho^\pm \pi^\mp$.

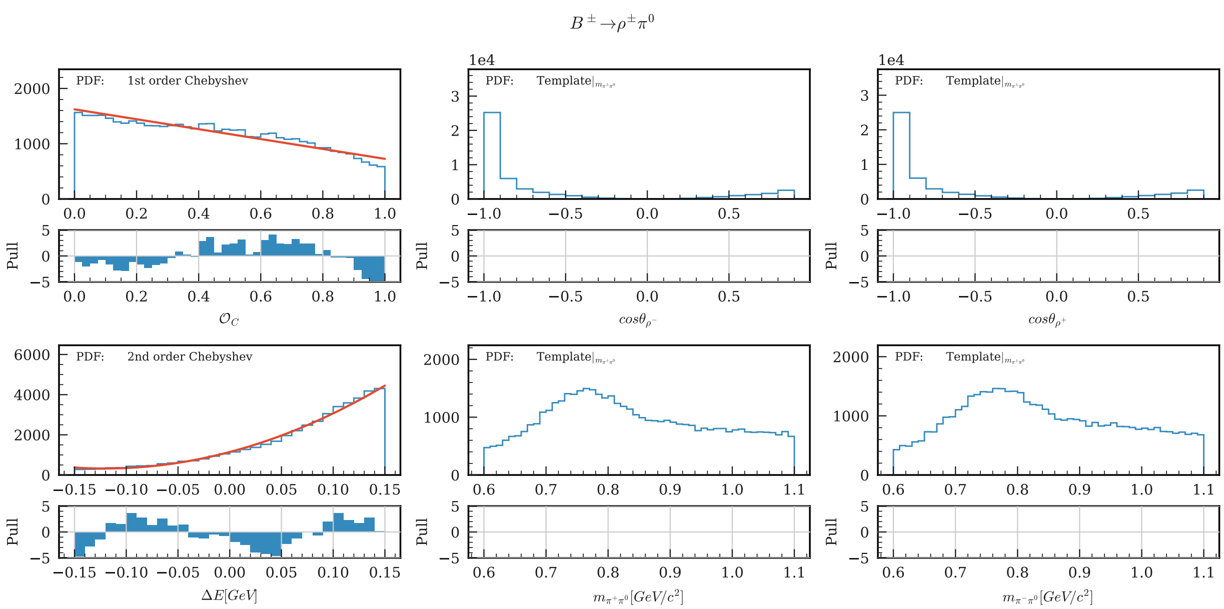


FIG. 145. Modeling fits for $B^\pm \to \rho^\pm \pi^0$.

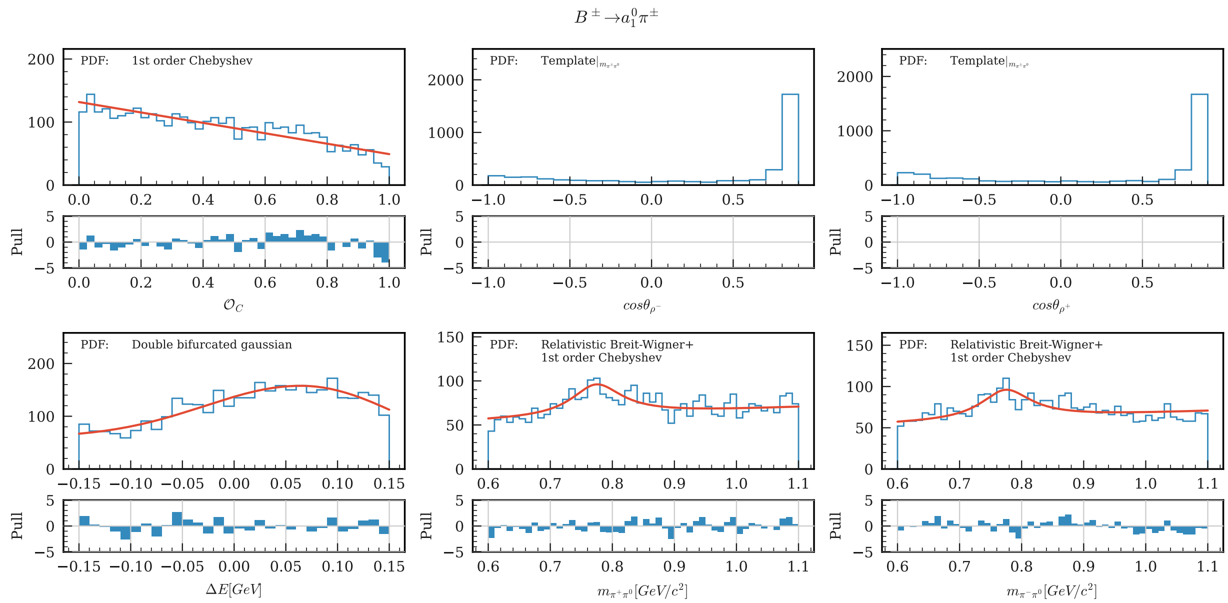


FIG. 146. Modeling fits for $B^{\pm} \to a_1^0 \pi^{\pm}$.

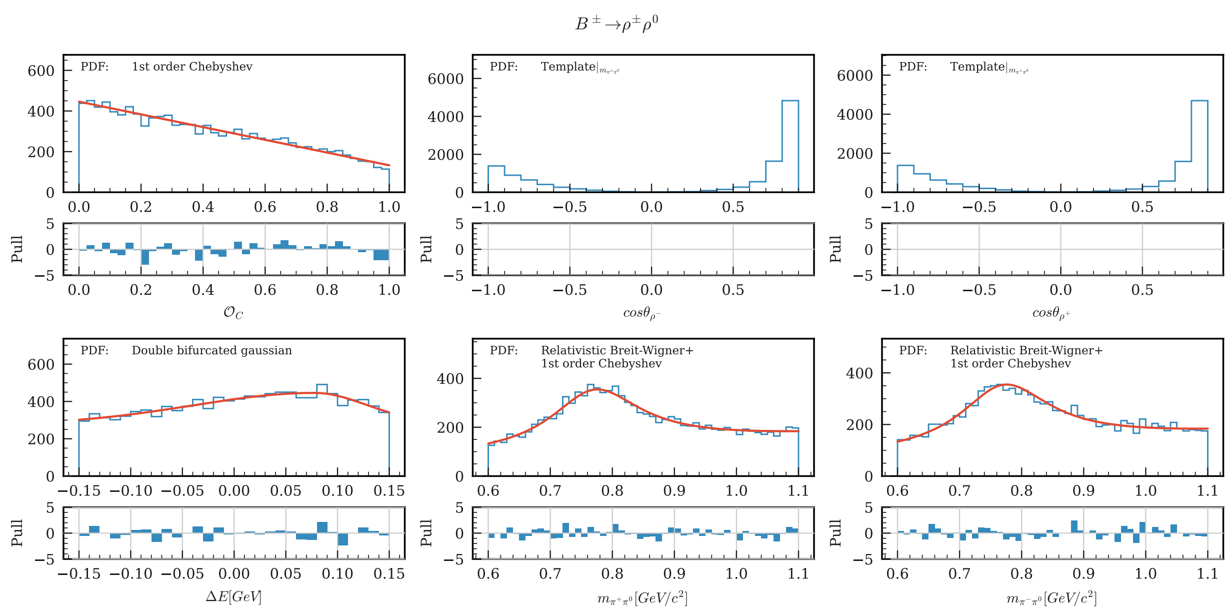


FIG. 147. Modeling fits for $B^{\pm} \to \rho^{\pm} \rho^0$.

