Doctoral Dissertation
博士論文

Measurement of Time-Dependent CP Violation in Decays of $B^0 \to \eta' K^0_S$ at the Belle II Experiment

(Belle II 実験における $B^0 \to \eta' K^0_S$ 崩壊過程での時間依存 CP 非対称度測定)

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

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

BAE Hanwook
裵　漢郁

Acknowledgments

My doctoral thesis was not written only by my endeavors but also by many guidance and support from many good people around me. I would like to express my sincere thank in this acknowledgment to those who have helped me on the journey for this dissertation and my academic life so far, and a genuine apology to all people whom I forgot to mention in this short letter.

First of all, words cannot express my gratitude to my academic supervisor Prof. Yutaka Ushiroda for his invaluable patience and feedback. I could not have undertaken my journey toward Ph.D. without his guidance, who generously provided knowledge and expertise.

Also, I am deeply indebted to a dissertation committee chaired by Prof. Yasuyuki Okumura since they gave me many positive criticisms about the dissertation. Thanks to their valuable comments and advice, I could improve the quality of my dissertation and finalize it.

I cannot begin to express my sincere gratitude to Dr. Koji Hara and Dr. Kazutaka Sumisawa, who gave enormous advice on my studies and practical solutions when I faced problems. It was highly precious to me, and I only finished my *CP* violation analysis with their assistance.

I would also like to extend my deepest thanks to my friends and colleagues in the Ushiroda laboratory. I received incredible inspiration and refreshment during the COVID-19 pandemic from an academic discussion, having dinner, drinking liquor, and any other memories with Yuma-chan and Dr. Hikaru Tanigawa. My colleagues in the laboratory, Dr. Yu Nakazawa, Dr. Yo Sato, Dr. Seokhee Park, Dr. Takuto Kunigo, Mr. Ryohei Sugiura, Mr. Zihan Wang, and all other members also gave me much help and refreshed me during my doctoral course.

Many thanks to members of the Japanese group for Time-Dependent *CP* Violation studies in Belle II for their helpful advice. Prof. Kenkichi Miyabayashi and Dr. Akimasa Ishikawa positively criticized my analysis and corrected me when I said the wrong thing. In addition, the Japanese members of Belle II corrected the wrong Japanese sentences in my presentation material for conferences or weekly meetings so that I could improve the overall quality of the material.

I also wish to thank a working group of Belle II for Time-Dependent *CP* Violation studies convened by Prof. Takeo Higuchi and Dr. Thibaud Humair for many comments on my analysis and for reading a draft of my internal note. Also, Prof. Stefano Lacaprara provided essential tools for event reconstruction with a detailed description. Thanks to all the working group members, I enhanced the quality of my analysis for this dissertation.

I am highly grateful to the Belle II trigger team members. Dr. Yoshihito Iwasaki and Dr. Taichiro Koga gave me great support and expertise on the trigger system at my initial stage in Belle II, and I was able to develop firmware for UT4 thanks to their assistants. Also, Mrs. Eunji Jang in the team provided me with her delicious food and some beverage when I felt tired from my studies in Japan.

I would like to thank the secretaries of the Belle II project for their administrative help so I could perform my research without any worries about official businesses. Also, officials of the National Institute for International Education in Korea provided many assistants for my scholarship. I could continue studying abroad in Japan thanks to their financial support.

I feel grateful to the members of Prof. Lee's laboratory at Kyungpook National University (KNU) during my master's course, Prof. Sehwook Lee, Dr. Wooyoung Jang, Dr. Sangil Park, Mr. Changgi Huh, Mrs. Bobae Kim, and Mr. Ryonghae Ye, who provided me with a lot of help and comments for my studies of the Geant4 simulation.

Thanks should also go to my friends who supported me while studying abroad in Japan. Dr. Kookhyun Kang gave some clues to solve problems in my analysis and refreshed me with his joke and beer when I was down. I thank Mr. Woonghyun Jeung, who provided me with delicious gourmet and lovely scenery around Yamanashi-ken. 저스티스 and my friends in middle and high school, the Department of Physics at KNU, and Center for Underground Physics at ibs provided great emotional support by calling me or playing games with me occasionally so that I feel less lonely even if I was far from them. Especially, Mr. Gwangsoo Kim, Mr. San Park, Mr. Sangwook Lee, Mr. Wonyeung Lee, Mr. Hyojune Lee, Mr. Doyeong Kim, Mr. Sungjune Lee, and Mr. Jaemin Bang visited and played with me in Japan; It will be hard to forget good memories with them during my difficult times.

Finally, I can't thank my family enough for their love and support for many years. I always knew that my father believed I could do my best and achieve my goal of studying abroad in Japan. I am also extremely grateful to my brother and sister; they supported not only me but also my father, who felt lonely due to my absence. I owed much to the Fujie family during my life in Japan and will never forget their devotional support. This journey would not have been possible without the support of all the people stated above (and all the people I might have forgotten to mention), and I would like to say to them here: 정말 고마워요, 모두들.

Sincerely,
Hanwook BAE

Tokyo, Japan
August 11, 2023.

Abstract

$B^0 \to \eta' K^0_S$ channel is a decay mode of neutral B-meson, which is mediated by the penguin process of $b \to sq\bar{q}$ of the Standard Model, and we can observe the CP violation phenomenon in the $B^0 \to \eta' K^0_S$ channel. The CP violation can be parameterized by the CP asymmetries, which consists of two real parameters of A_{CP} and S_{CP}. The parameters can be obtained in a systematical discussion for the difference in the temporal decay rate between B^0 and $\overline{B}^0$ by analyzing the difference in a time-dependent fashion. Also, the CP violation is deeply related to the matter-antimatter asymmetry, which is one of the unsolved problems that the Standard Model cannot describe. We expect a new theory, called New Physics, beyond the Standard Model to solve this problem, and New Physics might modify phenomena of the CP violation in the $B^0 \to \eta' K^0_S$ decay from those described by the Standard Model using the CKM matrix. We can explore New Physics by comparing measurements for parameters of the CP violation between those in $B^0 \to \eta' K^0_S$ decay mode and tree-dominating one such as $B^0 \to J/\psi K^0_S$ decays. In this aspect, extracting CP parameters (A_{CP}, S_{CP}) in $B^0 \to \eta' K^0_S$ is a key measurement of Belle II, an experiment for particle physics with asymmetrically accelerated electron and positron. We measure the CP parameters in a sub-decay mode of $\eta' \to \eta\pi^+\pi^-$, $\eta \to \pi^+\pi^-\pi^0$, and $K^0_S \to \pi^+\pi^-$ of $B^0 \to \eta' K^0_S$ channel at the Belle II experiment with experimental data accumulated between 2019 March and 2022 June corresponding integrated luminosity of $362\,\mathrm{fb}^{-1}$. For the measurement, we should consider the various factor of the reconstruction, such as the inaccuracy of decay time determination for B-mesons or effects from background events. We develop a resolution function and method to calibrate and validate the function parameters using the experimental data. In detail, we developed a resolution function that considers the inaccuracy in determining the decay time and is an essential component for CP asymmetry studies for all the subchannels of the $B^0 \to \eta' K^0_S$ channel. A method to calibrate the parameters of the resolution function is also devised, which has more advantages than conventional calibration methods and is an original methodology developed in this study firstly. The measured CP parameter (A_{CP}, S_{CP}) from this study is

$$A_{CP} = -0.111^{+0.319}_{-0.311}(\mathrm{stat}) \pm^{+0.032}_{-0.031}(\mathrm{syst}), \text{ and}$$
$$S_{CP} = +0.248^{+0.470}_{-0.526}(\mathrm{stat}) \pm^{+0.053}_{-0.053}(\mathrm{syst}).$$

We confirm that the measured CP parameters are consistent with the world average values of $B^0 \to \eta' K^0$ decay or that of $B^0 \to c\bar{c}K^0$ decay considered reference values of the Standard Model prediction, within the uncertainty. We also found that when measuring CP asymmetries using all subchannels in the $B^0 \to \eta' K^0_S$ channel by the method developed in this study, the statistical error of A_{CP} and S_{CP} at $50\,\mathrm{ab}^{-1}$, the target luminosity of Belle II, was 0.01 or less, which is lower than the current theoretical prediction of CP asymmetries in the channel.

Personal Contribution

This analysis is a result of collaborative work from the Belle II collaboration. All the dataset of experimental data and simulation samples used in this analysis is produced using the Belle II detector, peripheral components for the data collection, and computing cluster for data processing and simulations, summarized in Ch. 2. Also, I used the standard analysis software frameworks widely used in the Belle II collaboration, such as flavor tagger, vertex fitter, and toolkit for estimations of systematic uncertainties. I completed the analysis of *CP* parameter measurements in the $B^0 \to \eta' K^0_S$ decay with a subchannel of $\eta' \to \eta(\pi^+\pi^-\pi^0)\pi^+\pi^-$ and $K^0_S \to \pi^+\pi^-$ by developing a specific methodology optimized to measure *CP* asymmetries in the target channel based upon an understanding of theory for *CP* violation given in Ch. 1.

In this work, I developed an algorithm to reconstruct the decay from the experimental data described in Ch. 4 and optimized criteria to obtain an optimal reconstruction result for the *CP* asymmetry measurements from the experimental data, especially optimizing selection criteria for π^0 and K^0_S, which is described in App. E, and η' selection using the mass difference between η' and η, which are newly introduced in this analysis. Then, I devised a method to extract the probability of signal events with 5-D variables, described in the same chapter, and is very sensitive for signal and background events so that the sensitivity of our measurement is unaffected even with more background events, as discussed in App. L. On top of that, I configured a resolution function and a probability density function for background events and determined their parameters using the Belle II data or simulation samples, which is described in Ch. 5. Especially, the calibration method of the parameters of the resolution function using cosmic samples is the first attempt to understand the reason for the structure of the resolution function and correct the difference in the resolution between simulation and data. In addition, I built a procedure for validating the resolution parameters using the control channel of $B^\pm \to \eta' K^\pm$ by emulating the vertex resolution of $B^0 \to \eta' K^0_S$. Also, I improved a program to estimate the *CP* asymmetries so that the program can consider the effects of the motion of B-mesons in the center of mass frame of $\Upsilon(4S)$ in the estimation procedure, which reduces the systematic errors of the *CP* parameters measurement. This improvement will be included in the Belle II official software for *CP* violation studies. Finally, in Chs. 6 and 7, I established a method to estimate the systematic uncertainties of the measurement and predicted the expected sensitivity of *CP* violation measurement in this channel at integrated luminosity $50\,\mathrm{ab}^{-1}$, which is a target luminosity of Belle II, with quantitative discussions on prospects of Belle II. Based on the discussion, I examined the details of each item for systematic error and established a reasonable method for extrapolation. Then, I proved that Belle II can measure *CP* asymmetries in the $B^0 \to \eta' K^0_S$ channel at $50\,\mathrm{ab}^{-1}$ with a smaller uncertainty than the current world average and observe New Physics with high sensitivity comparable to the current theoretical prediction. Therefore, I contributed to the analysis to build a reliable methodology determining the *CP* asymmetries with $50\,\mathrm{ab}^{-1}$. I hereby hope and believe those improvements, developments, and discussions will be highly useful for Belle II in the future measurement of *CP* asymmetries in the $B^0 \to \eta' K^0_S$ channel as well as in other channels.

Outline of This Dissertation

In this dissertation, we introduce the physics motivation of the *CP* violation and New Physics with it in Ch. 1. After the introduction, we discuss the experimental apparatus to collect the experimental data in Ch. 2, then explain the current data-taking status of the Belle II experiment. In Ch. 3, we introduce an overview of the methodology for the estimation in the experimental environment using the theory discussed in the previous chapter. After that, we explain how to reconstruct our target decay for the analysis and estimate the yield of signal events in Ch. 4. Then, in the following chapter of Ch. 5, we prepare ingredients for a fitting procedure to determine A_{CP} and S_{CP} from the reconstruction results. In the next chapter of Ch. 6, we build the fitting procedure using the ingredients and validate it using various methods, and finally, we describe the *CP* violation parameters with systematic uncertainties from the Belle II experimental data. In the last two chapters of Ch. 7, we discuss the result and prospect of this analysis with the experimental data from Belle II in the future and the conclusion of this dissertation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Standard Model (SM) has successfully described many phenomena in nature. However, the Standard Model cannot answer problems such as dark matter, the mass of neutrinos, and the quantization of electric charges. Due to these unexplained problems, we suppose there is New Physics (NP) beyond SM. Observation of hints for New Physics is the most major goal for particle physicists to solve these remaining problems.

One big unsolved problem of the Standard Model is the known matter-antimatter asymmetry in the universe. In particle physics, matter can be converted into antimatter by the *CP* transformation; whether the *CP* symmetry is breaking (*CP* violation) or not is deeply related to the problem. This provides a good test bench to examine New Physics because although the *CP* violation in the quark sector is measured at the kaon system [1], and Kobayashi and Maskawa described this phenomenon with the third generation of quarks and foreseen a large *CP* violation at *B*-mesons[2, 3], which is actually measured in the neutral *B*-meson system[4], the estimated amount of the violation is insufficient to explain the asymmetry of matter and anti-matter. It is not enough to explain the asymmetry even though we take into account the *CP* violation from the lepton sector[5]. It implies we need an additional source of the *CP* violation to explain the matter-antimatter asymmetry in the universe. Meanwhile, if New Physics exists, it will affect *CP* violation and make a deviation of the measurement of the violation from the prediction of the Standard Model. Therefore, precisely examining the *CP* violation parameters will provide a clue to discovering New Physics beyond the Standard Model.

We describe how to measure parameters of the time-dependent *CP* violation and the result and discussion of the measurement in this dissertation. A decay process of $B^0 \to \eta' K^0_S$ with a subchannel of $\eta' \to \eta\pi^+\pi^-$ and $\eta \to \pi^+\pi^-\pi^0$ is considered. The main feature of $B^0 \to \eta' K^0_S$ decay channel is that there is a prediction from New Physics on *CP*-violating parameters on the channel, and we can examine the contribution of New Physics by comparing the *CP*-violating parameters between this channel and charmonium channels, such as $B^0 \to J/\psi\, K^0_S$. In this chapter, we describe the *CP* violation of the Standard Model and how it appears in the decay of $B^0 \to \eta' K^0_S$, then explain the effects of New Physics on the *CP* violation in the decay.

1.1 *CP* violation in the Standard Model

Time-dependent *CP* violation in the Standard Model is related to the Cabibbo-Kobayashi-Maskawa (CKM)[2, 6] matrix and the mixing of B^0-$\overline{B}^0$. We review the CKM matrix briefly and discuss how the *CP* violation occurs in the mixing and decay to the *CP*-eigenstate of neutral *B*-mesons, then we will discuss a specific situation for the *CP* violation in our target decay, $B^0 \to \eta' K^0_S$, in the following subsections.

1.1.1 Cabibbo-Kobayashi-Maskawa Matrix

Cabibbo-Kobayashi-Maskawa (CKM) matrix is a 3×3 unitary matrix which describes the quark mixing by the weak interaction. The weak eigenstates of the down-type quarks (d', s', b') relate to the mass eigenstate of the down-type quarks (d, s, b) via the matrix as follows:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.1}$$

Since the matrix is unitary and five phases can be absorbed into the relative phase of the wave function of six quarks, we get four free parameters of the matrix. The Wolfenstein parameterization is often used to expand the matrix in powers of sine value of Cabibbo angle, $\lambda \equiv \sin(\theta_C) \approx 0.22$ as follows:

$$V = \begin{pmatrix} 1 - \lambda/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{1.2}$$

where A, ρ, and η are the real parameters of which order is $\mathcal{O}(1)$[7].

Unitarity conditions for off-diagonal elements of the CKM matrix build six triangles on the complex plane. Among the triangles, a triangle indicated by the following condition is proper for the actual measurement as all the three terms are of the same order, $\mathcal{O}(A\lambda^3)$.

$$V_{ud}V^*_{ub} + V_{cd}V^*_{cb} + V_{td}V^*_{tb} = 0 \tag{1.3}$$

By dividing Eq. 1.3 by $V_{cd}V^*_{cb}$, we can get a triangle shown in Fig. 1.1 that is called a "Unitarity triangle".