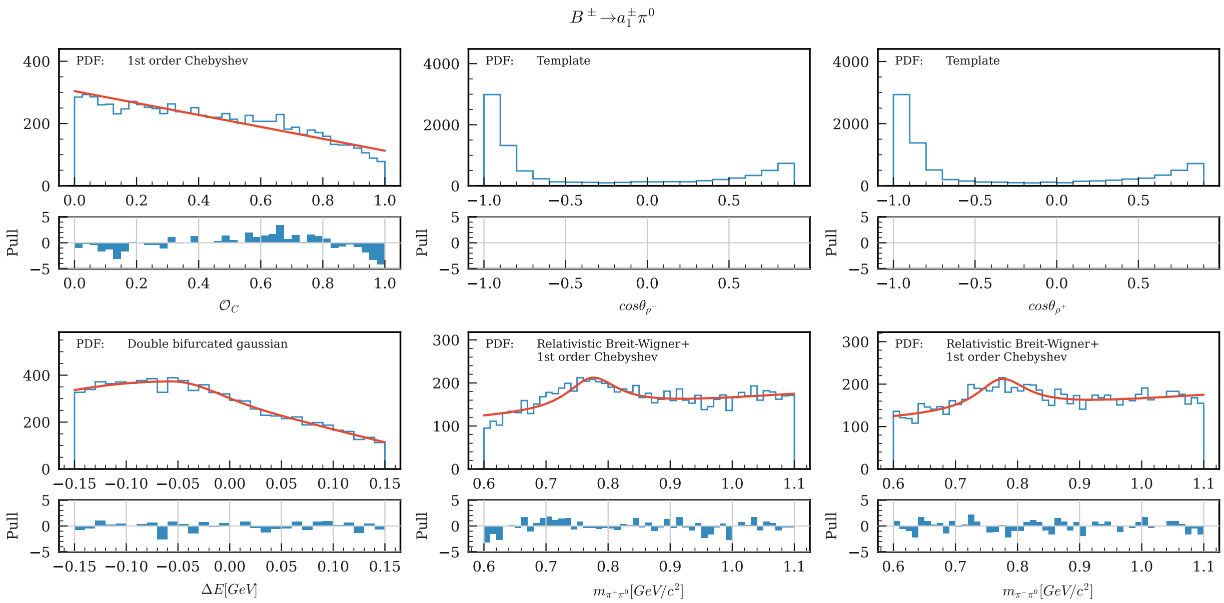


FIG. 148. Modeling fits for $B^{\pm} \to a_1^{\pm}\pi^0$.

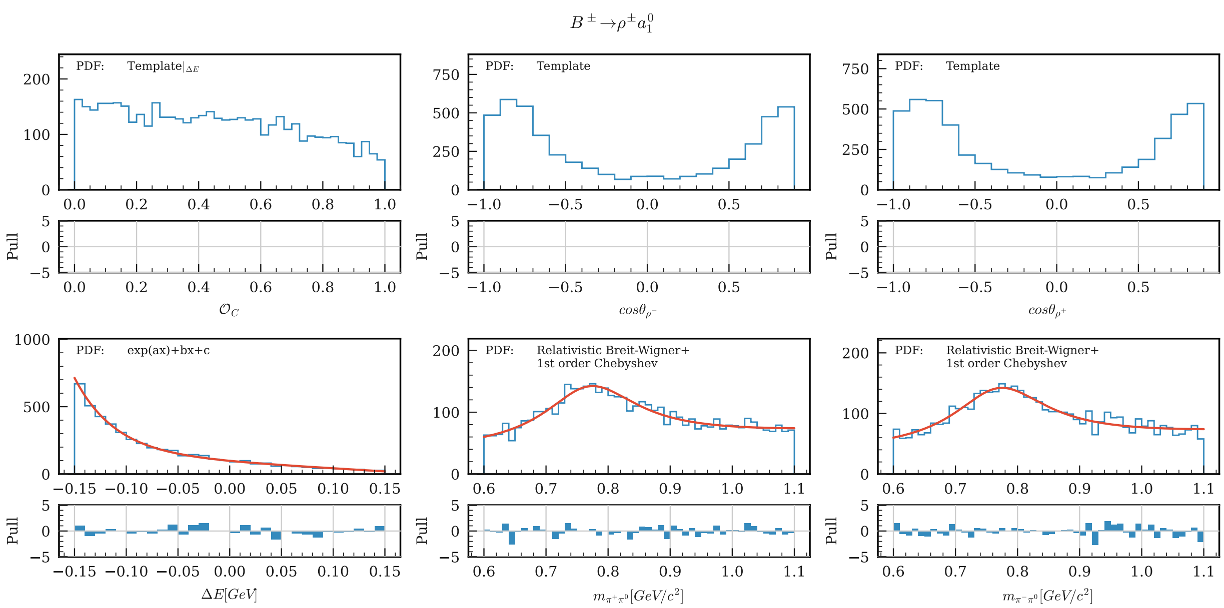


FIG. 149. Modeling fits for $B^{\pm} \to \rho^{\pm}a_1^0$.

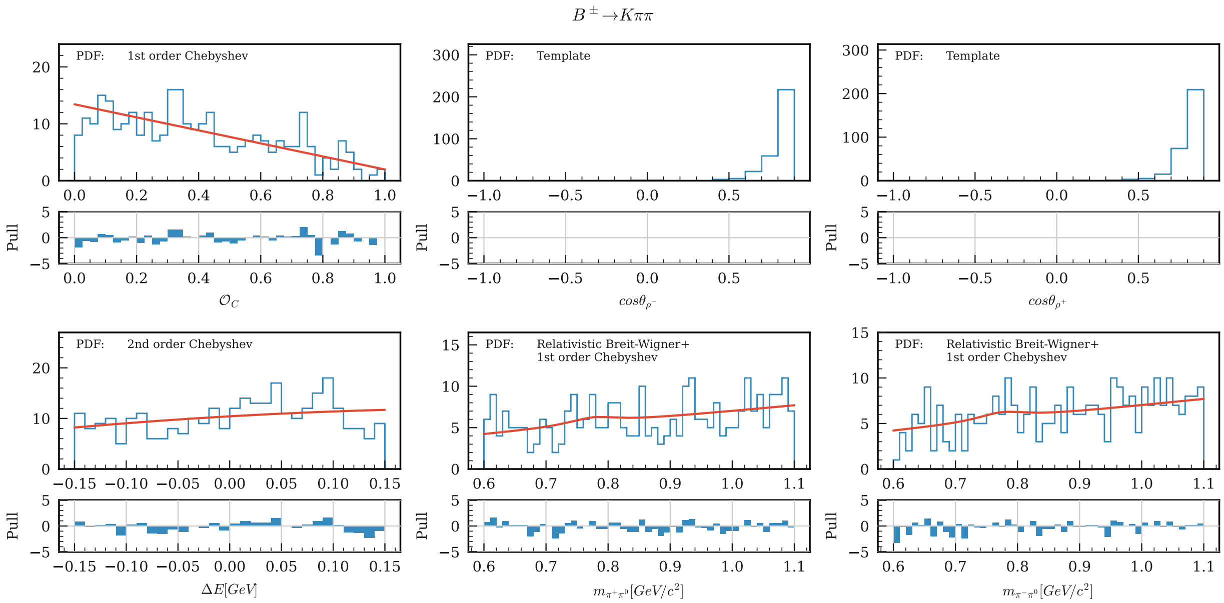


FIG. 150. Modeling fits for $B^{\pm} \to K\pi\pi$.

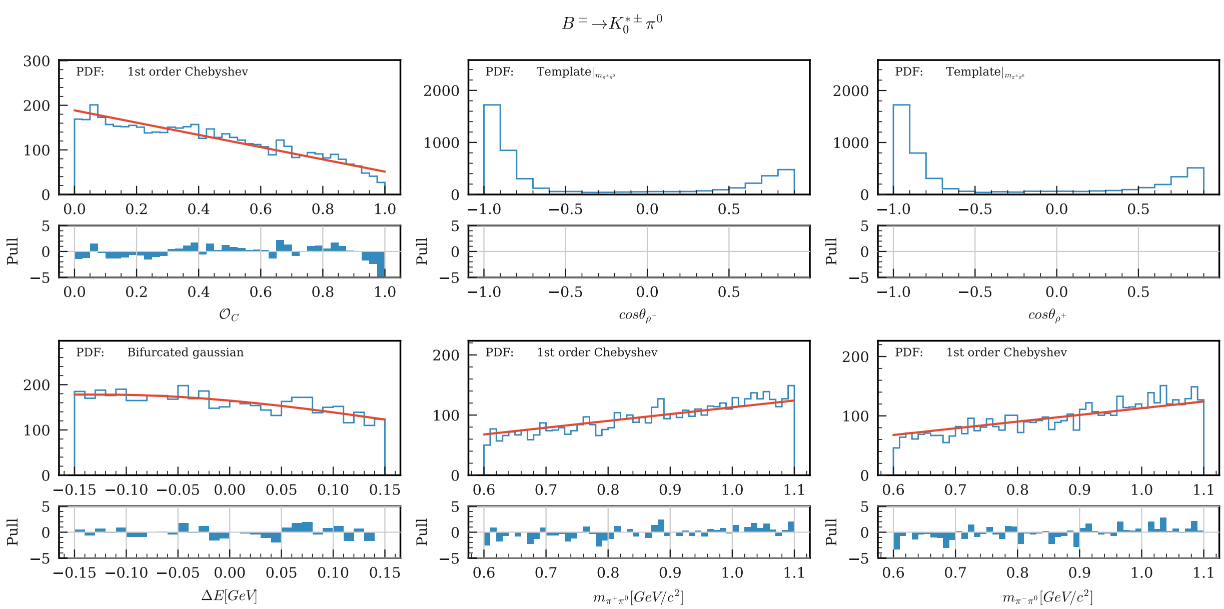


FIG. 151. Modeling fits for $B^{\pm} \to K_0^{*\pm}\pi^0$.