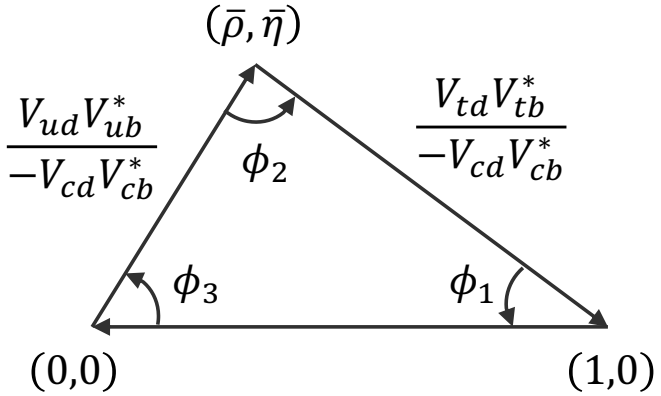


Figure 1.1: Unitarity triangle with interior angles of ϕ_1, ϕ_2, and ϕ_3.

The coordinate of apex $\bar{\rho}$ and $\bar{\eta}$, interior angles of ϕ_1, ϕ_2, and ϕ_3 are given by following equations:

$$\phi_1 = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \tag{1.4}$$

$$\phi_2 = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \tag{1.5}$$

$$\phi_3 = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \tag{1.6}$$

$$\bar{\rho} = \left(1-\frac{\lambda^2}{2}\right)\rho \tag{1.7}$$

$$\bar{\eta} = \left(1-\frac{\lambda^2}{2}\right)\eta \tag{1.8}$$

Among these three angles, our measurement for the channel of $B^0 \to \eta' K_S^0$ has sensitivity on ϕ_1 value that can be compared with the results from the decay of $B^0 \to J/\psi\, K_S^0$ that is the most precisely determined through the time-dependent CP violation. We can check the existence of New Physics by comparing the results.

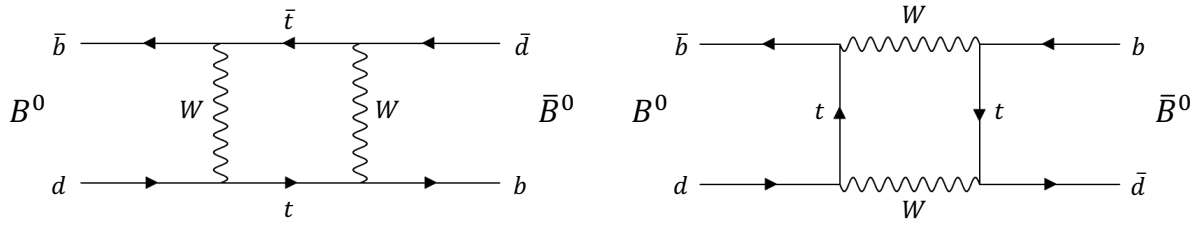


Figure 1.2: Two types of box diagram for $B^0 - \overline{B}^0$ mixing by the weak interaction.

1.1.2 Mixing and time evolution of neutral *B*-mesons

Neutral B-mesons of B^0 and $\overline{B}^0$ mix through the second-order weak interaction shown in Fig. 1.2. This means that the neutral B-mesons oscillate between B^0 state and $\overline{B}^0$ state over the course of time. Let us consider the time evolution of a B-meson system from an initial state with the superposition of B^0 and $\overline{B}^0$, which can be written as $|\psi(0)\rangle = a(0)\,|B^0\rangle + b(0)\,|\overline{B}^0\rangle$. If we limit our interests to the subspace of $|B^0\rangle$ and $|\overline{B}^0\rangle$, the state vector of the system at time t will be $|\psi(t)\rangle = a(t)\,|B^0\rangle + b(t)\,|\overline{B}^0\rangle$ and be governed by the following equation:

$$i\frac{d}{dt}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \mathcal{H}'\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \tag{1.9}$$

with the flavor eigenstate basis of $|B^0\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$ and $|\overline{B}^0\rangle = \begin{pmatrix} 0 & 1 \end{pmatrix}^T$. The matrix $\mathcal{H}'$ is an effective Hamiltonian to express the $B^0 - \overline{B}^0$ mixing and consists of two Hermitian matrices, which are the mass matrix $\mathbf{M}$ and the decay matrix $\mathbf{\Gamma}$, as follows:

$$\mathcal{H}' = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{21} - \frac{i}{2}\Gamma_{21} & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}. \tag{1.10}$$

where $M_{21} = M_{12}^*$ and $\Gamma_{21} = \Gamma_{12}^*$.

The $M_{11(22)}$ and $\Gamma_{11(22)}$ means the mass and decay width of the flavor eigenstate of B^0 ($\overline{B}^0$), respectively, and we can write $M_{11} = M_{22} = M$ and $\Gamma_{11} = \Gamma_{22} = \Gamma$ with assuming the CPT invariance. We can get the two mass eigenstates to solve the above equation by diagonalizing the matrix $\mathcal{H}'$. The eigenvector $|B_{1,2}\rangle$ and eigenvalue $\lambda_{1,2}$ are as follows:

$$|B_1\rangle = p\,|B^0\rangle + q\,|\overline{B}^0\rangle\,,\ \text{for}\ \lambda_1 = M - \frac{i}{2}\Gamma + \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}, \tag{1.11}$$

$$|B_2\rangle = p\,|B^0\rangle - q\,|\overline{B}^0\rangle\,,\ \text{for}\ \lambda_2 = M - \frac{i}{2}\Gamma - \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}, \tag{1.12}$$

where

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}, \tag{1.13}$$

and $\sqrt{|p|^2 + |q|^2} = 1$ for the normalization. Meanwhile, we adjust $\lambda_{1,2}$ using q/p as follows:

$$\lambda_1 = \left(M - \frac{i}{2}\Gamma\right) + \frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right) = M_1 - \frac{i}{2}\Gamma_1, \tag{1.14}$$

$$\lambda_2 = \left(M - \frac{i}{2}\Gamma\right) - \frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right) = M_2 - \frac{i}{2}\Gamma_2, \tag{1.15}$$

where $M_{1,2} = M \pm (q/p)M_{12}$ and $\Gamma_{1,2} = \Gamma \pm (q/p)\Gamma_{12}$. In this form, we can write the time evolution of the mass eigenstates of $|B_{1,2}(t)\rangle$ from an initial state of $|B_{1,2}(0)\rangle$ as follows:

$$
\begin{aligned}
|B_1(t)\rangle &= e^{-i\left(M_1 - \frac{i}{2}\Gamma_1\right)t} |B_1(0)\rangle \,, \text{ and} && (1.16)\\
|B_2(t)\rangle &= e^{-i\left(M_2 - \frac{i}{2}\Gamma_2\right)t} |B_2(0)\rangle \,. && (1.17)
\end{aligned}
$$

In this context, what we want is the time evolution of the "flavor" eigenstate, and this can be obtained by inverting Eqs. 1.11 and 1.12, and put Eqs. 1.16 and 1.17 to the inverted equations as follows:

$$
\begin{aligned}
|B^0(t)\rangle &= \frac{1}{2p}(|B_1(t)\rangle + |B_2(t)\rangle) \\
&= \frac{1}{2p}\left(e^{-i\left(M_1 - \frac{i}{2}\Gamma_1\right)t} |B_1(0)\rangle + e^{-i\left(M_2 - \frac{i}{2}\Gamma_2\right)t} |B_2(0)\rangle\right) \\
&= \frac{1}{2p}\left[e^{-i\left(M_1 - \frac{i}{2}\Gamma_1\right)t} \left(p\,|B^0\rangle + q\,|\overline{B}^0\rangle\right) + e^{-i\left(M_2 - \frac{i}{2}\Gamma_2\right)t} \left(p\,|B^0\rangle - q\,|\overline{B}^0\rangle\right)\right] \\
&= f_+(t)\,|B^0\rangle + \frac{q}{p} f_-(t)\,|\overline{B}^0\rangle \,,
\end{aligned}
\tag{1.18}
$$

and similarly, we get $|\overline{B}^0(t)\rangle$ as:

$$
\begin{aligned}
|\overline{B}^0(t)\rangle &= \frac{1}{2q}(|B_1(t)\rangle - |B_2(t)\rangle) \\
&= \frac{1}{2q}\left(e^{-i\left(M_1 - \frac{i}{2}\Gamma_1\right)t} |B_1(0)\rangle - e^{-i\left(M - \frac{i}{2}\Gamma_2\right)t} |B_2(0)\rangle\right) \\
&= \frac{1}{2q}\left[e^{-i\left(M_1 - \frac{i}{2}\Gamma_1\right)t} \left(p\,|B^0\rangle + q\,|\overline{B}^0\rangle\right) - e^{-i\left(M_2 - \frac{i}{2}\Gamma_2\right)t} \left(p\,|B^0\rangle - q\,|\overline{B}^0\rangle\right)\right] \\
&= \frac{p}{q} f_-(t)\,|B^0\rangle + f_+(t)\,|\overline{B}^0\rangle \,,
\end{aligned}
\tag{1.19}
$$

where

$$
\begin{aligned}
f_\pm(t) &= \frac{1}{2}\left(e^{-i\left(M_1 - \frac{i}{2}\Gamma_1\right)t} \pm e^{-i\left(M_2 - \frac{i}{2}\Gamma_2\right)t}\right) \\
&= \frac{e^{-i\left(M_1 - \frac{i}{2}\Gamma_1\right)t}}{2}\left(1 \pm e^{-i\left(\Delta m - \frac{i}{2}\Delta\Gamma\right)t}\right) ,
\end{aligned}
\tag{1.20}
$$

with $\Delta m \equiv M_2 - M_1$, and $\Delta\Gamma \equiv \Gamma_2 - \Gamma_1$, which indicates the difference of mass and decay width between the two mass eigenstates, respectively.

1.1.3 *CP* violation in the decay of neutral *B*-mesons into *CP* eigenstates

Here, we consider the decay of B-mesons into a CP eigenstate f with the CP eigenvalue of η_f from a pure B^0 or $\overline{B}^0$ state as an initial condition. Let us define the decay amplitudes from each initial state as follows:

$$
\begin{aligned}
B^0 \to f \quad &: \mathcal{A} \quad \equiv \langle f | \, \mathcal{H} \, | B^0 \rangle \,, && (1.21)\\
\overline{B}^0 \to f \quad &: \overline{\mathcal{A}} \quad \equiv \langle f | \, \mathcal{H} \, | \overline{B}^0 \rangle \,, && (1.22)\\
&\lambda_f \quad \equiv \frac{q}{p}\frac{\overline{\mathcal{A}}}{\mathcal{A}}, && (1.23)
\end{aligned}
$$

where $\mathcal{H}$ is a Hamiltonian for the transition. In the neutral B-meson system, $\Delta\Gamma/\Gamma \sim \mathcal{O}(10^{-3})$ thanks to the heavy mass of B-mesons. Thus, we can approximate Eq. 1.20 as $\Delta\Gamma \approx 0$ and $\Gamma_1 \approx \Gamma_2 \approx \Gamma$. Then, we can calculate the rate of decay $B^0 \to f$ at time t as:

$$
\begin{aligned}
\Gamma(B^0(t) \to f) &\propto |\, \langle f | \, \mathcal{H} \, | B^0(t) \rangle \,|^2 \\
&= |\mathcal{A}|^2 \left(|f_+(t)|^2 + |\lambda_f|^2 |f_-(t)|^2 + \lambda_f f_-(t) f_+^*(t) + \lambda_f^* f_-^*(t) f_+(t) \right) \\
&= |\mathcal{A}|^2 \frac{e^{-\Gamma t}}{2} ((1 + \cos(\Delta m t)) + \\
&\qquad |\lambda_f|^2 (1 - \cos(\Delta m t)) - 2 \sin(\Delta m t) \Im(\lambda_f)),
\end{aligned} \tag{1.24}
$$

and the decay rate of $\overline{B}^0 \to f$ can be calculated in the similar fashion:

$$
\begin{aligned}
\Gamma(\overline{B}^0(t) \to f) &\propto |\, \langle f | \, \mathcal{H} \, | \overline{B}^0(t) \rangle \,|^2 \\
&= |\mathcal{A}|^2 \left| \frac{p}{q} \right|^2 \left(|f_-(t)|^2 + |\lambda_f|^2 |f_+(t)|^2 + \lambda_f f_+(t) f_-^*(t) + \lambda_f^* f_+^*(t) f_-(t) \right) \\
&= |\mathcal{A}|^2 \left| \frac{p}{q} \right|^2 \frac{e^{-\Gamma t}}{2} ((1 - \cos(\Delta m t)) + \\
&\qquad |\lambda_f|^2 (1 + \cos(\Delta m t)) + 2 \sin(\Delta m t) \Im(\lambda_f)).
\end{aligned} \tag{1.25}
$$

In the neutral B-meson system, we can approximate as $|q/p| \approx 1$, then the above equations become

$$
\begin{aligned}
\Gamma(B^0(t) \to f) &\propto e^{-\Gamma t} |\mathcal{A}|^2 \left(1 - \frac{|\lambda_f|^2 - 1}{|\lambda_f|^2 + 1} \cos(\Delta m t) - \frac{2\Im(\lambda_f)}{|\lambda_f|^2 + 1} \sin(\Delta m t) \right), && (1.26)\\
\Gamma(\overline{B}^0(t) \to f) &\propto e^{-\Gamma t} |\mathcal{A}|^2 \left(1 + \frac{|\lambda_f|^2 - 1}{|\lambda_f|^2 + 1} \cos(\Delta m t) + \frac{2\Im(\lambda_f)}{|\lambda_f|^2 + 1} \sin(\Delta m t) \right). && (1.27)
\end{aligned}
$$

From the above two equations, we can calculate the time-dependent asymmetry $\mathcal{A}sym(t)$ of the decay width between $B^0 \to f$ and $\overline{B}^0 \to f$ as follows:

$$
\begin{aligned}
\mathcal{A}sym(t) &\equiv \frac{\Gamma(B^0(t) \to f) - \Gamma(\overline{B}^0(t) \to f)}{\Gamma(B^0(t) \to f) + \Gamma(\overline{B}^0(t) \to f)} \\
&= \frac{(|\lambda_f|^2 - 1)\cos(\Delta m t) + 2\Im(\lambda_f)\sin(\Delta m t)}{|\lambda_f|^2 + 1} \\
&\equiv A_{CP} \cos(\Delta m t) + S_{CP} \sin(\Delta m t)
\end{aligned} \tag{1.28}
$$

where A_{CP} and S_{CP} are defined as

$$A_{CP} \equiv \frac{|\lambda_f|^2 - 1}{|\lambda_f|^2 + 1}, \text{ and}$$
$$S_{CP} \equiv \frac{2\Im(\lambda_f)}{|\lambda_f|^2 + 1}. \tag{1.29}$$

According to Eq. 1.28, A_{CP} and S_{CP} are restricted within a circle with the unit radius[1]. A_{CP} is called the direct CP violation that arises if $|\mathcal{A}| \neq |\overline{\mathcal{A}}|$, which means the decay amplitudes from B^0 and $\overline{B}^0$ are different. S_{CP} is called the mixing-induced CP violation that occurs if λ_f has an imaginary phase, which means there is a phase difference in the amplitude between two decay path of $B^0 \to f$ and $\overline{B}^0 \to f$.

1.1.4 *CP* violation in the decay of $B^0 \to \eta' K^0_S$

In this subsection, we discuss a specific situation of CP violation in the decay of $B^0 \to \eta' K^0_S$. Among discussions in the previous subsections regarding CP violation in the mixing of the neutral B-mesons, the only part that depends on the decay mode is λ_f given in Eq. 1.23. Let us expand this equation to a concrete form for our target decay. A Feynman diagram for the decay is shown in Fig. 1.3.

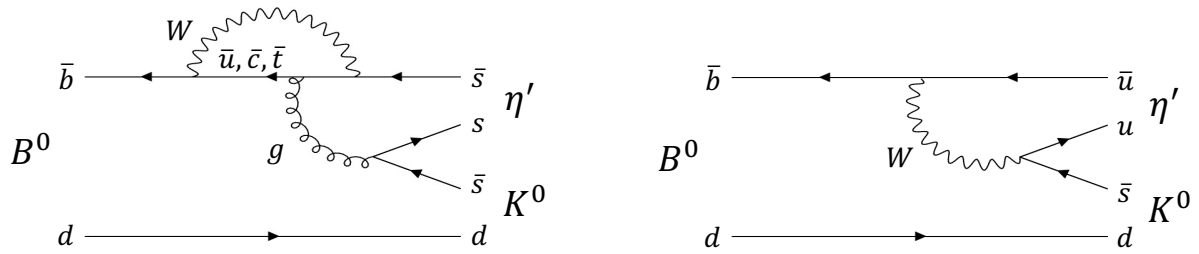


Figure 1.3: Two of the possible Feynman diagrams that represents $B^0 \to \eta' K^0_S$ decays. The diagram on the left shows a penguin diagram and the right one shows a tree-level diagram which is doubly Cabibbo suppressed.