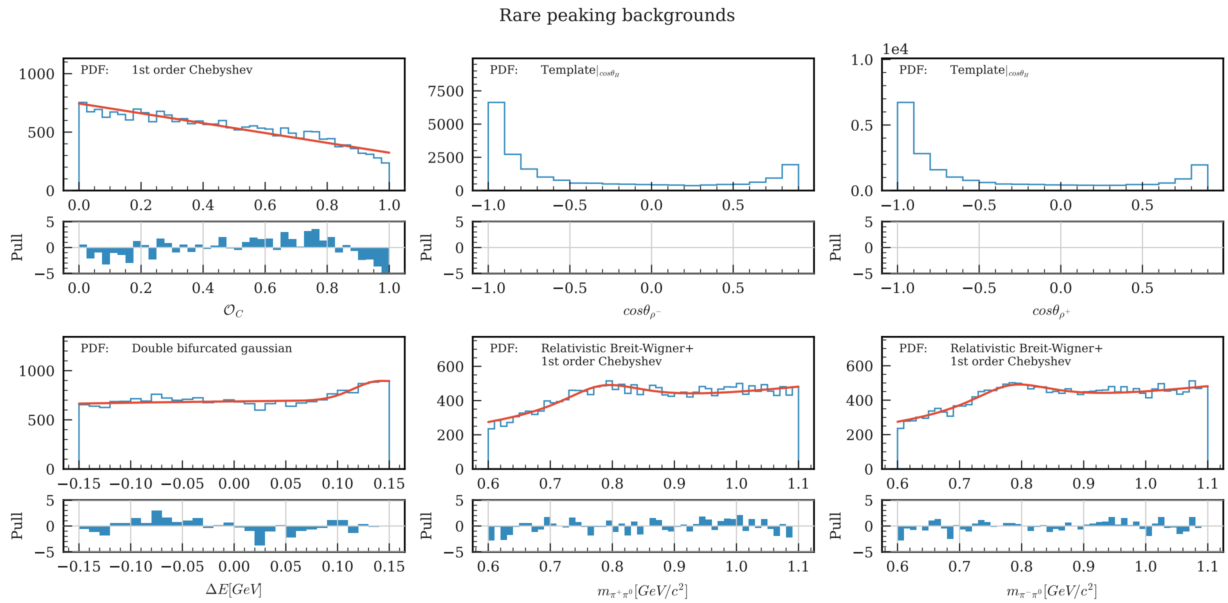


FIG. 152. Modeling fits for Rare peaking backgrounds.

In addition, several modes show correlations between the variables in the fitting, and this is taken into account for each mode. Figure 153 to Figure 157 shows the correlation of the fit variables of peaking backgrounds. The correlation between $cos\theta_{\rho^+}$ and $cos\theta_{\rho^-}$, $cos\theta_{\rho^\pm}$ and ΔE, $cos\theta_{\rho^\pm}$ and $m_{\pi^\pm\pi^0}$, $m_{\pi^\pm\pi^0}$ and $m_{\pi^\mp\pi^0}$, and $\mathcal{O}_C$ and ΔE is considered in the modelling. The correlation is considered by slicing the data and making a template 1D histogram for each component, and no 3D correlation is included.

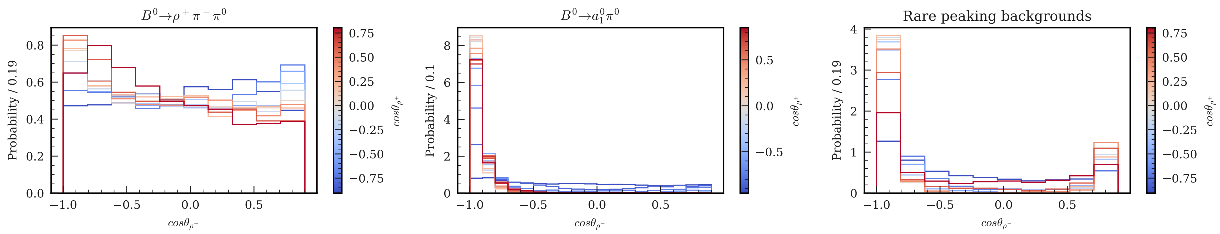


FIG. 153. Correlation between $cos\theta_{\rho^+}$ and $cos\theta_{\rho^-}$ of peaking backgrounds

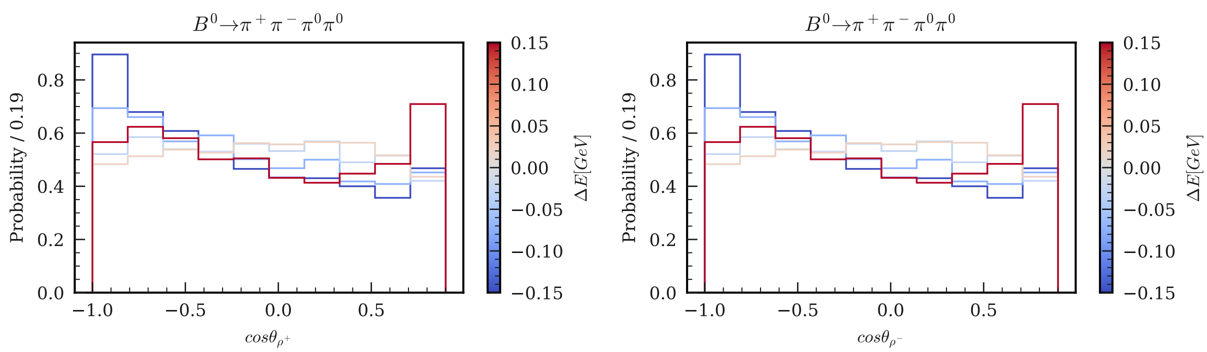


FIG. 154. Correlation between $cos\theta_{\rho^\pm}$ and ΔE of peaking backgrounds