First, we will calculate the q/p value. Among three box diagrams that represent $B^0 - \overline{B}^0$ mixing (with u, c, or t as an intermediate quark in the diagram), one with the top quarks shown in Fig. 1.2 has a dominant contribution in the mixing amplitude because of the heavy mass of top quarks. Also, the top quark cannot participate in the process of an on-shell intermediate state since its mass is larger than b quark. Thus we can represent $M_{12} \propto (V_{tb}V^*_{td})^2 m_t^2$ and $\Gamma_{12} \propto (V_{tb}V^*_{td})^2 m_b^2$ where m_t is the mass of top quark and m_b is the mass of bottom quark as stated in literature [3, 8]. Therefore, we can approximate Eq. 1.13 as follows:

[1] $\mathcal{Asym}(x)$ can be rewritten as $\mathcal{Asym}(x) = \sqrt{{A_{CP}}^2 + {S_{CP}}^2}(\sin(\alpha)\cos(x) + \cos(\alpha)\sin(x)) = \sqrt{{A_{CP}}^2 + {S_{CP}}^2}\sin(x+\alpha)$ with $\sin(\alpha) = \frac{A_{CP}}{\sqrt{{A_{CP}}^2+{S_{CP}}^2}}$ and $\cos(\alpha) = \frac{S_{CP}}{\sqrt{{A_{CP}}^2+{S_{CP}}^2}}$. If $|\mathcal{Asym}(x)| \leq 1$, then $\sqrt{{A_{CP}}^2 + {S_{CP}}^2} \leq 1$ holds.

$$\frac{q}{p} \approx \sqrt{\frac{M_{12}^*}{M_{12}}} + \mathcal{O}\left(\frac{\Gamma_{12}}{M_{12}}\right) \approx \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}. \tag{1.30}$$

Next, let us discuss the calculation of $\mathcal{A}$ in Eqs. 1.21 and 1.22 that express the amplitude of the transitions of $b \to sq\bar{q}$. The transitions can be proceeded by the penguin or tree diagram shown in Fig. 1.3, but the tree diagram is doubly Cabibbo suppressed in our case, so here we consider only the penguin diagram. We can factorize the amplitude explicitly only with CKM factors as follows:

$$\begin{aligned} \mathcal{A} &= V_{tb}^* V_{ts} P_t + V_{cb}^* V_{cs} P_c + V_{ub}^* V_{us} P_u \qquad (1.31) \\ &= V_{cb}^* V_{cs}(P_c - P_t) + V_{ub}^* V_{us}(P_u - P_t) \\ \overline{\mathcal{A}} &= \eta_{CP}(V_{tb} V_{ts}^* P_t + V_{cb} V_{cs}^* P_c + V_{ub} V_{us}^* P_u) \qquad (1.32) \end{aligned}$$

where $\eta_{CP} = -1$ is the CP-eigenvalue of $\eta' K_S^0$ final state, $P_{u,c,t}$ is the parameters that are not related directly to CKM factors, and the unitarity condition given in Eq. 1.3 is used. The difference of $P_i (i = u, c, t)$ is the type of propagator quark; the order of magnitude of terms $(P_c - P_t)$ and $(P_u - P_t)$ will be the same. On top of that, according to Eq. 1.2, $V_{cb}^* V_{cs}$ is order of $\lambda^2 \approx (0.22)^2$ and $V_{ub}^* V_{us}$ is order of $\lambda^4 \approx (0.22)^4$, so that the second term is negligible. Hence, we can approximate $\overline{\mathcal{A}}/\mathcal{A}$ to $(V_{cb} V_{cs}^*)/(-V_{cb}^* V_{cs})$.

In the final step, we should include a factor from the $K^0 - \overline{K}^0$ mixing as we consider the final state only for K_S^0, which is given by $V_{cs} V_{cd}^* / V_{cs}^* V_{cd}$ [9]. Finally, we can get λ_f by multiplying the above three calculation results as follows:

$$\lambda_f \approx \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \cdot \left(-\frac{V_{cb} V_{cs}^*}{V_{cb}^* V_{cs}}\right) \cdot \frac{V_{cs} V_{cd}^*}{V_{cs}^* V_{cd}} = -\frac{V_{tb}^* V_{td}}{V_{cb}^* V_{cd}} \frac{V_{cb} V_{cd}^*}{V_{tb} V_{td}^*} \tag{1.33}$$

With the graphical interpretation of the unitarity condition shown in Fig. 1.1, we can get a relation between λ_f and the first inner angle ϕ_1 of the unitarity triangle as $\lambda_f \approx -e^{-2i\phi_1}$. Therefore, we can connect A_{CP} and S_{CP} given in Eq. 1.29 with the actual situation in the Standard Model for our target decay as follows:

$$A_{CP} \approx 0 \quad \text{and} \quad S_{CP} \approx \sin(2\phi_1). \tag{1.34}$$

1.2 Effects of New Physics in the decay of $B^0 \to \eta' K_S^0$

In the aspect of exploring New Physics with CP violation in $b \to q\bar{q}s$ decay, the $B^0 \to \eta' K_S^0$ decay is prominent thanks to the following features:

1. The branching ratio is relatively higher than other $b \to q\bar{q}s$ process. (Table 1.1)
2. Low theoretical uncertainty for SM-only prediction because of the low contamination of the tree diagram. (Table 1.2[10])

3. New Physics can contribute to the decay amplitude through the penguin diagram due to the loop of W bosons in the diagram. (Fig. 1.3)

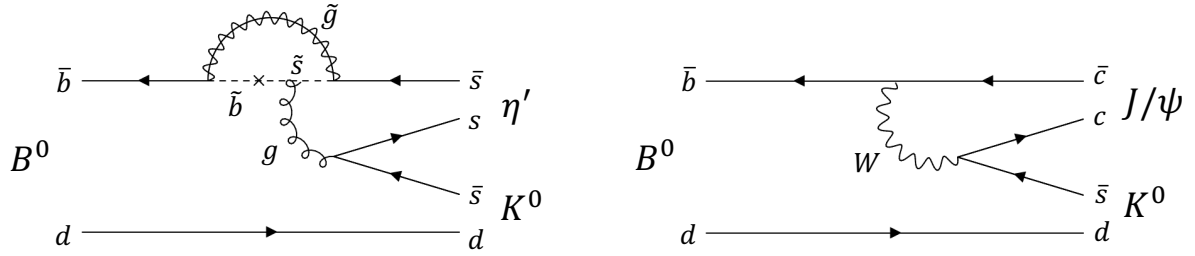


Figure 1.4: An example Feynman diagram with New Physics of SUSY for $B^0 \to \eta' K_S^0$ (left) and example Feynman diagram with charmonium transition of $B^0 \to J/\psi\, K_S^0$ (right).

Decay mode from B^0	Measured branching fraction[11]
$\eta' K^0$	$(6.6 \pm 0.4) \times 10^{-5}$
$\pi^0 K^0$	$(9.9 \pm 0.5) \times 10^{-6}$
ϕK^0	$(7.3 \pm 0.7) \times 10^{-6}$
ωK^0	$(4.8 \pm 0.4) \times 10^{-6}$
$\rho^0 K^0$	$(3.4 \pm 1.1) \times 10^{-6}$
ηK^0	$(1.23^{+0.27}_{-0.24}) \times 10^{-6}$

Table 1.1: Measured branching fraction of decay channels to the charmless two-body final states

Especially, the third item is important in the aspect of exploring New Physics because a change of amplitude by New Physics enables us to determine the existence of New Physics through *CP* asymmetry measurement. For example, SUSY can contribute to the penguin diagram as shown on the left of Fig. 1.4 and create additional contributions that can change λ_f. In the Feynman diagram with SUSY, a loop with the W bosons and the internal quarks is replaced with a loop of gluino and internal squarks of which the flavor transits along the process, which is a supersymmetric counterpart of gluons and quarks, respectively. The transition for the flavor of squarks in SUSY can introduce a complex phase[14], and it can affect measured *CP* asymmetries. If any New Physics that changes the amplitude exists, the amplitude in Equations 1.21 and 1.22 will change. We used the parameterization in literature [15] to get the model-independent relation of A_{CP} and S_{CP} between the Standard Model and New Physics. The decay amplitude of $\mathcal{A}$ and $\overline{\mathcal{A}}$ changes as

$$\begin{aligned} \mathcal{A}' &= |\mathcal{A}_{SM}| e^{i\phi_{SM}} e^{i\delta_{SM}} + |\mathcal{A}_{NP}| e^{i\phi_{NP}} e^{i\delta_{NP}}, \text{ and} \\ \overline{\mathcal{A}}' &= \eta_{CP}(|\mathcal{A}_{SM}| e^{-i\phi_{SM}} e^{i\delta_{SM}} + |\mathcal{A}_{NP}| e^{-i\phi_{NP}} e^{i\delta_{NP}}) \end{aligned} \tag{1.35}$$

Decay mode from B^0	Expected deviation [12]	Data [13]
$\eta' K^0_S$	$0.01^{+0.01}_{-0.01}$	$-0.05 \pm +0.06$
$\pi^0 K^0_S$	$0.07^{+0.05}_{-0.04}$	$-0.11^{+0.17}_{-0.17}$
ϕK^0_S	$0.02^{+0.01}_{-0.01}$	$0.06^{+0.11}_{-0.13}$
ωK^0_S	$0.13^{+0.08}_{-0.08}$	$0.03^{+0.21}_{-0.21}$
$\rho^0 K^0_S$	$-0.08^{+0.08}_{-0.12}$	$-0.14^{+0.18}_{-0.21}$
ηK^0_S	$0.10^{+0.11}_{-0.07}$	–

Table 1.2: Theoretical expectation value within the Standard Model. The prediction is given as a deviation of S_{CP} in each decay mode having charmless two-body final states from the value in the charmonium mode ($B^0 \to c\bar{c}K^0$). The averaged values of the experimental measurement are also given.

where $|\mathcal{A}_{NP(SM)}|$ is the magnitude of an amplitude, $\phi_{NP(SM)}$ is a weak phase which violates *CP* symmetry, $\delta_{NP(SM)}$ is a strong phase which conserves *CP* symmetry from New Physics (or from the Standard Model). Then, we can set as $\phi_{SM} = 0$ and use an approximation of $q/p \approx (V^*_{tb}V_{td})/(V_{tb}V^*_{td}) = e^{-2i\phi_1}$ because V_{cx} $(x = d, s, b)$ in Equation 1.33 does not have any complex phase as stated in Equation 1.2. Using these items, the modified λ'_f is obtained as the following equation:

$$\begin{aligned} \lambda'_f = \frac{q}{p}\frac{\overline{\mathcal{A}}'}{\mathcal{A}'} &= -e^{-2i\phi_1}\frac{|\mathcal{A}_{SM}|e^{i\delta_{SM}} + |\mathcal{A}_{NP}|e^{-i\phi_{NP}}e^{i\delta_{NP}}}{|\mathcal{A}_{SM}|e^{i\delta_{SM}} + |\mathcal{A}_{NP}|e^{i\phi_{NP}}e^{i\delta_{NP}}} \\ &= -e^{-2i\phi_1}\frac{1 + re^{-i\phi_{NP}}e^{i\delta}}{1 + re^{i\phi_{NP}}e^{i\delta}} \end{aligned} \tag{1.36}$$

where $r = |\mathcal{A}_{NP}|/|\mathcal{A}_{SM}|$ and $\delta = \delta_{NP} - \delta_{SM}$. The modified A'_{CP} and S'_{CP} due to the insertion of New Physics is calculated as follows:

$$\begin{aligned} A'_{CP} &= \frac{|\lambda'_f|^2 - 1}{|\lambda'_f|^2 + 1} = \frac{4r\sin(\delta)\sin(\phi_{NP})}{1 + r^2 + 2r\cos(\delta + \phi_{NP})} \cdot \frac{1 + r^2 + 2r\cos(\delta + \phi_{NP})}{2(1 + r^2) + 4r\cos(\delta)\cos(\phi_{NP})} \\ &= \frac{2r\sin(\delta)\sin(\phi_{NP})}{1 + r^2 + 2r\cos(\delta)\cos(\phi_{NP})} \end{aligned} \tag{1.37}$$

$$\begin{aligned} S'_{CP} &= \frac{2\Im(\lambda'_f)}{|\lambda'_f|^2 + 1} = \frac{-2\Im(e^{-2i\phi_1}(1 + r^2e^{-2i\phi_{NP}} + 2r\cos(\delta)e^{-i\phi_{NP}}))}{2(1 + r^2) + 4r\cos(\delta)\cos(\phi_{NP})} \\ &= \frac{\sin(2\phi_1) + r^2\sin(2\phi_1 + 2\phi_{NP}) + 2r\cos(\delta)\sin(2\phi_1 + \phi_{NP})}{1 + r^2 + 2r\cos(\delta)\cos(\phi_{NP})} \end{aligned} \tag{1.38}$$

According to the above equations for A'_{CP} and S'_{CP}, we can constrain the three parameters of r, δ, and ϕ_{NP} by measuring the *CP* parameters and examining the difference of measured values from the SM expectation of $\sin(2\phi_1)$.

Since *CP* asymmetries in the tree-level charmonium decay are not susceptible to New Physics, we can take *CP* asymmetries in the channel as a reference value for the SM

expectation. For example, the decay of $B^0 \to J/\psi K_S^0$ of which a Feynman diagram is shown on the right in Fig. 1.4 has small theoretical uncertainty. Furthermore, because the decay is a tree process, the channel has a large branching ratio so that small statistical uncertainty also can be achieved. Therefore, if there is any difference in *CP* asymmetries between charmonium channels like $B^0 \to J/\psi K_S^0$ and $B^0 \to \eta' K_S^0$, it will be solid evidence of the existence of New Physics.

1.3 The current status of *CP* violation measurements

Belle and BaBar reported the *CP* violation measurements in our target decay channel using their dataset corresponding to the integrated luminosity of $711\,\mathrm{fb}^{-1}$ ($N(B\bar{B}) = 772\mathrm{M}$) and $426\,\mathrm{fb}^{-1}$ ($N(B\bar{B}) = 467\mathrm{M}$), respectively, as follows:

$$\left.\begin{aligned} A_{CP} &= 0.03 \pm 0.05(\mathrm{stat}) \pm 0.03(\mathrm{syst}) \\ S_{CP} &= 0.68 \pm 0.07(\mathrm{stat}) \pm 0.03(\mathrm{syst}) \end{aligned}\right\} = \mathrm{Belle}[16], \quad \left.\begin{aligned} A_{CP} &= 0.08 \pm 0.06(\mathrm{stat}) \pm 0.02(\mathrm{syst}) \\ S_{CP} &= 0.57 \pm 0.08(\mathrm{stat}) \pm 0.02(\mathrm{syst}) \end{aligned}\right\} = \mathrm{BaBar}[17].$$

The comprehensive results for A_{CP} and S_{CP} in the other channels are shown in Fig. 1.5. Especially, the internal note of Belle reports A_{CP} and S_{CP} only with our target subchannel, and reported asymmetries are $A_{CP} = -0.058 \pm 0.181$ and $S_{CP} = 0.800 \pm 0.259$[18].

For the *CP* asymmetries in $B^0 \to J/\psi K_S^0$, Belle, BaBar, and LHCb reported the value. The averaged A_{CP} and S_{CP} values are $A_{CP} = 0.000 \pm 0.020(\mathrm{stat})$ and $S_{CP} = 0.695 \pm 0.019(\mathrm{stat})$[19]. Also, averaged A_{CP} and S_{CP} values over all charmonium channels are $A_{CP} = 0.005 \pm 0.015(\mathrm{stat})$ and $S_{CP} = 0.699 \pm 0.017(\mathrm{stat})$[19]. We will use values from the charmonium channels as a reference value for the SM expectation on *CP* asymmetries.