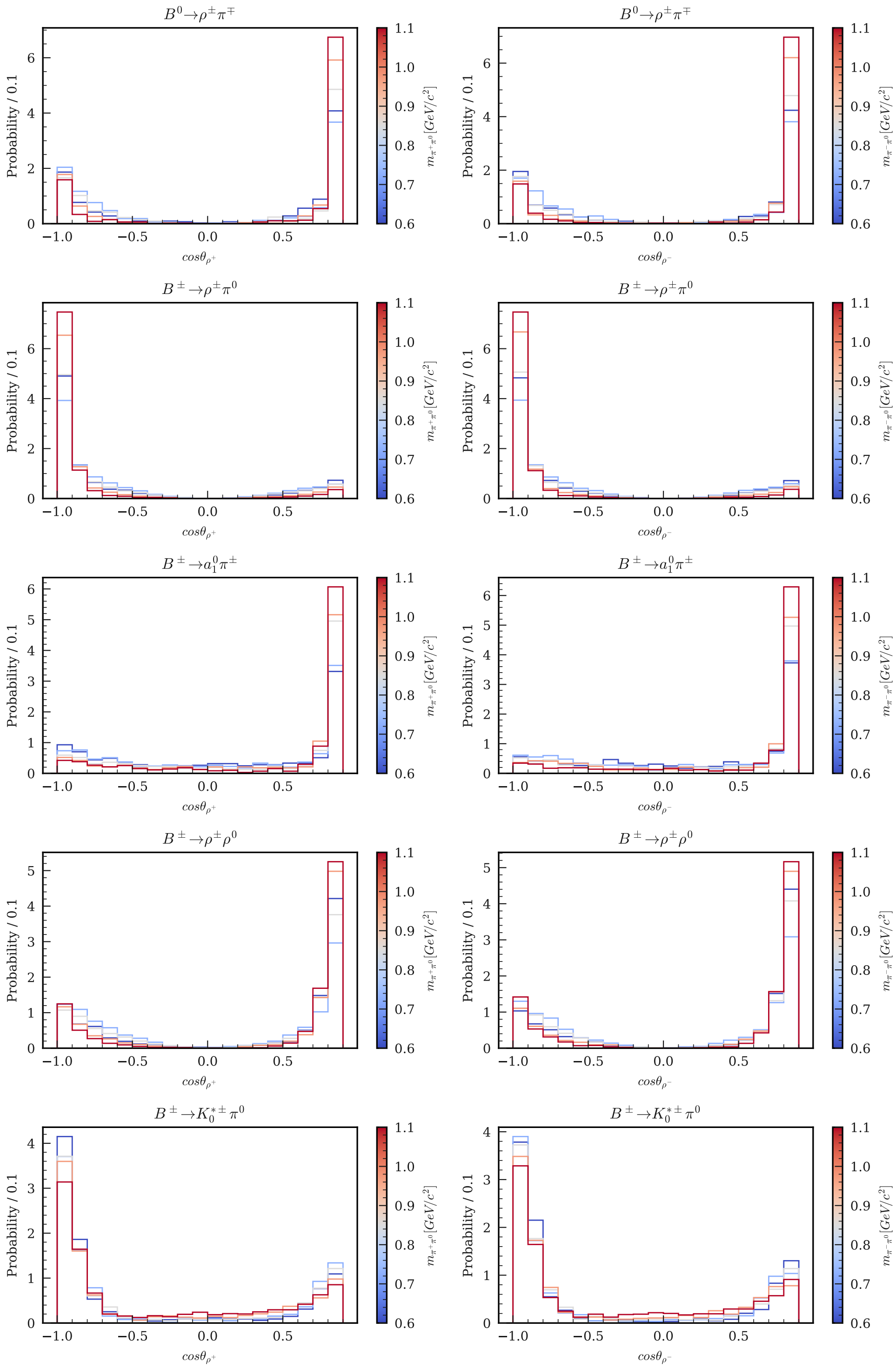


FIG. 155. Correlation between $cos\theta_{\rho^\pm}$ and $m_{\pi^\pm\pi^0}$ of peaking backgrounds

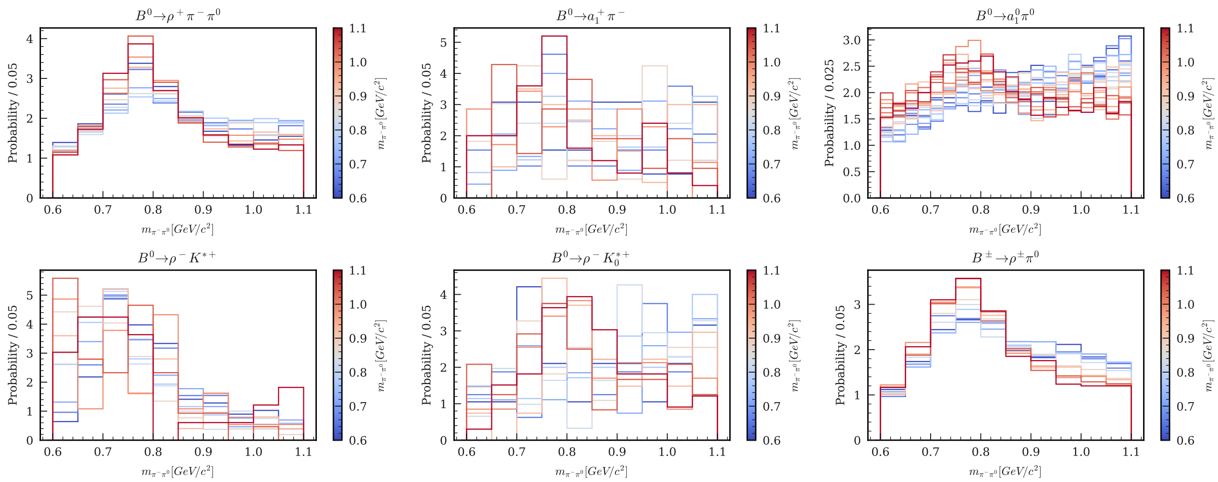


FIG. 156. Correlation between $m_{\pi^\pm\pi^0}$ and $m_{\pi^\mp\pi^0}$ of peaking backgrounds

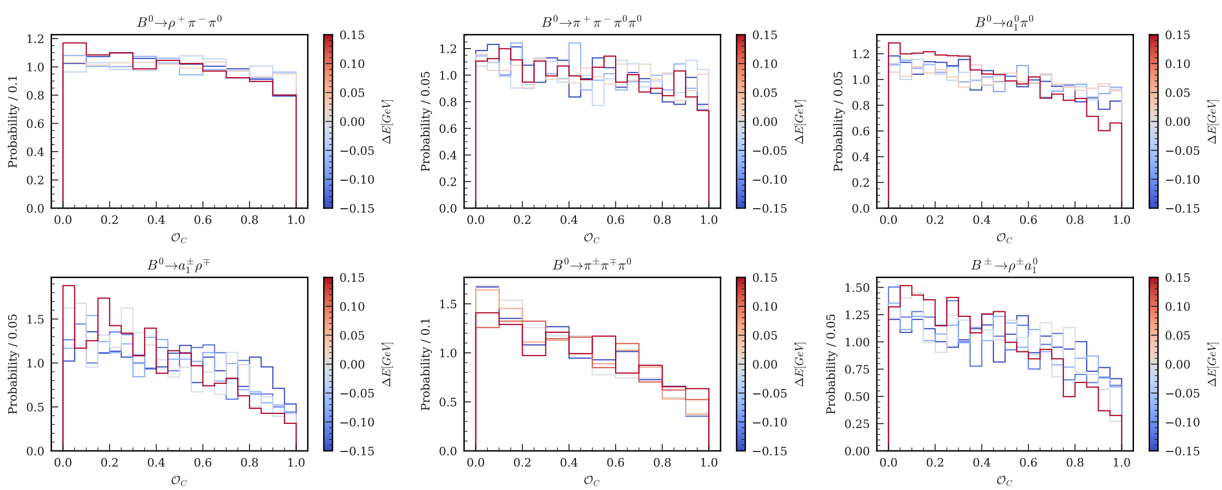


FIG. 157. Correlation between $\mathcal{O}_C$ and ΔE of peaking backgrounds

Appendix 3: CSMVA inputs

We report the distribution of CSMVA inputs. Figure 158 shows the comparison of the signal and $q\bar{q}$ background using MC15rd, and Figure 159 shows the Data-MC comparison of the CSMVA inputs using M_{bc} sideband region ($5.24 < M_{bc} < 5.26$) of $B^0 \to \rho + \rho^-$ signal.