1.4 Overview and goal of this dissertation

In this dissertation, we are going to elaborate on the detailed methodology and result for *CP* violation measurement in the decay channel of $B^0 \to \eta' K_S^0$ with a subchannel of $\eta' \to \eta\pi^+\pi^-$ and $\eta \to \pi^+\pi^-\pi^0$ using the experimental data of the Belle II experiment until June 2022, which corresponds to the amount of $362\,\mathrm{fb}^{-1}$. The $B^0 \to \eta' K_S^0$ decay is a key mode in exploring New Physics using *CP*-violation since the decay proceeds via the penguin diagram and has a large branching ratio. The subchannel that we aim for is hard to fully reconstruct due to it requiring six charged pions and two gammas for the reconstruction of π^0. Using the subchannel, we develop a universal methodology to measure *CP* asymmetries in all the subchannels of $B^0 \to \eta' K_S^0$, including the method for the validation. Also, we discuss the prospect of *CP*-violation measurements in the channel of $B^0 \to \eta' K_S^0$ with the methodology developed in this study at the target luminosity of Belle II, $50\,\mathrm{ab}^{-1}$. Therefore, the main goal of this dissertation is to establish

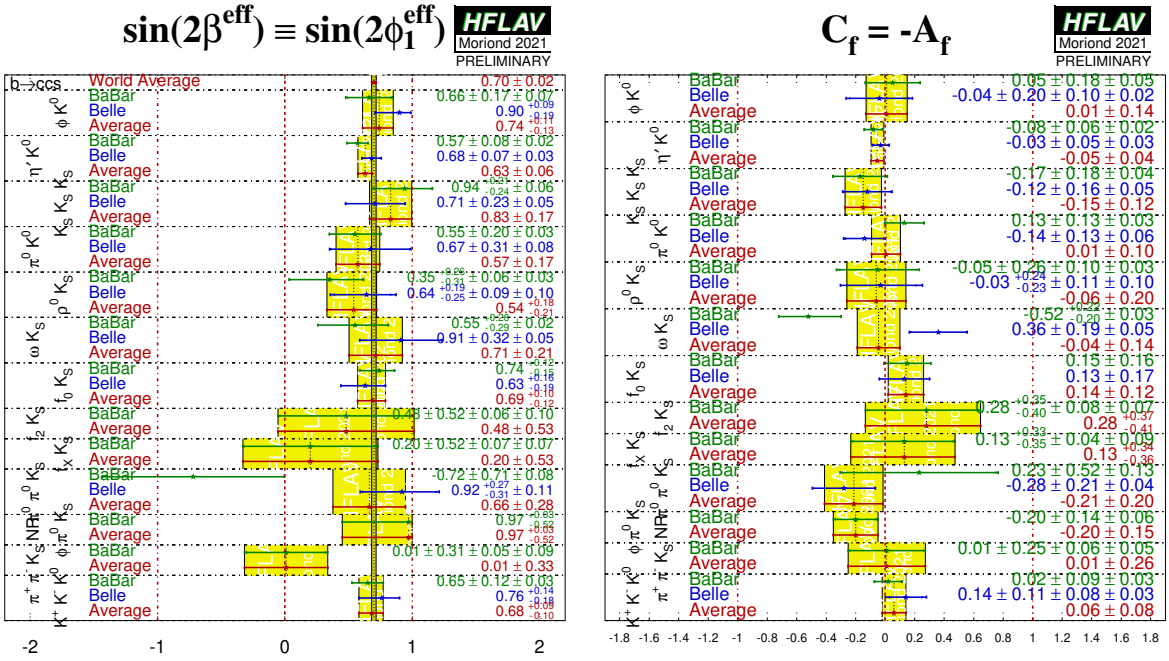


Figure 1.5: Average of *CP* asymmetries for ϕ_1 from the various decays. The left plot shows $S_{CP} = \sin(2\beta^{\text{eff}})$ results, and the right one shows $A_{CP} = A_f$ results[19].

a methodology to measure *CP* asymmetries in Belle II that can be applied to the channel of $B^0 \to \eta' K^0_S$ and discuss the validation and prospect of the *CP* measurement in the channel with the methodology using the experimental data.

After this chapter, we will first discuss the experimental apparatus to collect the experimental data in Ch. 2, then explain the procedure to extract parameters for the *CP* asymmetry. In Ch. 3, we will introduce an overview of the methodology for the estimation in the experimental environment using the theory discussed in this chapter. After that, we will explain how to reconstruct our target decay for the analysis in Ch. 4. Then, in the following chapter of Ch. 5, we will prepare ingredients for a fitting procedure to determine A_{CP} and S_{CP} from the reconstruction results. In the next chapter of Ch. 6, we build the fitting procedure using the ingredients and validate it using various methods, and finally, we describe the *CP* violation parameters with systematic uncertainties from the Belle II experimental data. The last two chapters of Ch. 7 discuss the result and prospect of this analysis with the experimental data from Belle II in the future and the conclusion of this dissertation. To sum up, we will discuss the whole procedure, including the reconstruction, signal extraction, resolution function, extraction of *CP* parameters, and the systematic uncertainty that is capable of measurement even with high statistics of $50\,\text{ab}^{-1}$ in the future in this dissertation.

Chapter 2

Experimental Apparatus

In this section, we will explain the experimental apparatus for the Belle II experiment. Belle II is an experiment for particle physics that aims to discover New Physics through the flavor physics in B-mesons. Since one of the main concepts of the experiment is focused on B-mesons, the experiment is often called B-factories. The experiment is at High Energy Accelerator Research Organization (KEK) in Tsukuba-shi, Japan. It covers a vast solid angle for the interaction point (IP) using its subdetectors to record physical phenomena at the IP arising from collisions of asymmetrically accelerated electrons and positrons from SuperKEKB. SuperKEKB accelerator, the successor of KEKB accelerator, is tuned to produce plenty of B-mesons and has the highest luminosity in the world.

2.1 SuperKEKB

SuperKEKB is an electron-positron collider that accelerates electrons to 7 GeV and positrons to 4 GeV. It comprises six components as shown in Fig. 2.1[20]: an electron source, linear accelerator, positron generator, damping ring, and two main rings for electrons and positrons. The main ring for electrons is called High Energy Ring (HER), where electrons move in the clockwise direction, and that for positrons is called Low Energy Ring (LER), where positrons move in the opposite direction to the electrons. Two beams collide at the interaction point where is the center of the Belle II detector, and the center of mass energy of the collision is tuned 10.58 GeV, which is a resonance energy of $\Upsilon(4S)$ that decays mainly to a pair of B-mesons.

SuperKEKB evolved in many aspects from KEKB to achieve 30 times higher luminosity than its predecessor. For example, the accelerator adopts the ”nano-beam scheme“, which squeezes the vertical beta function into 0.3 mm at the IP, which is 20 times smaller than that of KEKB. Two beams cross in SuperKEKB with a large angle of 83 mrad, which helps to avoid the degrading of luminosity (hourglass effect) and allows to place of the focusing magnet to be closer to the IP. On top of that, the beam current for both rings will be doubled, increasing the luminosity further[20].

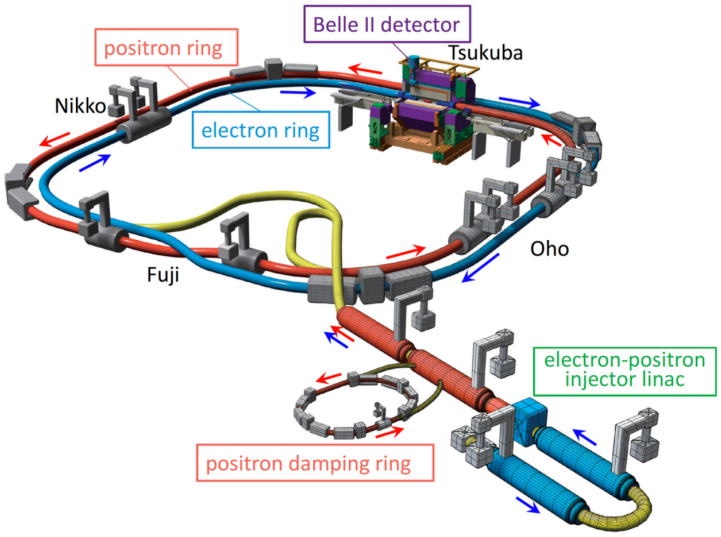


Figure 2.1: Configuration of SuperKEKB with Belle II.

On the other hand, there are challenges in *CP* violation measurement with improved SuperKEKB. The nano-beam scheme requires a narrow dynamic aperture around the IP, and it makes the shorter lifetimes by the Touschek effects. To mitigate the problems, the beam energy of LER is increased, and the asymmetry of beam energies became smaller than from KEKB[21]. Due to the smaller asymmetry, the boost factor ($\beta\gamma$) in the center of the mass system is reduced to 0.287, which is a 2/3 level of KEKB. This might negatively affect the resolution for the decay time determination of B-mesons. Also, more intensive beam backgrounds due to the highly focused beam and increased beam current are another challenge that Belle II should deal with for measurements of *CP* asymmetries. The following section will explain how Belle II copes with these challenges.

2.2 The Belle II detector

The Belle II detector is a general-purpose detector that consists of data acquisition systems (DAQ), triggers, superconducting solenoid magnets, and the following subdetectors:

- Vertex detector: This subdetector consists of PXD (Pixed Detector) and SVD (Silicon Vertex Detector), located closest to the interaction point, precisely measures the positions of charged particles.
- Central Drift Chamber (CDC): This wire chamber measures the trajectory of charged particles and provides information on their momenta and energy losses in the detector volume.
- Electromagnetic Calorimeter (ECL): This calorimeter is designed to measure the energy of electrons and photons and help identify these particles.
- Time-of-Propagation counter (TOP): This counter provides particle identification of charged K and π in the barrel part of the Belle II detector.

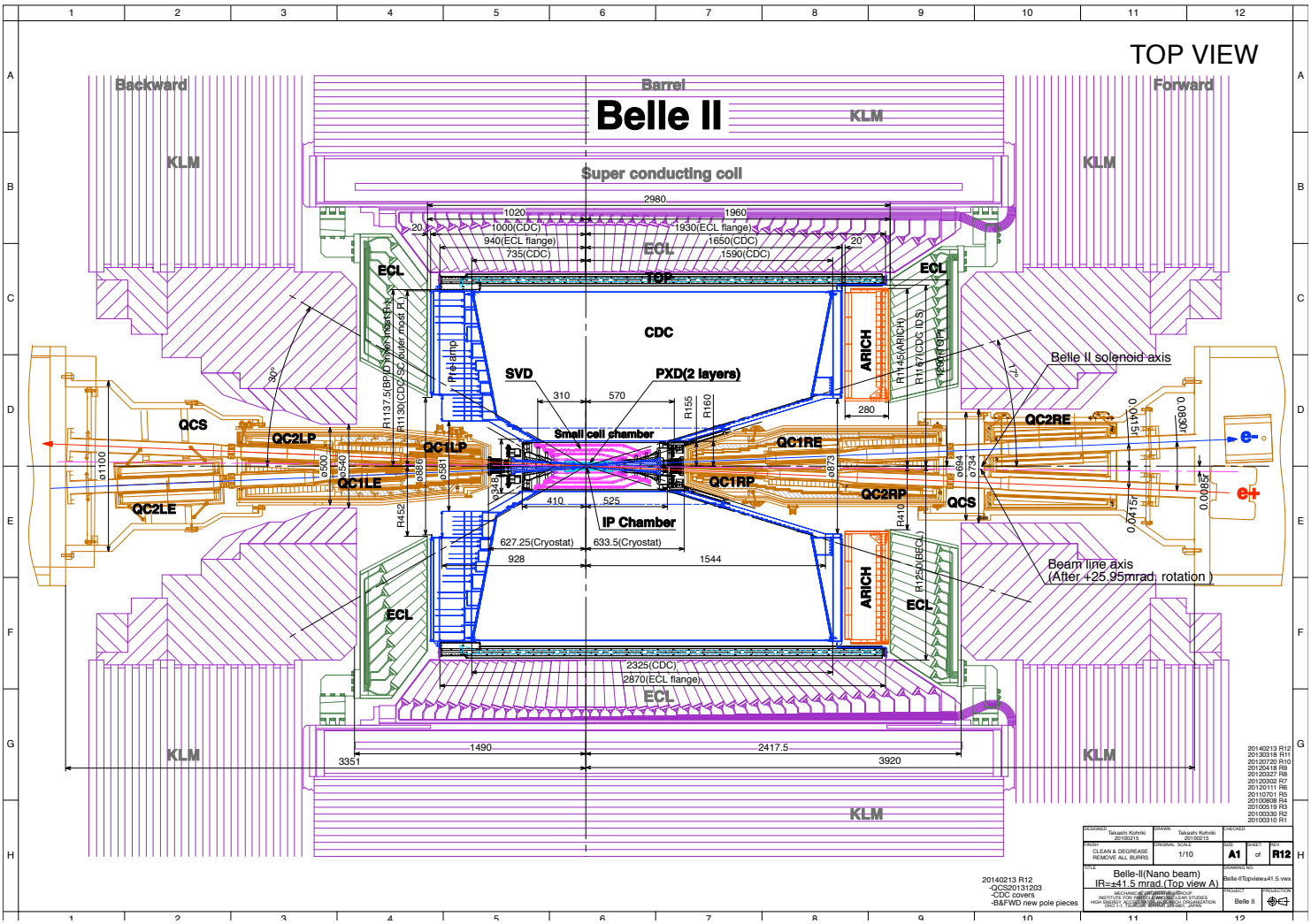


Figure 2.2: Top view of the Belle II detector.

- Aerogel Ring Image Cherenkov detector (ARICH): This ring image counter allows identification of charged K and π at the forward end-cap part of the Belle II detector.
- K_L^0 and μ detector (KLM): This subdetector is the outermost part of the Belle II detector and is dedicated to identifying K_L^0 and μ that can reach the subdetector.

Figure 2.2[22] shows a top-view scheme of the Belle II detector. Those subdetectors will be reviewed in the following subsections, and here we will explain the peripheral components of the Belle II detector rather than the subdetectors.

A superconducting solenoid magnet provides a strong magnetic field of which intensity is about 1.5 T along the median of the HER and LER beam line for the Belle II detector, which is extended to the inside of KLM. This magnetic field bends the trajectory of charged particles, allowing for the determination of the momenta of the particles. The detector is in the first period of long-shutdown (LS1) to upgrade its subdetectors from July 2022. Belle II and SuperKEKB will resume their operation at the end of 2023.

Coordinate system of the Belle II detector We set the Cartesian coordinate system (x, y, z) for the laboratory frame in our analysis. The unit vector $\hat{x}$ indicates the radial outside direction of the main ring, $\hat{y}$ is equal to the vertically upward direction, and $\hat{z}$ lies on the median line of two beam axes of HER and LER and its direction can be determined from right-hand law, which is the same as the right direction in Fig. 2.2. The origin point of $(x, y, z) = (0, 0, 0)$ of the Belle II detector is the nominal point of the IP.

Parameterization for the trajectory of charged particles In the Belle II experiment, the trajectory of a charged particle is a helix due to the magnetic field of the Belle II detector. The trajectory is called "Track", and the helix of a track is described by five parameters: $(d_0, z_0, \phi_0, \omega, \tan(\lambda))$. We need a pivot which can be any point on the helix for the parameterization with those parameters. We define the pivot as the Point Of Closest Approach (POCA), which is the closest point of the helix from the origin point of the Belle II detector. The description of each parameter is shown in the following list, and Fig. 2.3 gives the graphical illustration of the parameters

- d_0: The signed distance of POCA to the z-axis in the x-y plane.
- z_0: z coordinate of POCA
- ϕ_0: The azimuthal angle of the transverse momentum of the track at POCA
- ω: Signed curvature of the helix in the x-y plane. The particle charge and direction of the magnetic field determine the sign.
- $\tan(\lambda)$: Tangent of the trajectory in ϕ-z plane. ϕ is a 2-D polar angle on the x-y plane.