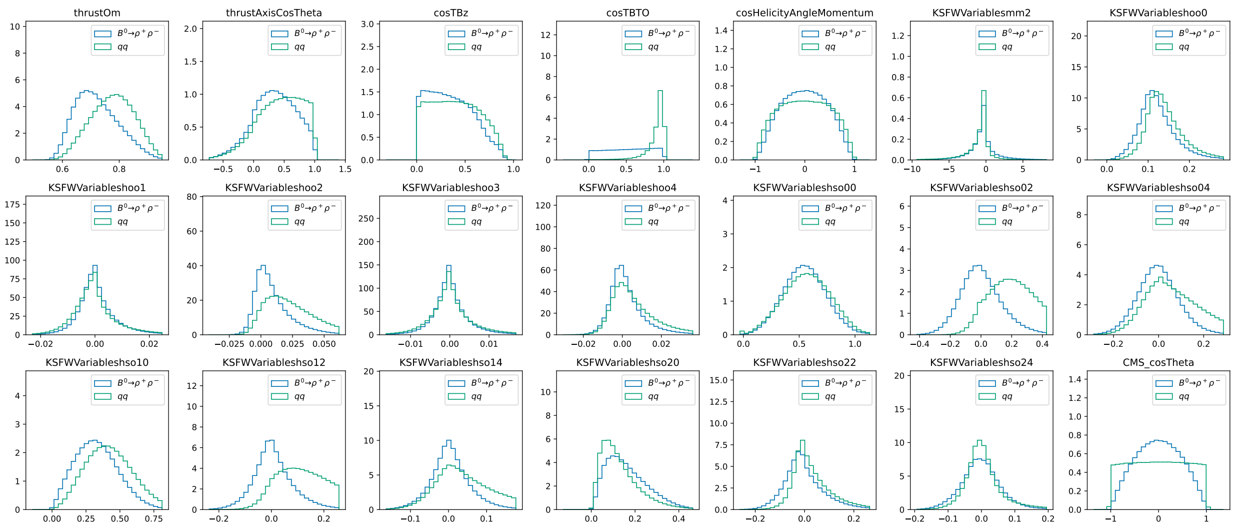


FIG. 158. Distritution of the CSMVA inputs of the signal and $q\bar{q}$

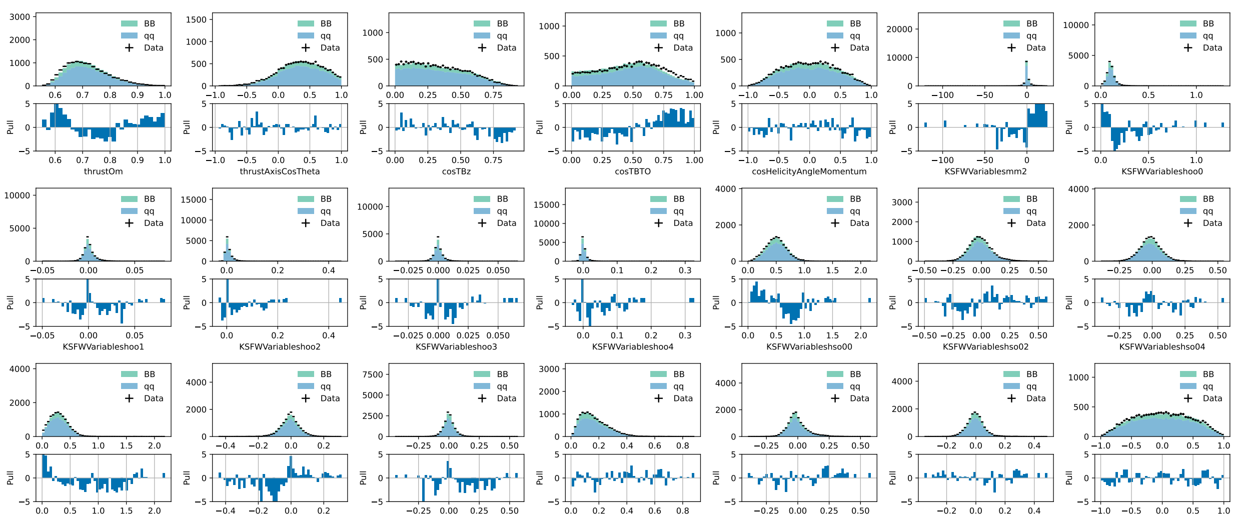


FIG. 159. Data-MC comparison of the CSMVA inputs in the M_{bc} sideband region

Appendix 4: CSMVA output comparison with the control mode

In this analysis we used $B^0 \to D^*\pi$ for the control mode of the CSMVA validation, and we compared the CSMVA output shape of control mode with the signal. Figure 160 shows comparison of the CSMVA output and efficiency.

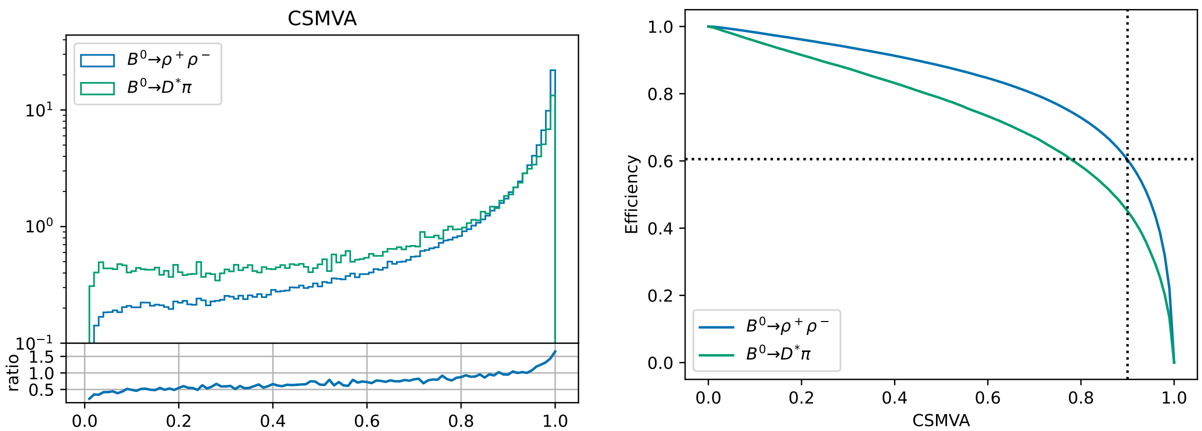


FIG. 160. Comparison of the CSMVA output between $B^0 \to \rho^+\rho^-$ and $B^0 \to D^*\pi$ using MC15rd signalMC samples (Left: CSMVA output, Right: CSMVA efficiency)

Appendix 5: Best candidate selection

The BCS in our analysis is with a sum of the χ^2 of the pi^0 mass kFit and vertex fit, and we compared the performance of it with different BCS. The first one is random BCS, which selects the candidate randomly. Second is BCS with $\sum |m(2\gamma) - m(\pi^0_{\text{true}})|^2$ which was used in the Belle analysis. The final one is BCS with $\prod \text{Prob}(\chi^2)$ which was requested for testing during the review. Table 88 shows the comparison of the BCS performance, and BCS with $\Sigma\chi^2$ which are used in this analysis showed the best performance.

	Number of the signal	Number of SxF	Purity (%)
Before BCS	195965	99349	66.4
$\Sigma\chi^2$	183251	59296	75.6
Random	178742	63241	73.9
$\sum \|m(2\gamma) - m(\pi^0_{\text{true}})\|^2$	181041	61601	74.6
$\prod \text{Prob}(\chi^2)$	183075	59437	75.5

TABLE 88. Comparison of the BCS performance using MC15rd signal MC samples

Appendix 6: qr asymmetry in genericMC samples

Figure 161 shows qr asymmetry in BB and qq MC samples. It is consistent with 0, but there are about 10% or 20% of the uncertainty, and it can be the cause of the bias in the fitter validation for the CP fitting. In this analysis, we assumed there are no CPV effect in BB and qq backgrounds. To remove the effect of the qr bin asymmetry, we performed the bootstrap study while changing the sign of the qr for qq and BB background randomly. Table 89 shows the result of the bootstrap study with random qr sign for qq and BB, and the bias observed in the CP parameters are reduced. The bias observed in the fitter validation for the CP fitting is due to the limited statistics of the generic MC samples, not mis-modelling or issues with the fitter.