2.2.1 Vertex Detector

Vertex detection of the Belle II detector is one of the most critical features in the time-dependent *CP* violation measurement, as the analysis requires precise vertex information to reconstruct the proper decay time difference of the neutral *B*-mesons. To measure the vertex precisely, the vertex detector of Belle II is installed at the innermost place and consists of two parts of semiconductor detectors: PXD(Pixel Detector) and SVD(Silicon Vertex Detector). In the nano-beam scheme, we have a smaller radius of the beam pipe and more background incidents in the innermost layer of the vertex detector, so we cannot realize a high-precision vertex detector only with strip detectors due to its high occupancy, which the Belle experiment adopted.

PXD is introduced to deal with these problems, which has a much larger number of silicon pixel channels, so it has a much lower occupancy. This semiconductor detector is located just around the beam pipe and can measure the 2D position where a charged particle passes through the detector with a spatial resolution of a few micrometers[21].

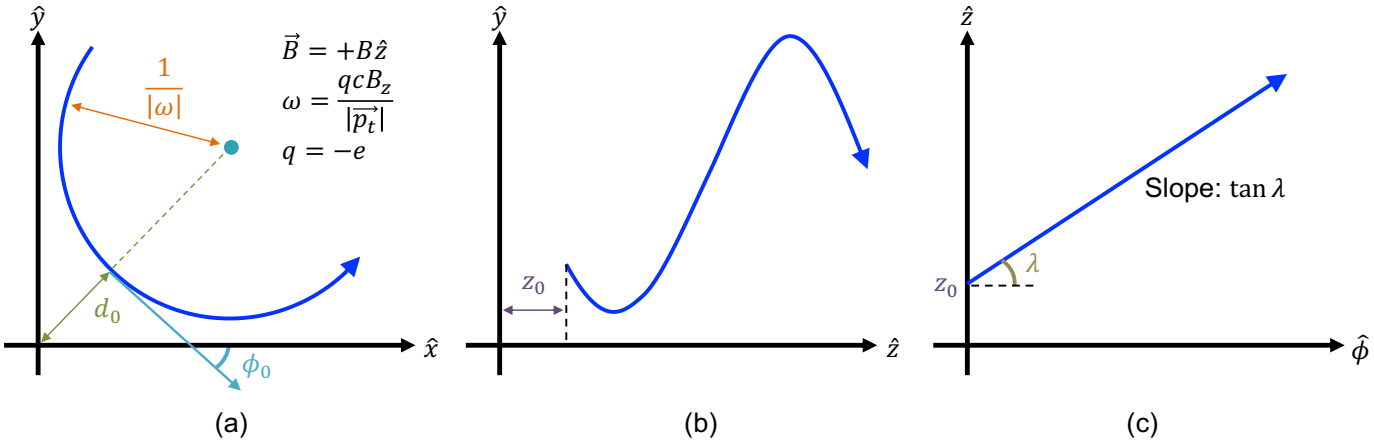


Figure 2.3: Definition of helix parameters in case of negatively charged particles. A blue solid line indicates the projected trajectory of the particle. (a) describes definition of d_0, ϕ_0, and ω, and (b) shows the definition of z_0. (c) illustrates definition of $\tan(\lambda)$. ϕ coordinate of (c) is the polar angle of the helix in the r-ϕ (x-y) plane.

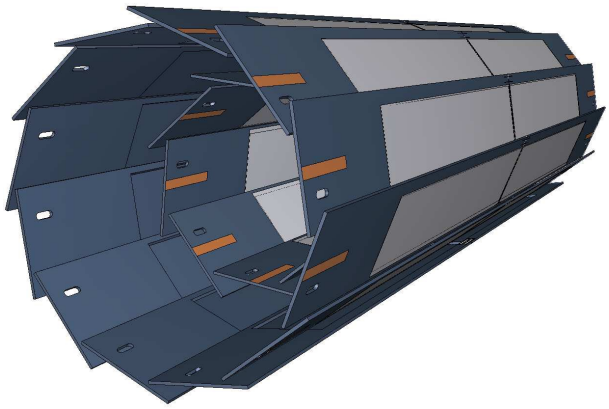
PXD has two layers; the innermost layer is fully-equipped, but the outer layer has only two module ladders. During the LS1, the entire PXD subdetector will be replaced with the fully functional second layer.

SVD is a Double-sided Silicon Strip Detector(DSSD) and is installed just outside of PXD with four layers, which has the spatial resolution of $\sim 20\,\mu$m for the charged particles that pass through the detector[23]. SVD provides not only the position of charged particles but also the region of interest for the PXD detector. PXD has many channels; we cannot gather signals from all channels of the detector due to the limited bandwidth. Therefore, the front-end electronics of PXD read out data only within the region of interest determined from an extrapolation of tracks from SVD onto PXD to reduce the amount of data to be processed.

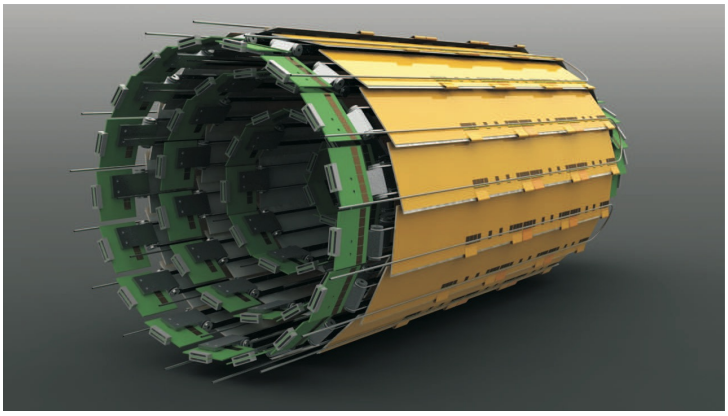
The vertex detector is placed in a windmill structure, as shown in Fig. 2.4d with a radius of 14 mm, 22 mm, 39 mm, 80 mm, 104 mm, and 135 mm for each layer. 3D modelings of PXD, SVD, and the entire VXD structure are given in Figs. 2.4a, 2.4b, and 2.4c, respectively. With this configuration, the vertex detector of the Belle II experiment can have the tracking resolution with the order of $\mathcal{O}(10\,\mu\text{m})$.

2.2.2 Central Drift Chamber

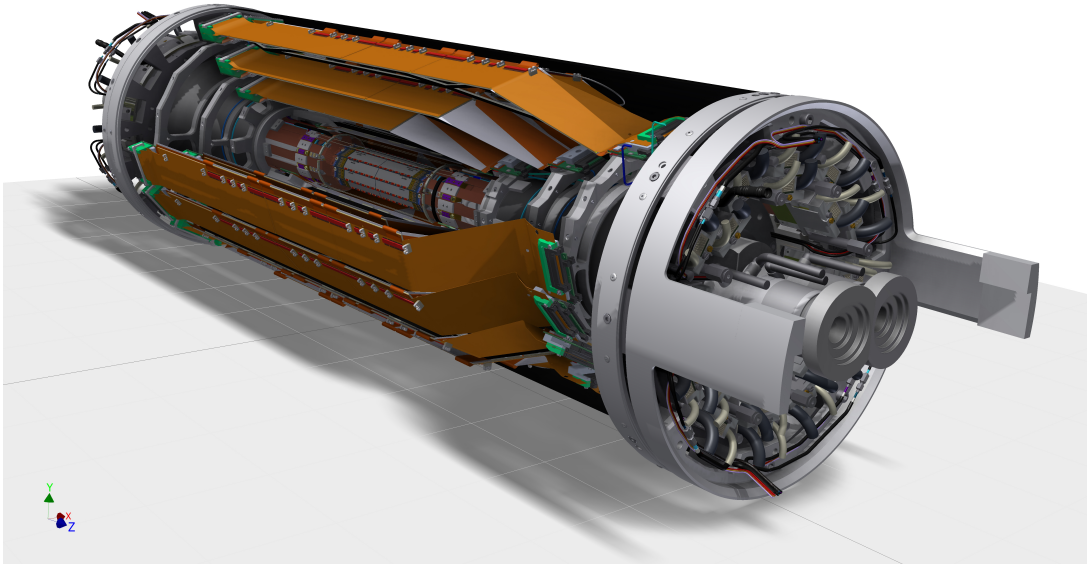
The Central Drift Chamber (CDC) is a drift chamber filled with a gas mixture of 50% helium and 50% ethane by their weight and is important for our analysis. It is one of the main parts for tracking the charged particle, such as the reconstruction of charged tracks, measuring their momenta with high precision, and providing trigger signals for events of B-mesons with the coverage of $17° < \theta < 150°$, θ is the polar angle along the z-axis.



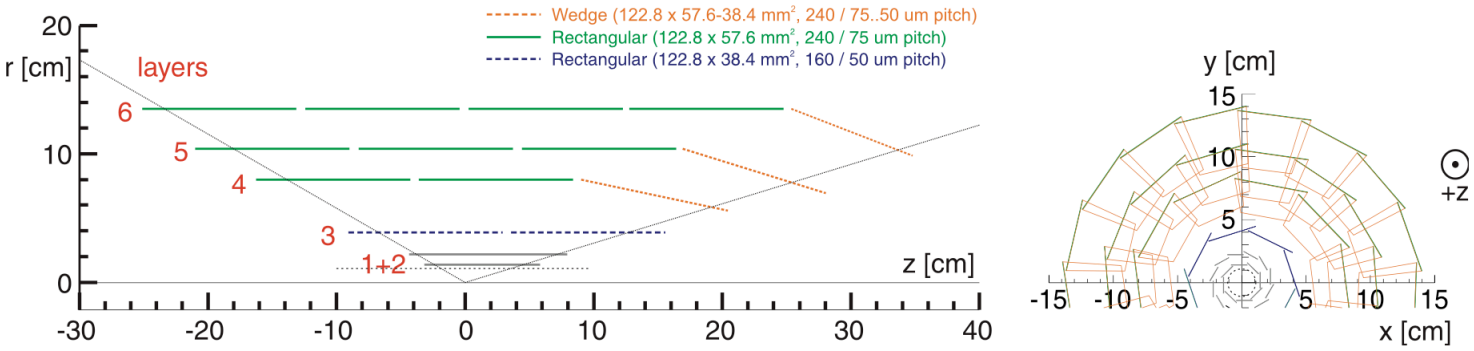
(a) Enlarged 3D modeling of 2-layer PXD[21]



(b) 3D modeling of 4-layer SVD[21]



(c) 3D modeling of the entire VXD with peripheral components, such as beam pipe and supports[24]. The coordinate system of the figure is unrelated to the lab frame of Belle II.



(d) The windmill stricture for VXD[25].

Figure 2.4: Geometrical configurations for the vertex detector. Note that the size of PXD in the 3D modeling is exaggerated.

On top of that, it measures the energy loss ($\mathrm{d}E/\mathrm{d}x$) within its gas mixture volume and provides useful information for particle identification(PID). Especially CDC can identify the charged tracks with low momentum, which cannot reach the additional subdetectors for PID located outside of the CDC. In those aspects, CDC is crucial for our analysis because the target sub-channel requires six charged tracks and their momenta for the reconstruction and good PID performance to reject background events.

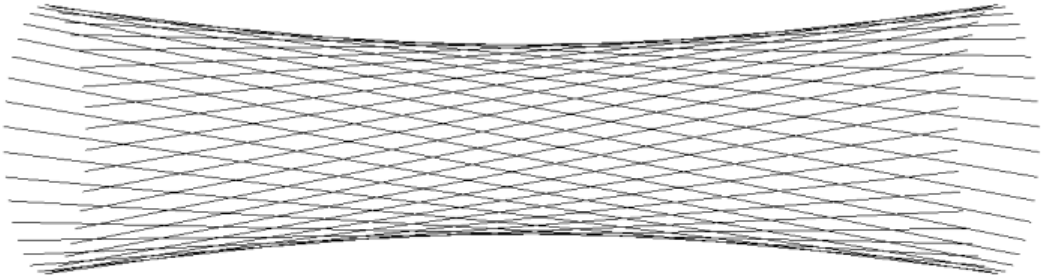
CDC consists of two types of wires: one is the field wire to make an electric field, and another one is the sensor wire that catches the signal from the charged particle. It has 14336 sense wires over 56 layers, and the layers are grouped into a super layer(SL); the first SL consists of 8 layers and otherwise contains 6 layers. Also, there are two types of SL: the axial and stereo layers. An axial layer is a group of wires that are parallel to the beamline (the z-axis), while a stereo layer is a group of wires twisted at a specific angle in the polar angle direction relative to the axis of the CDC cylinder. A graphical description of each type of the super layers is given in Fig. 2.5. The axial layers are sensitive to the distance r from the IP on the r-ϕ (x-y) plane. Stereo layers are introduced to measure the z coordinate from CDC. Thanks to the twisted wire configuration relative to the cylindrical axis of CDC, we can measure the z position of trajectory points of charged particles using the hit-time information from the stereo layers.

2.2.3 Electromagnetic Calorimeter

Electromagnetic Calorimeter(ECL) is an array of scintillation crystals of CsI(Tl), and Belle II reused the ECL detector of Belle. It detects photons with high efficiency and measures their energy and angular coordinates, and provides information to identify the electron by using $E/|\vec{p}|$ value, which is deposited energy on ECL over the measured momentum of a charged particle from tracking devices(VXD and CDC). Also, it generates a proper signal for the ECL trigger and is used to measure the on and off-line luminosity calculation using Bhabha scattering or pair annihilation of e^+e^- pair in the collision. The array in the barrel contains 6624 CsI(Tl) crystals of which the cross-section toward the IP is $6 \times 6\mathrm{cm}^2$, and the length is 30cm ($16.1X_0$). The end-cap part consists of 2112 crystals. The scintillation lights are digitized by photodiodes glued to each crystal, and relevant electronics that are renewed to deal with the pile-up events of CsI(Tl) crystals at the higher background environment.

Since our target channel contains π^0 in the final state and decays into a pair of photons, the performance of this detector is important for our analysis. According to the study from Belle, the invariant mass resolution for $\pi^0 \to \gamma\gamma$ is $4.75 \pm 0.08\,\mathrm{MeV}/c^2$ [27], which is adequate for our analysis.

(a) An axial wire layer - sense wires are parallel to the beamline



(b) A stereo wire layer - sense wires are skewed to the beamline (exaggerated)

Figure 2.5: Two types of the super layers in CDC[26].

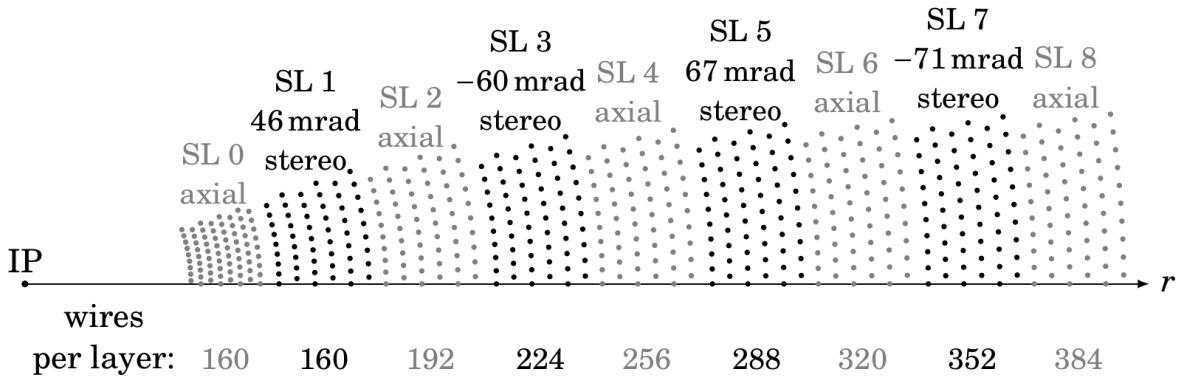


Figure 2.6: Wire configuration for CDC of Belle II [26].

2.2.4 Identification of π/K

Belle II has two subdetectors to distinguish the charged K and π, which one is TOP for the barrel and another one is ARICH for the end-cap part. These systems play an important role in the measurement of CP asymmetries with our target subchannel because it needs six charged pions to reconstruct the decay of B^0 meson through the subchannel, and for the successful reconstruction of the signal decay, all of them should be true pions and reject all of the fake pions. Also, the subdetectors aid in identifying the flavor of B-meson, which is essential information for our analysis.

ARICH covers the forward end-cap area and consists of the following four parts[21]:

- Aerogel radiator: In here, photons by Cherenkov radiation are produced when charged particles pass through with the velocity v of $v > c/n$ where c is the speed of light, and n is the relative index of refraction of the radiator,
- Photon detectors: An array of Hybrid Avalanche Photo-Detector (HAPD) that can detect single photons with high efficiency even in a strong magnetic field of the Belle II detector and measure the position of the photons with good resolution, and
- Expansion volume: Gap between the radiator and photon detectors to enable photons from the radiator to form rings on the surface of the photon detectors.
- Readout systems: An electronic system to digitize and encode signals from the photon detector and transfer the data to the DAQ system.

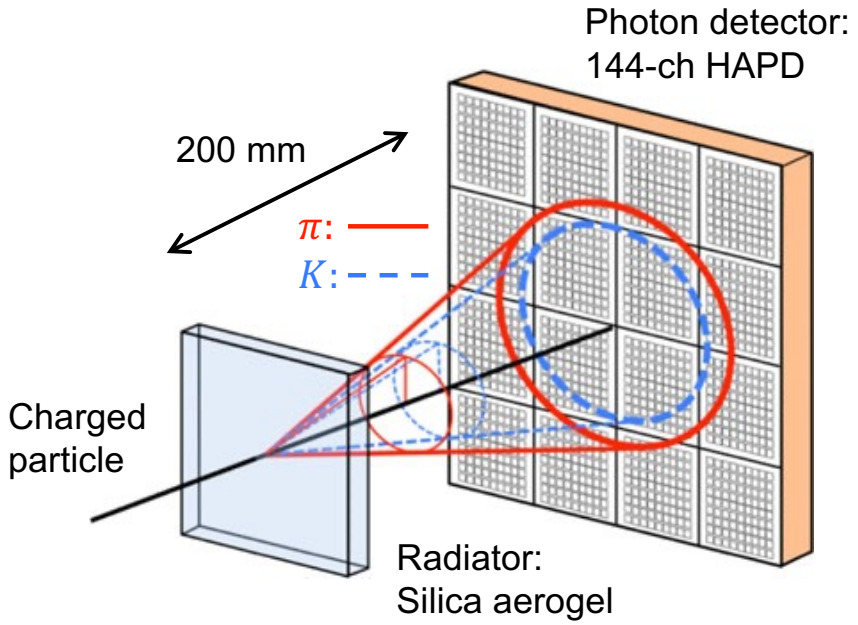


Figure 2.7: Geometry of ARICH and the principle of π/K identification in the ARICH counter. The solid and dotted lines with cones illustrate the emitted Cherenkov light from charged pions and kaons, respectively. Redrawn based on a figure in literature[28].

This subdetector reconstructs the ring image from the Cherenkov photons and determines the Cherenkov angle from the ring image. Since the angle depends on the mass of charged particles, we can calculate the mass of the incident particles with the angle and the momentum of the particle estimated in the tracking system of Belle II so that ARICH can identify the charged kaons and pions. This principle is illustrated in Fig. 2.7. According to literature[28], π and K identification efficiency was estimated at 97.4% and 4.9% at the particle momentum of 3.5 GeV/c, respectively. Study of literature[29] reports that π identification efficiency was estimated at $(93.5 \pm 0.6)\%$ and the fake rate for the charged pions is $(10.9 \pm 0.9)\%$ using the decay of the channel of $D^{*+} \to D^0(K^-\pi^+)\pi^+$ from the e^+e^- collision data.

TOP provides information to separate $\pi^\pm$ and $K^\pm$ in the barrel part of the Belle II detector. It consists of a bar-shaped radiator of quartz to generate Cherenkov lights, a mirror attached at the end of the radiator, and a Micro-Channel Plate Photomultiplier (MCP-PMT) with a prism as the light guide installed at the other end. This subdetector identifies the charged pions and kaons using Cherenkov lights. Figure 2.8 shows the basic geometry of TOP and provides a basic idea to identify the $\pi^\pm$ and $K^\pm$ using Cherenkov lights in TOP.

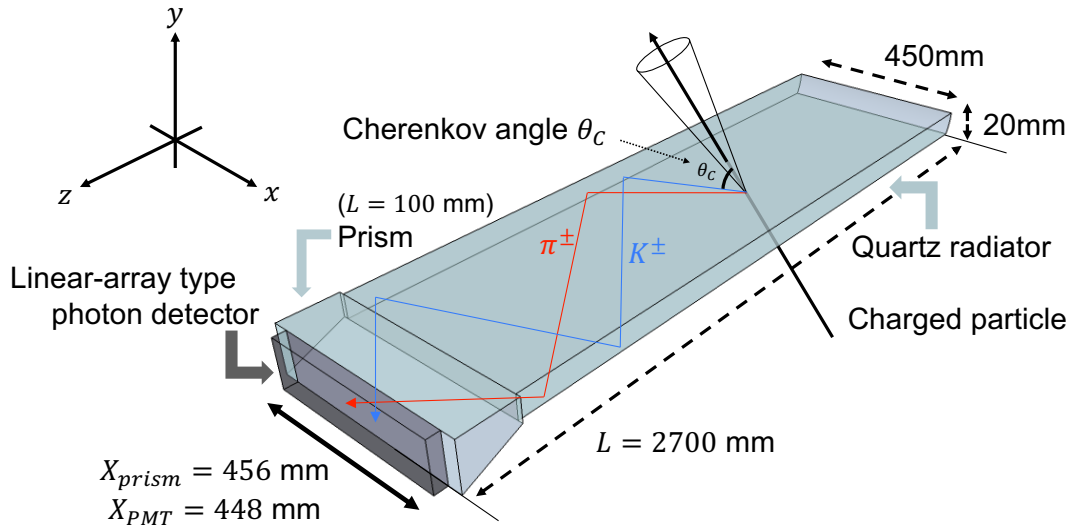


Figure 2.8: Basic geometry of TOP with charged particle and Cherenkov lights in case of charged K and π. Redrawn based on a figure in literature [21, 30]. Note that this figure's coordinate is unrelated to the lab frame discussed in the previous subsection.

TOP measures the photon with 2D coordinates of (x, y) shown in Fig. 2.8 and the propagation time of Cherenkov photons with better resolution than 100 ps. The distribution of position and propagation time of Cherenkov photons in TOP is compared with pre-evaluated PDF for particle hypotheses based on tracking information from the inner detector. TOP calculates the likelihood of hits with PDFs for each hypothesis and pro-

vides their ratios to determine the type of incident particles. Figure 2.9 shows an example hit distribution and calculated 2D PDF of an incident charged K track from the decay of $D^{*+} \to D^0(K^-\pi^+)\pi^+$.

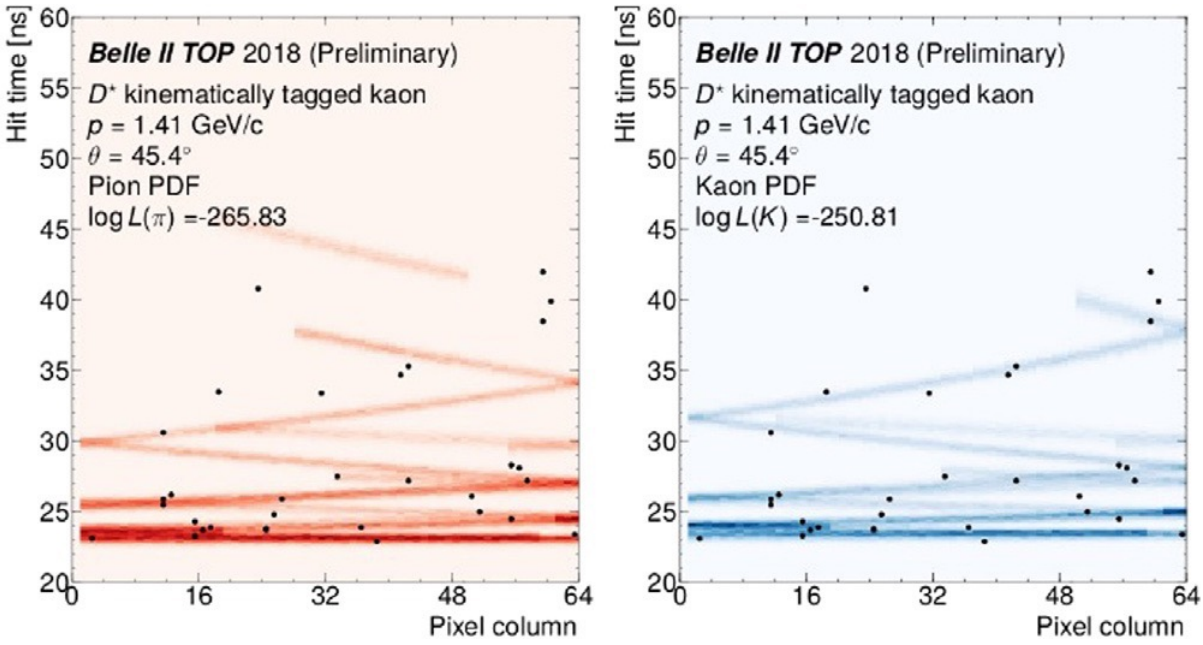


Figure 2.9: 2D PDF with a hypothesis of charged π and K overlapped with the 2D distribution of hits by position (Pixel column) and the time of the incident of the Cherenkov light from a charged K track. The Kaon hypothesis fits much better[30].

2.2.5 K_L^0 and μ Detector

K_L^0 and μ detector (KLM) is installed at the outermost part and surrounds all parts, including the solenoid of the Belle II detector, and identifies the two particles that can fly to outside of the Belle II detector due to their long lifetime. This subdetector consists of an alternating sandwich structure with the resistive plate chamber (RPC) or the plastic scintillators as an active detector and the 4.7cm-thick iron plates as an absorber. As shown in Fig. 2.2, KLM is separated into three parts of the barrel and forward or backward end-cap parts. The barrel part has an acceptance of $45° < \theta < 125°$ and the forward and backward end-caps provide an extended acceptance so that the total acceptance is $20° < \theta < 155°$.

The active volume of the subdetector in the barrel and end-cap part is different. KLM in the barrel part utilizes the RPC as an active part inherited from Belle, but the innermost two layers use the scintillator as an active material. On the other hand, the beam background in the end-cap region is severe, so RPC cannot deal with such an environment due to its long dead time. Therefore, KLM in the region adopted an active layer of scintillators equipped with wavelength-shifting fiber and Silicon Photomultiplier (SiPM) for its readout.

Particle identification of KLM starts from the information from the tracking system in the inner part, mainly CDC. μ is a charged particle that leaves signals in the tracking system and makes a track. This track is extrapolated to KLM and matched to a KLM hit cluster with penetrating depth. In the situation of K_L^0, the particle cannot make tracks because it is a neutral particle, so the hit cluster in KLM cannot be matched with the extrapolated tracks, and we identify the two particles from this feature.

2.2.6 Trigger and Data Acquisition System

The trigger and data acquisition systems are essential for all physics analysis in Belle II. The Belle II trigger system comprises hardware-based level-1 (L1) and software-based high-level triggers (HLT). L1 trigger processes data from subdetectors of CDC, ECL, TOP, and KLM in real time to reject backgrounds, detect physically interesting events and issue the L1 signal to record the corresponding event. A brief schematic for the L1 trigger system is shown in Fig. 2.10. Information from four subdetectors is combined in Global Reconstruction Logic (GRL), and Global Decision Logic (GDL) makes a final decision to issue the L1 trigger signal to record the corresponding event. Since events from B-mesons usually make tracks of more than three with large opening angles and significant energy deposition on ECL, the trigger system uses information from ECL and CDC as primary sources for the determination of the L1 signal, and KLM and TOP systems are used as auxiliary systems to determine the event timing. The goal of the trigger rate is 30 kHz, which is determined by DAQ system capability, and it is also required that the whole process for the trigger decision should be made within 5 μs due to requirements from ASICs of SVD electronics.

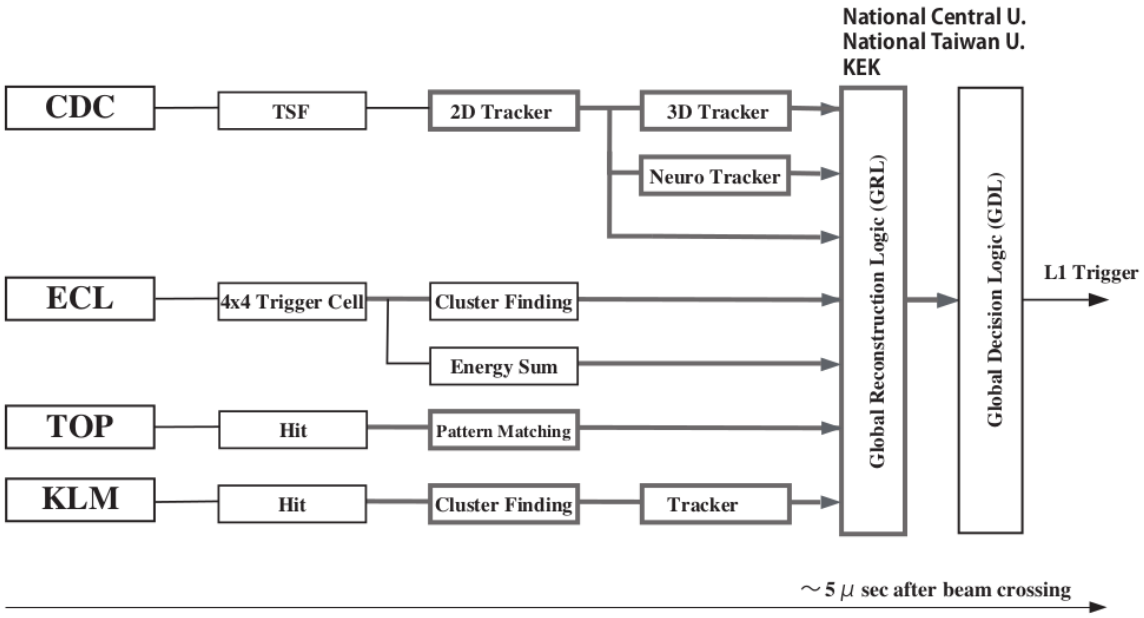


Figure 2.10: A block diagram of the Belle II trigger system.

HLT is performed on a computing cluster with a specially designed software framework named the Belle II Analysis Software Framework (basf2), a set of executables and libraries for the Belle II experiment[31]. If the L1 trigger is issued, the trigger signal is distributed by the front-end timing switch module (FTSW) to deliver the signal with proper timing to all destinations (e.g., front-end electronics for all subdetector). Then, subdetectors except for PXD will send data to HLT using the COPPER module, a standard module for data transferring in Belle II. Data from PXD are transferred by a unique path to select the meaningful PXD hits using track information extrapolated from SVD. On the computing cluster for HLT, the final hit selection of PXD and the complete reconstruction of an event from the event builder is performed. This process is illustrated in Fig. 2.11 and reduces the size of data sent by L1 triggers to 20% level.

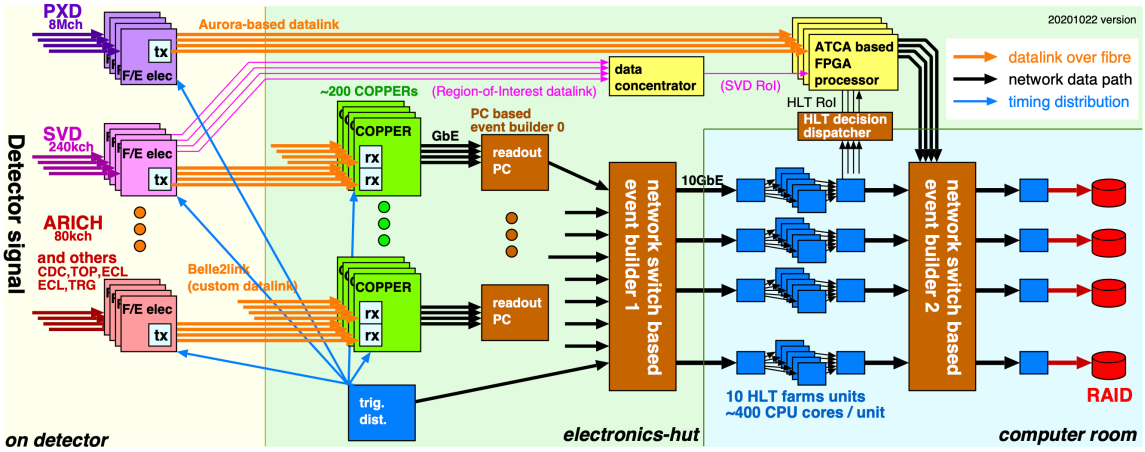


Figure 2.11: A scheme for the Belle II data acquisition[32].