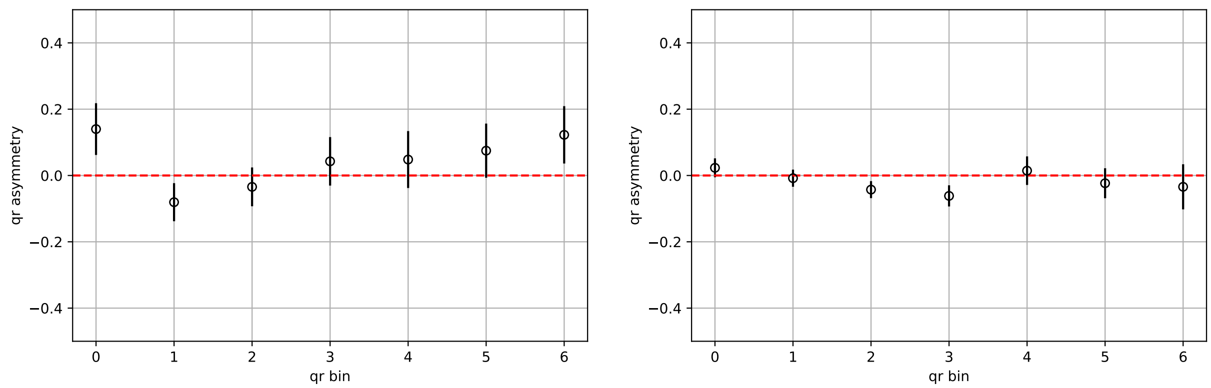


FIG. 161. qr bin asymmetry observed in the run-dependent MC samples. (Left: BB, Right:qq)

Parameter	Mean	Inputs	Mean(Pull)	Std.(Pull)
$\mathcal{B}(\times 10^{-6})$	27.5 $\pm$ 0.1	27.7	-0.088 $\pm$ 0.046	1.021 $\pm$ 0.033
f_L	0.995 $\pm$ 0.001	0.990	0.089 $\pm$ 0.046	1.015 $\pm$ 0.033
S_{cp}	-0.136 $\pm$ 0.006	-0.140	0.030 $\pm$ 0.046	1.020 $\pm$ 0.033
A_{cp}	0.003 $\pm$ 0.004	0.000	0.019 $\pm$ 0.046	1.008 $\pm$ 0.033
$N_{q\bar{q}}$	6278.4 $\pm$ 3.8	6277.1	0.011 $\pm$ 0.046	1.021 $\pm$ 0.033
$N_{B\bar{B}}$	1425.0 $\pm$ 2.9	1444.7	-0.225 $\pm$ 0.047	1.035 $\pm$ 0.034
$N_{B^0\to\rho^+\pi^-\pi^0}$	58.6 $\pm$ 4.4	34.8	0.172 $\pm$ 0.056	1.231 $\pm$ 0.040
$N_{B^0\to\pi^+\pi^-\pi^0\pi^0}$	-28.5 $\pm$ 3.0	11.5	-0.588 $\pm$ 0.058	1.273 $\pm$ 0.042
$N_{B^0\to\pi^\pm\pi^\mp\pi^0}$	11.4 $\pm$ 0.9	12.3	-0.092 $\pm$ 0.048	1.057 $\pm$ 0.034
$N_{B^0\to a_1^0\pi^0}$	84.3 $\pm$ 1.8	42.4	0.749 $\pm$ 0.047	1.037 $\pm$ 0.034
$N_{Rarepeakingbackgrounds}$	54.2 $\pm$ 2.7	45.1	0.075 $\pm$ 0.047	1.039 $\pm$ 0.034

TABLE 89. result of the bootstrap study with random qr sign for qq and BB

Appendix 7: Validation related to resolution function

1 Bias on $S_{\rm cp}$ and $A_{\rm CP}$ due to the difference in the resolution between $B^0 \to \rho^+\rho^-$ and $B^0 \to D^*\pi$

We performed bootstrap tests with the resolution function determined from the residual fit to $B^0 \to \rho^+\rho^-$ signal function and the resolution function from the residual fit to$B^0 \to D^{*-}\pi^+$, to test the difference of the resolution between signal mode and control mode. Figure 162 shows the result of the test, and the difference was less than 2 % of the statistical uncertainty.

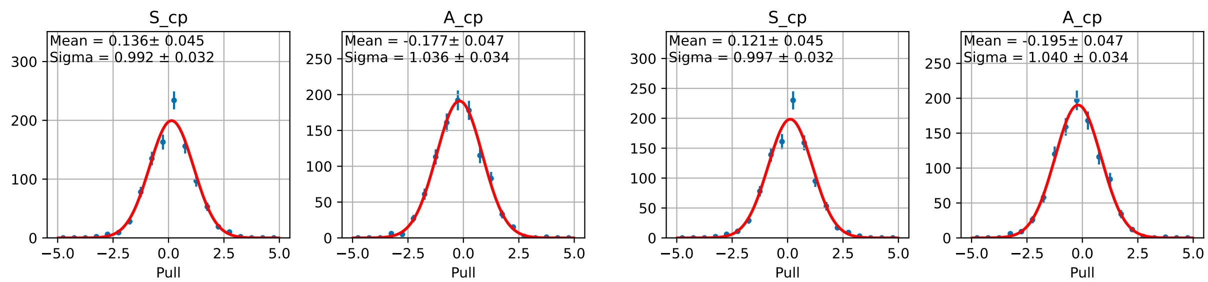


FIG. 162. Result of the bootstrap test. The left one shows the result with a default resolution, and right one shows the result with the resolution determined from $B^0 \to D^*\pi$. The sample is signal MC for the signal and peaking backgrounds, and generic MC(1444 fb^{-1}) for qq and BB backgrounds.

2 Bias on $S_{\rm cp}$ and $A_{\rm CP}$ due to the method of the validation for the resolution

In this analysis, the Δt resolution function parameters are determined by the fitting to Δt distribution of the $B^0 \to D^*\pi$. To validate the procedure, we performed additional tests. Firstly, we performed a test for the resolution function shape measurement. In this test, we performed fitting to the bootstrap samples of $B^0 \to D^*\pi$ MC15rd, and checked the fitted resolution function parameters. In this validation, the background components have a very low statistics, so we generated TOYMC samples for it. Figure 163 shows the result of the bootstrap and TOY test for resolution function fitting. The fitted parameters are slightly biased. To check the impact on CP parameters due to the procedure to validate the resolution function shape, we performed a bootstrap and TOY study using the default resolution and the resolution obtained from the bootstrap and TOY study of $B^0 \to D^*\pi$, and the result is shown in Figure 164.

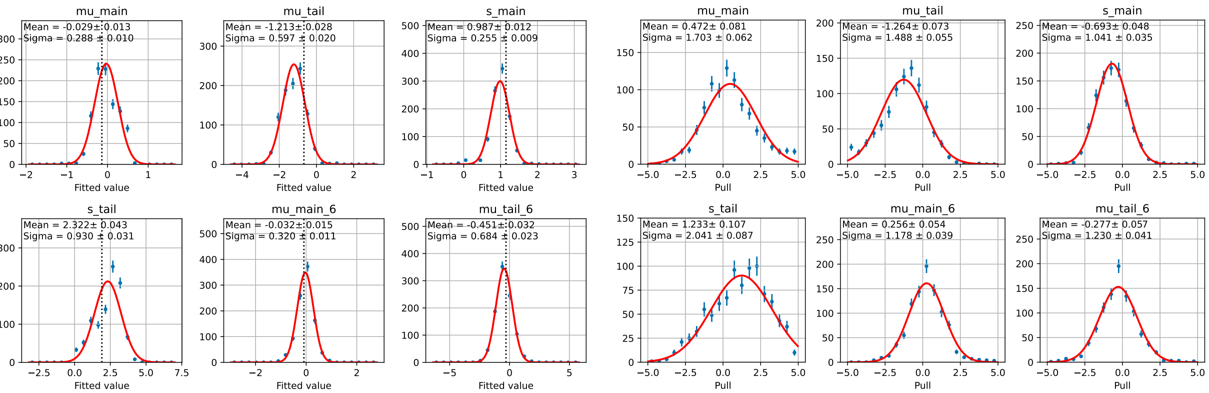


FIG. 163. Result of the bootstrap + TOY study for the Δt fitting for $B^0 \to D^*\pi$. (Left: Fitted value, black dotted line shows the fit result to the $B^0 \to \rho^+\rho^-$ signal MC samples. Right: Pull distribution of the right plot.)