Universal Trigger board 4 (UT4) The experimental condition of Belle II requires a large variety of calculations within the limited time interval to the trigger system of Belle II, and the system copes with such a complicated problem with a universal electronic board with an FPGA called Universal Trigger (UT) board. UT4, the fourth generation of the board, was introduced in 2019 with largely improved performance of the FPGA and component devices like optical interfaces from its predecessor of UT3, as given in Tab. 2.1. We develop firmware for basic functionality to utilize the boards for the experiment, such as input or output through VME bus, clock or temperature monitoring, and device management, and contribute to extending the capability of the trigger system of Belle II.

2.3 Data-taking status of the Belle II experiment

The Belle II started the Run 1 physics data taking in April 2019 and continued it until June 2022. Currently, the Belle II operation is stopped as Belle II is having the first Long-Shutdown (LS1) to perform the subdetector upgrades, such as installing the complete

Items		UT3	UT4
FPGA	Model	Xilinx Vertex 6 XC6VHX380 (565T)	Virtex UltraScale 7 XCVU080 (190)
	Number of logic gates	382K (580K)	975K (2026K)
Bandwidth of the optical interface		530 Gbps	1300 Gbps
Internal RAM		None	DDR4 32 GiB

Table 2.1: Comparison of Universal Trigger 3 and 4

second layer of PXD, adding radiation shields around VXD, and replacing PMTs of TOP to a long-lifetime version. Until starting LS1, SuperKEKB achieved its new world record of the instantaneous luminosity of $4.7 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, and Belle II collected $424 \pm 3\,\mathrm{fb}^{-1}$ of data from e^+e^- collision by SuperKEKB. The detailed status of integrated luminosity is shown in Fig. 2.12. Especially, SuperKEKB produced the $\Upsilon(4S)$ on-resonance collisions of the integral luminosity of $(362 \pm 2)\,\mathrm{fb}^{-1}$ in the pre-LS1 operation, which corresponds to $(387 \pm 6) \times 10^6$ pairs of B-mesons. We used the full dataset of the B-mesons for analysis in this dissertation.

Belle II is planning to accumulate a 50 times larger integrated luminosity than Belle. To achieve the goal, Belle II will have another long-shutdown around 2027 to upgrade the final beam-focusing magnets installed around the IP. The summary for the luminosity plan is shown in Fig. 2.13.

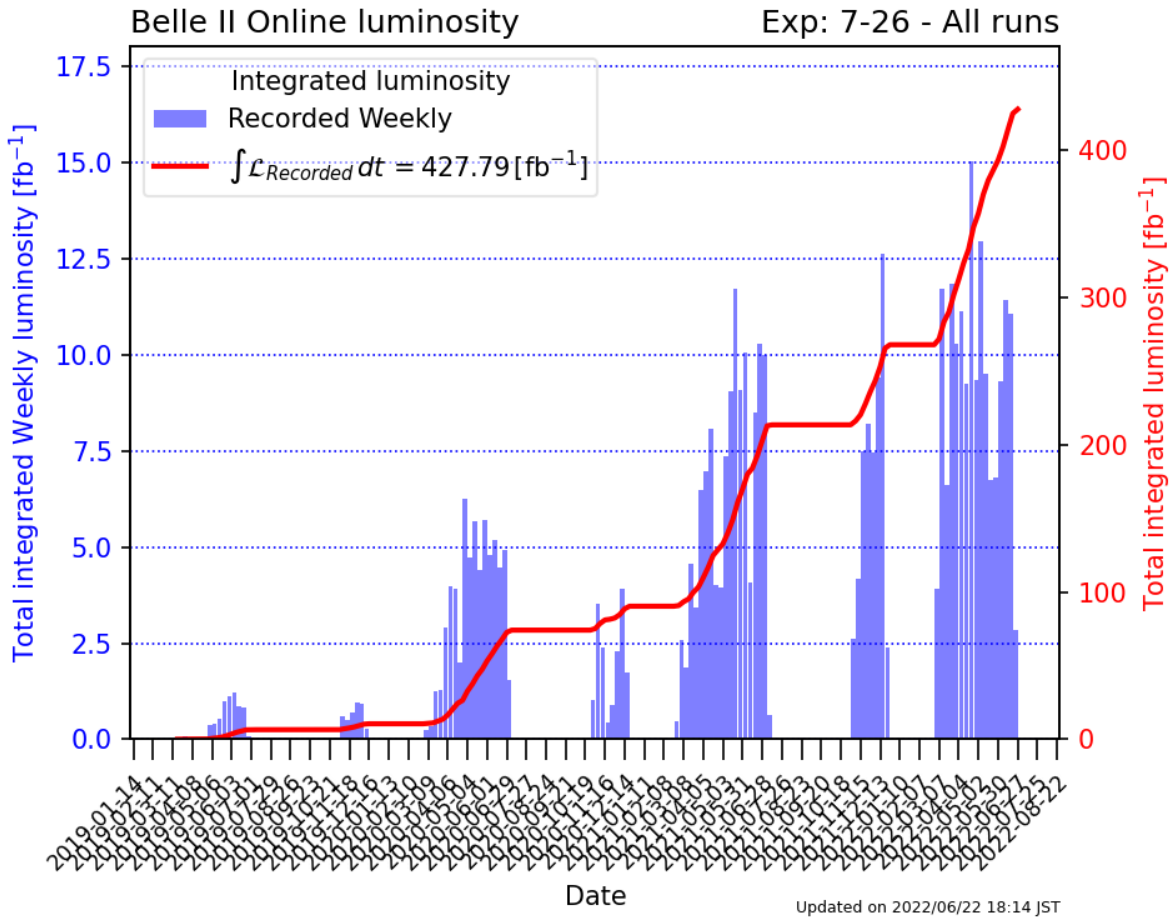


Figure 2.12: The total recorded luminosity before starting LS1[33].

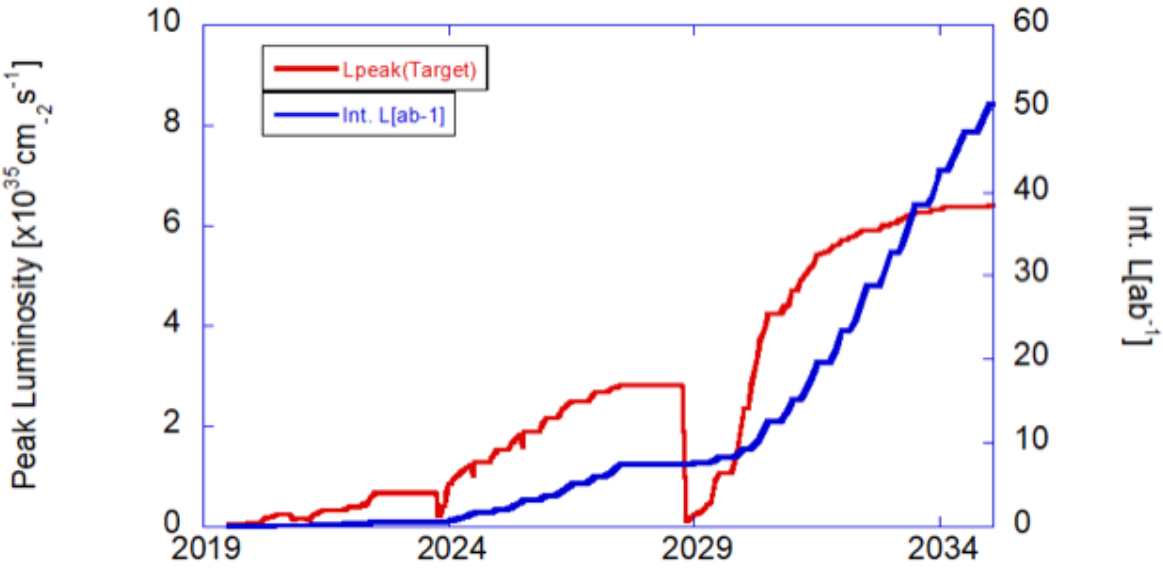


Figure 2.13: Future plan for luminosity for Belle II [33].

Chapter 3

Overview of the Analysis

In this chapter, we will discuss the theory of measurement for the *CP* violation in a specific scenario for the time evolution of a $B^0\overline{B}^0$ pair from decays of $e^+e^- \to \Upsilon(4S) \to B^0\overline{B}^0$ in the Belle II experiment and an overview of the strategy for our analysis for the measurement.

3.1 Time evolution of $B^0\overline{B}^0$ system from $\Upsilon(4S)$

In the Belle II experiment, a pair of the neutral B-mesons (B^0 and $\overline{B}^0$) is produced from $\Upsilon(4S)$ vector resonance created in the collision of asymmetrically accelerated electrons (7 GeV) and positrons (4 GeV). Although the pair oscillates, as discussed in Ch. 1, they never have the same flavors as B^0B^0 or $\overline{B}^0\overline{B}^0$ at the same time because they are produced at the $C = -1$ eigenvalue and preserve it. Using this feature, if we know the flavor of the B-meson on one side (we will call it as tag-side B meson or B_{tag}.) at a certain time t, we can identify the flavor of the B-meson on the opposite side (we will call it as *CP*-side B meson or B_{CP}.) as it will have the opposite flavor from its counterpart at the time t, even if we cannot identify the flavor of *CP*-side B mesons directly (i.e. *CP* side goes to a final state of *CP* eigenstates).

Let us consider the situation stated above quantitatively. If there is no asymmetry in the decay rate to the *CP*-eigenstate from B^0 and $\overline{B}^0$, an initial state at their production at the time of $t_{CP} = t_{tag} = 0$ with $C = -1$ eigenvalue is given as

$$|B_{CP}(t_{CP} = 0)B_{\text{tag}}(t_{tag} = 0)\rangle = \frac{1}{\sqrt{2}}\left(|B^0_{CP}\rangle\,|\overline{B}^0_{\text{tag}}\rangle - |\overline{B}^0_{CP}\rangle\,|B^0_{\text{tag}}\rangle\right). \tag{3.1}$$

Then, we can expand the above equation at a certain time of t_{CP} for B_{CP} and t_{tag} for B_{tag} with Eqs. 1.18 and 1.19 as follows:

$$
\begin{aligned}
|B_{CP}B_{\rm tag}(t_{CP},t_{tag})\rangle &= \frac{1}{\sqrt{2}}\left(|B^0_{CP}(t_{CP})\rangle\,|\overline{B}^0_{\rm tag}(t_{tag})\rangle - |\overline{B}^0_{CP}(t_{CP})\rangle\,|B^0_{\rm tag}(t_{tag})\rangle\right)\\
&= \frac{1}{\sqrt{2}}\Bigg(\frac{p}{q}[f_+(t_{CP})f_-(t_{tag}) - f_-(t_{CP})f_+(t_{tag})]\,|B^0_{CP}B^0_{\rm tag}\rangle\\
&\qquad + [f_+(t_{CP})f_+(t_{tag}) - f_-(t_{CP})f_-(t_{tag})]\,|B^0_{CP}\overline{B}^0_{\rm tag}\rangle\\
&\qquad + [f_-(t_{CP})f_-(t_{tag}) - f_+(t_{CP})f_+(t_{tag})]\,|\overline{B}^0_{CP}B^0_{\rm tag}\rangle\\
&\qquad + \frac{q}{p}[f_-(t_{CP})f_+(t_{tag}) - f_+(t_{CP})f_-(t_{tag})]\,|\overline{B}^0_{CP}\overline{B}^0_{\rm tag}\rangle\Bigg)
\end{aligned}
\tag{3.2}
$$

using Eq. 1.20, we can get the following relations:

$$
\begin{aligned}
f_+(t_1)f_-(t_2) - f_-(t_1)f_+(t_2) &= \frac{1}{4}e^{-iM_1(t_1+t_2)-\frac{\Gamma}{2}(t_1+t_2)}\big((1+e^{-i\Delta mt_1})(1-e^{-i\Delta mt_2})-\\
&\qquad (1-e^{-i\Delta mt_1})(1+e^{-i\Delta mt_2})\big)\\
&= \frac{1}{2}e^{-i(t_1+t_2)(M_1+\frac{\Delta m}{2})-\frac{\Gamma}{2}(t_1+t_2)}\left(e^{i\Delta m\frac{t_2-t_1}{2}} - e^{i\Delta m\frac{t_1-t_2}{2}}\right)\\
&= \frac{1}{2}e^{-i\frac{M_1+M_2}{2}(t_1+t_2)-\frac{\Gamma}{2}(t_1+t_2)}\left(-2i\sin\left(\frac{\Delta m(t_1-t_2)}{2}\right)\right)\\
&= -ie^{-\frac{\Gamma}{2}T}e^{-i\frac{M_1+M_2}{2}T}\sin\left(\frac{\Delta m\Delta\tau}{2}\right)
\end{aligned}
\tag{3.3}
$$

where $T = t_1 + t_2$ and $\Delta\tau = t_1 - t_2$. Similarly, we can get also

$$
f_+(t_1)f_+(t_2) - f_-(t_1)f_-(t_2) = e^{-\frac{\Gamma}{2}T}e^{-i\frac{M_1+M_2}{2}T}\cos\left(\frac{\Delta m\Delta\tau}{2}\right). \tag{3.4}
$$

Then, we can calculate Eq. 3.2 with the Eqs. 3.3 and 3.4 with the substitution of $t_1 = t_{CP}$ and $t_2 = t_{tag}$ as follows:

$$
\begin{aligned}
|B_{CP}B_{\rm tag}(\Delta\tau,T)\rangle = \frac{1}{\sqrt{2}}e^{-\frac{\Gamma}{2}T}e^{-i\frac{M_1+M_2}{2}T}\Bigg(&\cos\left(\frac{\Delta m\Delta\tau}{2}\right)\left(|B^0_{CP}\overline{B}^0_{\rm tag}\rangle - |\overline{B}^0_{CP}B^0_{\rm tag}\rangle\right)\\
&-i\sin\left(\frac{\Delta m\Delta\tau}{2}\right)\left(\frac{p}{q}\,|B^0_{CP}B^0_{\rm tag}\rangle - \frac{q}{p}\,|\overline{B}^0_{CP}\overline{B}^0_{\rm tag}\rangle\right)\Bigg).
\end{aligned}
\tag{3.5}
$$

The above equation proves that the pair of B-mesons cannot have the same flavor at $\Delta\tau = 0$, which means the same time ($t_{CP} = t_{tag}$). Then, we will consider a situation of the decay of the B_{CP} and the $B_{\rm tag}$. Let f_{tag} is a final state from $B^0_{\rm tag}$ and $\bar{f}_{tag}$ is a final state from $\overline{B}^0_{\rm tag}$ and define their amplitude as $a \equiv \langle f_{tag}|\,\mathcal{H}\,|B^0_{\rm tag}\rangle$ and $\bar{a} \equiv \langle \bar{f}_{tag}|\,\mathcal{H}\,|\overline{B}^0_{\rm tag}\rangle$. We will assume that f_{tag} cannot be reached from $\overline{B}^0_{\rm tag}$ so that $\langle f_{tag}|\,\mathcal{H}\,|\overline{B}^0_{\rm tag}\rangle = \langle \bar{f}_{tag}|\,\mathcal{H}\,|B^0_{\rm tag}\rangle = 0$.[1]

[1]Generally, this is not true, but the deviation on A_{CP} and S_{CP} due to this effect is sufficiently small at the per-mile level, so we can consider the effects as a systematic uncertainty, which is described in Sec. 6.3.10.