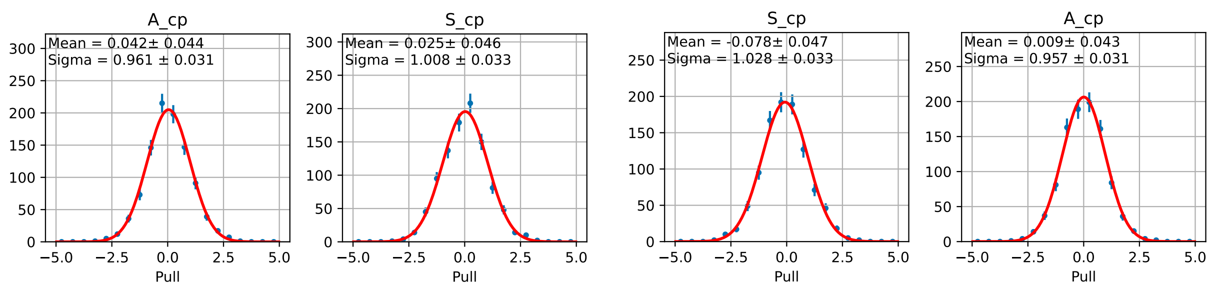


FIG. 164. Fitter validation using bootstrap for the signal and TOYMC for backgrounds. Left one is the result With the default resolution function, and right one shows the result with the resolution function determined in the bootstrap+TOY study for $B^0 \to D^*\pi$.

Appendix 8: Punzi effect due to the difference in the $\sigma(\Delta t)$ distribution between signal and background

In the fitting, the PDF of the $\sigma(\Delta t)$ is not included while the resolution function shape is changed depending on it. It can be the cause of the bias due to the Punzi effect. Figure 165 shows the difference in $\sigma(\Delta t)$ distribution to the signal samples. To validate the effects from the Punzi effect, we performed a bootstrap test with a default settings and with a $\sigma(\Delta t)$ PDF. Figure 166 shows the result, and the bias was very small. Table 90 shows the result and bias due to the Punzi effect from $\sigma(\Delta t)$ distribution, and it was very small.

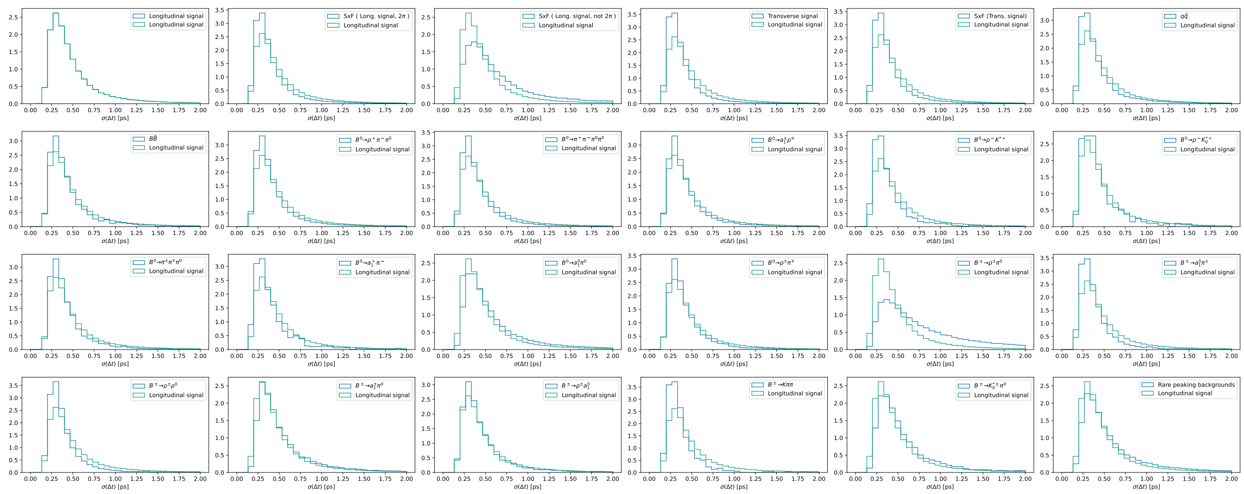


FIG. 165. The difference in the $\sigma(\Delta t)$ distrubution using MC15rd samples.

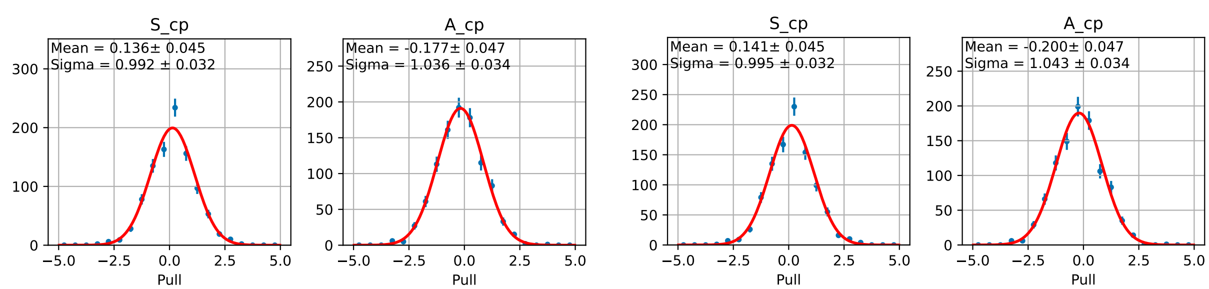


FIG. 166. Result of the bootstrap test. The left one shows the result with a default settings, and the right one is the bootstrap with a $\sigma(\Delta t)$ PDF. The sample is signal MC for the signal and peaking backgrounds, and generic MC(1444 fb^{-1}) for qq and BB backgrounds.

Parameter	With $\sigma(\Delta t)$ PDF	Default	Bias
S_{CP}	-0.1154	-0.1172	0.0018
A_{CP}	-0.0247	-0.0213	0.0034

TABLE 90. Comparison of the bootstrap test result with and without $\sigma(\Delta t)$ PDF.

Appendix 9: Bias due to $\cos\theta^*_B$ PDF

1 Punzi effect due to the difference in the $\cos\theta^*_B$ PDF

In the fitting, the PDF for $\cos\theta^*_B$ is not considered, but several fit components show a difference from the longitudinal signals as shown in Figure 167. To understand the bias due to the Punzi effect by ignoring the $\cos\theta^*_B$ PDF, we performed a bootstrap study with and without $\cos\theta^*_B$ PDF. Table 91 shows the bias due to the Punzi effect from $\cos\theta^*_B$ PDF, and they were negligible compared to the expected statistical uncertainty.

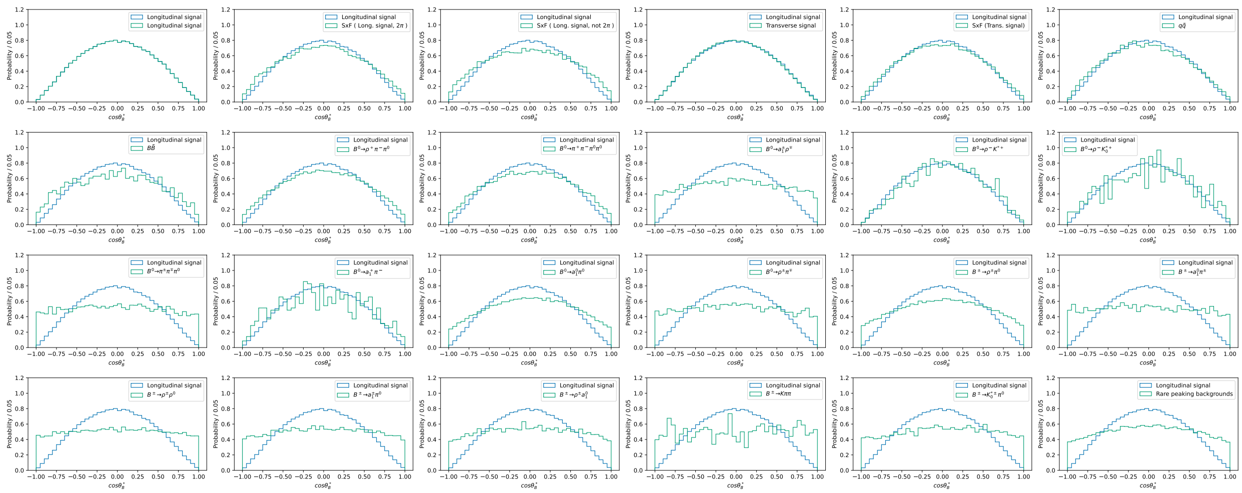


FIG. 167. Difference in the $cos\theta^*_B$ distribution in MC15rd samples

Parameter	With $\cos\theta^*_B$ PDF	Default	Bias
$S_{\rm CP}$	-0.1210	-0.1172	0.0038
$A_{\rm CP}$	-0.0202	-0.0213	0.0011

TABLE 91. Bias due to Punzi effect by ignoring $\cos\theta^*_B$ PDF.

2 Bias due to $\cos\theta^*_B$ PDF shape

In this analysis, we assume $\cos\theta^*_B$ PDF shape to be Equation 91, but the $\cos\theta^*_B$ showed a small difference from the PDF due to the event selection as shown in 36. To understand the bias due to the difference, we tested the fitter with a default $\cos\theta^*_B$ PDF and $\cos\theta^*_B$ PDF shape taken from the distribution of the signal MC. Table 92 shows the difference, and the bias due to the $\cos\theta^*_B$ PDF shape difference is negligible.

$$\mathcal{P}(\cos\theta^*) = \frac{3}{4}(1-\cos^2\theta^*) \tag{91}$$

Parameter	With corrected $\cos\theta_B^*$ PDF	Default	Bias
S_{CP}	-0.1173	-0.1172	0.0001
A_{CP}	-0.0211	-0.0213	0.0002

TABLE 92. Bias due to approximation of the $\cos\theta_B^*$ PDF

Appendix 10: Additional tests for the $cos\theta_\rho$ PDF

This section summarizes the checks related to $cos\theta_\rho$ which was done before the box opening.

1 CSMVA cuts

In the pre-box opening step, we found that a relatively large number of the $\tau\tau$ background can contribute to the signal and M_bc sideband regions. It comes from $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$, $\tau^- \to \pi^-\pi^0\nu_\tau$, and $\tau^- \to \pi^-\pi^0\pi^0\nu_\tau$, which account for more than 90% of the $\tau\tau$ background. The angular distributions of them have not been measured and can make a large bias in the $cos\theta_\rho$ distributions. To suppress them, we tighten the continuum-suppression cut from $\mathcal{O}_C > 0.9$ to $\mathcal{O}_C > 0.96$.

	$\mathcal{O}_C > 0.9$	$\mathcal{O}_C > 0.96$
qq	10039.8	3529.7
BB	2329.3	1305.7
$\tau\tau$	336.6	71.6

TABLE 93. Number of events in the M_{bc} sideband region (5.24 GeV $< \mathrm{M_{bc}} <$ 5.27 GeV) estimated using MC15rd samples

2 $cos\theta_\rho$ PDF correction

In the fitting, we corrected the $cos\theta_\rho$ for qq background using the data-MC ratio of the P_π distribution observed in the sideband. In the correction, we assumed that there are no Data-MC differences in the BB backgrounds.

We also tested three alternative $cos\theta_\rho$ corrections to see the impact of the correction. The data-MC corrections for $cos\theta_\rho$ are

1. qq correction using the sideband data-MC ratio
2. qq and BB correction using the sideband data-MC ratio
3. no correction
4. qq correction using the off-resonance data

The first one is the default correction. The second one is an alternative correction using the sideband data, and we assumed that there is the same data-MC discrepancy in both

qq and BB, and we weighted the PDF shape using the data-MC ratio in P_π distribution. The third one means no correction, all $cos\theta_\rho$ PDF shapes are fixed to MC. The fourth one is the qq correction using an off-resonance sample, we corrected $cos\theta_\rho$ PDF shape using the data-MC ratio in P_π distribution. As discussed 6.5, the data-MC ratio in P_π observed in off-resonance is consistent with the data-MC ratio observed in the sideband data assuming there are no data-MC discrepancies in BB.

We performed fittings to 364 fb^{-1} data while changing the data-MC corrections masking the fit results, and only checking the likelihood. Table 94 shows the likelihood for each correction, and the default correction was the best. In principle, the qq correction using the off-resonance should also have a better likelihood, however, it shows the worst because of the poor statistics. Comparing the default correction with the second and third corrections, we found that the major source of the discrepancy comes from qq. The fourth one has a large statistical uncertainty in the correction. Therefore, we decided to use the qq correction from the sideband data.

Correction	-2$log\mathcal{L}$
qq correction (sideband)	-55122.7
$qq + BB$ correction (sideband)	-55117.8
No correction (Fix to MC)	-55114.1
qq correction (off-resonance)	-55113.8

TABLE 94. Likelihood of the signal-extraction fitting to 364 fb^{-1} dataset for each data-MC corrections

[1] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
[2] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).
[3] J. Charles, A. Höcker, H. Lacker, S. Laplace, F. R. Le Diberder, J. Malclés, J. Ocariz, M. Pivk, and L. Roos, The European Physical Journal C **41**, 1–131 (2005).
[4] J. Charles, S. Descotes-Genon, Z. Ligeti, S. Monteil, M. Papucci, and K. Trabelsi, Phys. Rev. D **89**, 033016 (2014).
[5] J. Dalseno *et al.* (Belle Collaboration), Phys. Rev. D **88**, 092003 (2013).
[6] M. Gronau and D. London, Phys. Rev. Lett. **65**, 3381 (1990).
[7] J. Charles, O. Deschamps, S. Descotes-Genon, and V. Niess, The European Physical Journal C **77** (2017), 10.1140/epjc/s10052-017-5126-9.
[8] R. L. Workman *et al.* (Particle Data Group), PTEP **2022**, 083C01 (2022).
[9] T. Abe *et al.*, "Belle ii technical design report," (2010), arXiv:1011.0352 [physics.ins-det].
[10] B. I. Collaboration, "A new graph-neural-network flavor tagger for belle ii and measurement of $\sin 2\phi_1$ in $b^0 \to j/\psi k^0_\mathrm{s}$ decays," (2024), arXiv:2402.17260 [hep-ex].
[11] F. Abudinén *et al.*, The European Physical Journal C **82** (2022), 10.1140/epjc/s10052-022-10180-9.
[12] S. O. Arik and T. Pfister, "Tabnet: Attentive interpretable tabular learning," (2020), arXiv:1908.07442 [cs.LG].
[13] R. Storn and K. Price, Journal of Global Optimization **11**, 341 (1997).
[14] A. Hocker *et al.*, arXiv (2007), arXiv:physics/0703039.
[15] B. Aubert *et al.* (BaBar), Phys. Rev. D **74**, 031104 (2006), arXiv:hep-ex/0605024.
[16] F. Abudinén *et al.* (Belle II Collaboration), Phys. Rev. D **107**, L091102 (2023), arXiv:2302.12791 [hep-ex].
[17] T. Humair *et al.*, B2Note (2021).
[18] G. Punzi, "Comments on likelihood fits with variable resolution," (2004), arXiv:physics/0401045 [physics.data-an].
[19] M. Pivk and F. R. Le Diberder, Nucl. Instrum. Meth. **A555**, 356 (2005), arXiv:physics/0402083 [physics.data-an].
[20] S. Choudhury and et al. (Belle Collaboration), Phys. Rev. D **107**, L031102 (2023).
[21] M. Gronau and D. London, Phys. Rev. Lett. **65**, 3381 (1990).
[22] Vanhoefer *et al.* (Belle Collaboration), Phys. Rev. D **93**, 032010 (2016).
[23] R. Aaij *et al.*, Journal of High Energy Physics **2016** (2016), 10.1007/jhep12(2016)087.
[24] S. Choudhury *et al.* (Belle), Phys. Rev. D **107**, L031102 (2023), arXiv:2207.01194 [hep-ex].
[25] T. Aushev *et al.*, "Physics at super b factory," (2010), arXiv:1002.5012 [hep-ex].
[26] A. Lenz, U. Nierste, J. Charles, S. Descotes-Genon, A. Jantsch, C. Kaufhold, H. Lacker, S. Monteil, V. Niess, and S. T'Jampens (CKMfitter Group), Phys. Rev. D **83**, 036004 (2011).
[27] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," (2019), arXiv:1907.10902 [cs.LG].