Also, we will redefine Eqs. 1.21, 1.22, and 1.23 for a final state f_{CP} which is a CP eigenstate: $\mathcal{A} \equiv \langle f_{CP}| \mathcal{H} |B^0\rangle$ and $\bar{\mathcal{A}} \equiv \eta_{CP} \langle f_{CP}| \mathcal{H} |\overline{B}^0\rangle$ where η_{CP} is the CP-eigenvalue of the final state. With these definitions and assumptions, we can obtain the decay amplitude for the tag-side flavor $B^0_{\rm tag}$ as follows:

$$
\begin{aligned}
|\langle f_{CP} f_{tag}| \mathcal{H} |B_{CP} B_{\rm tag}(\Delta\tau, T)\rangle|^2 &= \frac{1}{2} e^{-\Gamma T} \left| \left(a\bar{\mathcal{A}} \cos\left(\frac{\Delta m \Delta\tau}{2}\right) + a\mathcal{A}\frac{p}{q} i \sin\left(\frac{\Delta m \Delta\tau}{2}\right)\right)\right|^2 \\
&= \frac{1}{2} e^{-\Gamma T} |a\mathcal{A}|^2 \left|\frac{p}{q}\right|^2 \left| \left(\lambda_f \cos\left(\frac{\Delta m \Delta\tau}{2}\right) + i \sin\left(\frac{\Delta m \Delta\tau}{2}\right)\right)\right|^2 \\
&\approx \frac{1}{2} e^{-\Gamma T} |a\mathcal{A}|^2 \left(|\lambda_f|^2 \cos^2\left(\frac{\Delta m \Delta\tau}{2}\right) + \sin^2\left(\frac{\Delta m \Delta\tau}{2}\right) + \cos\left(\frac{\Delta m \Delta\tau}{2}\right) \sin\left(\frac{\Delta m \Delta\tau}{2}\right) (i\lambda_f^* - i\lambda_f)\right) \\
&\approx \frac{1}{2} e^{-\Gamma T} |a\mathcal{A}|^2 \left(|\lambda_f|^2 \frac{1+\cos(\Delta m \Delta\tau)}{2} + \frac{1-\cos(\Delta m \Delta\tau)}{2} + \Im(\lambda_f) \sin(\Delta m \Delta\tau)\right) \\
&\approx \frac{1}{4} e^{-\Gamma T} |a\mathcal{A}|^2 \left((|\lambda_f|^2 + 1) + (|\lambda_f|^2 - 1)\cos(\Delta m \Delta\tau) + 2\Im(\lambda_f)\sin(\Delta m \Delta\tau)\right)
\end{aligned}
\tag{3.6}
$$

where the approximation of $|q/p| \approx 1$ is used. The equivalent equation for the $\overline{B}^0_{\rm tag}$ can be obtained similarly.

$$
\begin{aligned}
|\langle f_{CP} \bar{f}_{tag}| \mathcal{H} |B_{CP} B_{\rm tag}(\Delta\tau, T)\rangle|^2 &\approx \frac{1}{2} e^{-\Gamma T} \left| \left(\bar{a}\mathcal{A} \cos\left(\frac{\Delta m \Delta\tau}{2}\right) - \bar{a}\bar{\mathcal{A}}\frac{q}{p} i \sin\left(\frac{\Delta m \Delta\tau}{2}\right)\right)\right|^2 \\
&\approx \frac{1}{2} e^{-\Gamma T} |\bar{a}\mathcal{A}|^2 \left| \left(\cos\left(\frac{\Delta m \Delta\tau}{2}\right) - i\lambda_f \sin\left(\frac{\Delta m \Delta\tau}{2}\right)\right)\right|^2 \\
&\approx \frac{1}{4} e^{-\Gamma T} |\bar{a}\mathcal{A}|^2 \left((|\lambda_f|^2 + 1) - (|\lambda_f|^2 - 1)\cos(\Delta m \Delta\tau) - 2\Im(\lambda_f)\sin(\Delta m \Delta\tau)\right)
\end{aligned}
\tag{3.7}
$$

In this context, $\Delta\tau = t_{CP} - t_{tag}$ becomes the difference of decay time between CP and tag-side B mesons, and T is the summation of their lifespan. Experimentally, we cannot measure the lifespan of each B-meson with a reliable method; T cannot be determined in our case. If we integrate the two above equations for T over the range of $0 < t_{CP} < \infty$ and $0 < t_{tag} < \infty$, we can get rid of T from the equation and get the following decay rate which is a function of $\Delta\tau$ and the flavor of tag-side B meson.

$$
\Gamma(\Delta\tau, q^{true}_{tag}) \propto e^{-\frac{|\Delta\tau|}{\tau_{B^0}}} \left(1 + q^{true}_{tag}(\mathcal{A}_{CP}\cos(\Delta m \Delta\tau) + \mathcal{S}_{CP}\sin(\Delta m \Delta\tau))\right) \tag{3.8}
$$

where q^{true}_{tag} is a sign function that indicates the true flavor of $B_{\rm tag}$ of which value is $+1$ for $B^0_{\rm tag}$ and -1 for $\overline{B}^0_{\rm tag}$, and $\tau_{B^0} \equiv 1/\Gamma$ is the lifetime of neutral B-mesons. The actual probability for extraction of CP-violating parameters from this situation will be discussed in Ch. 6.2

3.2 Overview of the measurement procedure

We will discuss how to extract A_{CP} and S_{CP} in Eq. 3.8 from the measurement results of the Belle II experiment in this section. Experimentally, we should reconstruct the decay of $B^0 \to \eta' K^0_S$, which is a final state of the CP eigenstate and need two pieces of information to extract CP asymmetries from the reconstruction result:

- Reconstructed proper decay time difference Δt, which is a estimated $\Delta\tau$, and
- The flavor of tag-side B meson (q_{tag}), which is a estimation result for q^{true}_{tag}.

Since we cannot know the "true" information of $\Delta\tau$ and q^{true}_{tag} but can estimate the two variables and get Δt and q_{tag}, we should consider the incompleteness of the estimation in extracting the CP asymmetries. Thus, we first reconstruct the target decay, then determine Δt and q_{tag}, and finally, extract CP-violating parameters by fitting the PDF to $(\Delta t, q_{tag})$ distribution with consideration of the incompleteness in the reconstruction or estimation of Δt and q_{tag}. This procedure is called "CP Fitting" in this dissertation, and its overview is shown in Fig. 3.1.

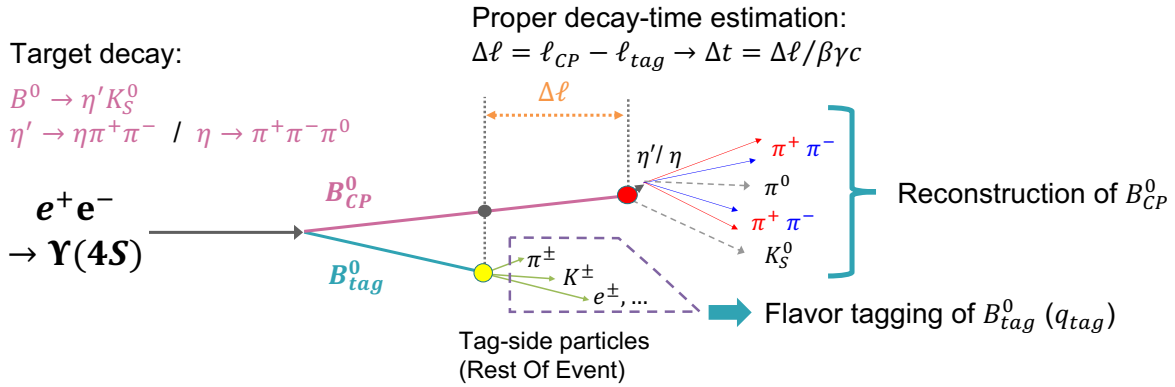


Figure 3.1: Definition of tag-side particles and a simplified view of the estimation procedure to determine the parameters of time-dependent CP violation.

We use a subdecay of $B^0 \to \eta' K^0_S$ with the final state particles of six charged pions and one neutral pion, which is $\eta' \to \eta\pi^+\pi^-$, $K^0_S \to \pi^+\pi^-$ and $\eta \to \pi^0\pi^+\pi^-$. Reconstruction of the subdecay is highly challenging since we need six charged particles and two gammas to reconstruct π^0, and details of the reconstruction will be explained in the later chapters. In the following subsections, we will discuss the techniques briefly which determine Δt and q_{tag} for our analysis.

3.2.1 Vertex fitting

To determine the proper time difference Δt, we should estimate the decay vertex of two neutral B-mesons. The vertex for B_{CP} and $B_{\rm tag}$ is determined by performing the

kinematical fittings with some constraints, called "Vertex fitting". From the vertex fitting, we measure the relative difference in lifespan between B-mesons by utilizing the difference of the z-coordinate of the decay vertex of the particles along the boost direction. From these two lengths, we calculated Δt as follows:

$$\Delta t = \frac{\ell_{CP} - \ell_{tag}}{\beta\gamma c} \tag{3.9}$$

where ℓ is the flight length of B-mesons along the boost direction, c is the speed of light, and γ and β is the Lorentz factor of $\Upsilon(4S)$ in the lab frame, which can be determined from the beam parameter. There are remaining things to take into account Δt properly, such as relativistic effects, moving of B-mesons in the frame of $\Upsilon(4S)$, and detector response on Δt reconstruction. We will consider those effects in the resolution function on the basis of event-by-event by using conditional variables. This matter will be discussed in a later chapter.

We utilized the treeFit algorithm[34] to determine the decay vertex of B_{CP}. It reconstructs the vertex position of B_{CP} and all intermediate particles in the decay chain using the helix information of charged particles in the final state. We used the vertex and momenta of B_{CP} and χ^2 value from the vertex fitting for our analysis. For the CP-side vertex fitting, two types of constraint are used. One of them is the kinematical constraint, which is that the momentum of intermediate-state particles of η', π^0, and K_S^0 is constrained so that its invariant mass is the value reported by the Particle Data Group[11]. Another one is IP constraint, which requires the virtual trajectory of B_{CP} starting from the IP, and its direction is the reconstructed momentum of B_{CP}, as shown on the left side of Fig. 3.2.

For the tag-side vertex fitting, we used KFit algorithm[35] with iterative removal treatment[36] with BTube constraint[37]. Especially, we do not know any kinetic constraints on the tag-side, such as the decay chain for the decay of B_{tag}. Thus, we perform a kinematic fit using all tracks with a constraint that the tracks originate from a single B_{tag} vertex. The BTube constraint is a requirement that the reconstructed vertex of B^0 decay should lie within a certain region of space known as the BTube. The BTube is defined as a cylinder around the axis of momentum of tag-side B meson, with dimensions determined by the estimated production point of B_{CP}^0 and its uncertainty, and the direction of the momentum of CP-side B meson. This procedure is illustrated on the right side subplot in Fig. 3.2. We can improve the resolution of tag-side vertex fitting and perform the vertex fitting even if there is only one reconstructed track on the tag-side. The iterative removal method is a method to reject tracks originating from long-lived intermediate states of charm quarks. If reduced χ^2 from vertex fitting is above threshold(20), then the most contributing tracks to χ^2 will be removed from vertex fitting. This procedure will be continued until the reduced χ^2 reaches the threshold or no more track in the input list.

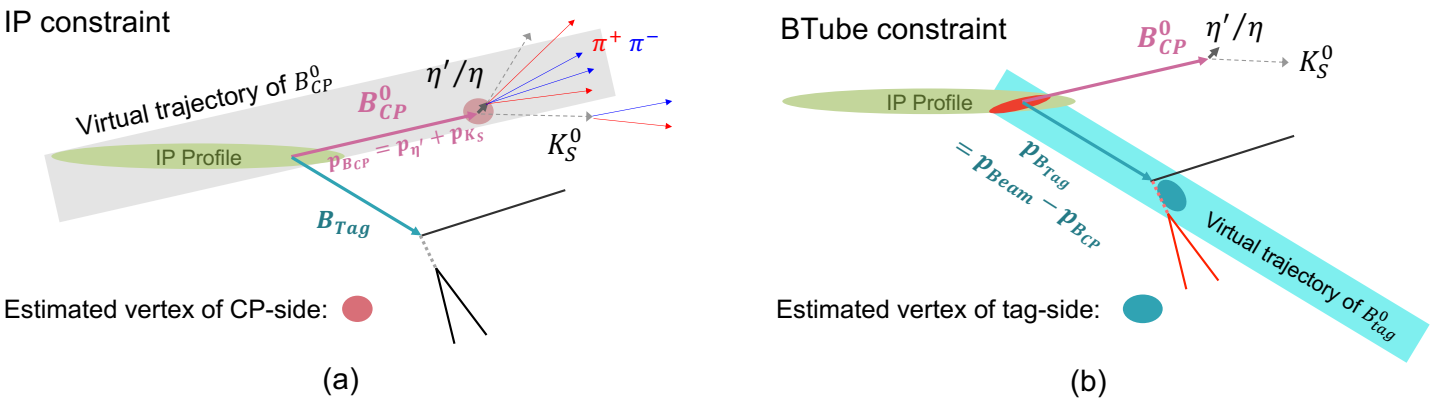


Figure 3.2: Graphical description of constraints in the vertex fitting for each side. (a) indicates IP constraint for *CP*-side vertex fitting, and (b) shows BTube constraint for tag-side vertex fitting.

3.2.2 Flavor tagging

Tagging the flavor of tag-side B meson is also essential information in *CP* fitting. We determined that information using Flavor Tagger[38] developed by the Belle II collaboration. The flavor tagger identifies the flavor of B_{tag} using the information of tag-side particles with the category-based method, as shown in 3.3. Each category is a Boosted-Decision Tree (BDT) that identifies the characteristic decay products from the CKM-flavored decays shown in the figure and calculates a score indicating an intermediate flavor tagging result. Then, the combiner that is also a BDT calculates the final output for the flavor tagging as a real number $q_{tag} \cdot r$, which distributes in the range of $[-1, 1]$ and this value is a product of two resultant values of q_{tag} and r. q_{tag} means the flavor of B^0_{tag} and r called "dilution factor" indicates the certainty of tagging result of q_{tag} and has the value in range of $[0, 1]$. If the resultant value is nearby 1, then the result of flavor tagging is highly probable to be the true flavor of B^0_{tag}, and vice versa. All reconstruction results will have the result since we need at least one track is needed for the tag-side vertex fitting, and the flavor tagger can produce a result for the flavor tagging even with only one input track.

The performance of Flavor Tagger can be parameterized by wrong tag fraction w, which suggests the probability of making the wrong decisions in flavor tagging and its difference of Δw by the true flavor of B^0_{tag}. Such parameters are used for our *CP* fitting, and their definition is as follows:

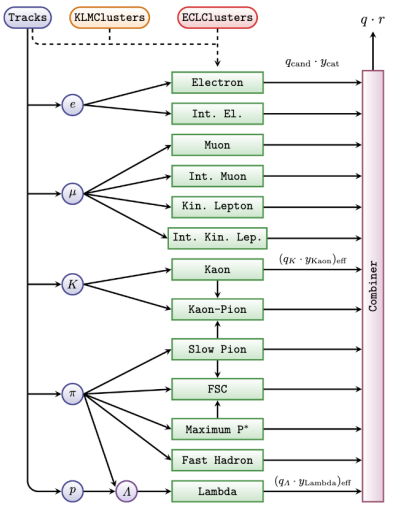
$$w = \frac{w_+ + w_-}{2}, \quad \Delta w = w_+ - w_- \tag{3.10}$$

$$\epsilon = \frac{\epsilon_+ + \epsilon_-}{2}, \quad \mu = \frac{\epsilon_+ - \epsilon_-}{\epsilon_+ + \epsilon_-} \tag{3.11}$$

where w_+ and w_- is wrong tagging probability when q^{true}_{tag} is +1(B^0) and -1($\overline{B}^0$), respec-

tively, $\epsilon_{+,-}$ is the tagging efficiency. w (or ϵ) is an average of $w_{\pm}$ (or $\epsilon_{\pm}$), Δw is the difference of $w_{\pm}$, and μ is the asymmetry of $\epsilon_{\pm}$. The effective efficiency ϵ_{eff} given by $\epsilon_{eff} = \epsilon \cdot (1 - 2w)$ summarizes the performance of Flavor Tagger as it simultaneously indicates the efficiency of Flavor Tagger and the dilution by the wrong decision for the flavor of B_{tag} in the estimation of CP asymmetries.

In practice, w_i and Δw_i where i is an index of a bin are provided in 7-bins for dilution factor r, which is binned with bin boundaries of $(0, 0.1, 0.25, 0.45, 0.6, 0.725, 0.875, 1)$. This binning for the dilution factor is called "r-bin" in this dissertation. All the w_i and Δw_i values along the binning are calibrated using the experimental data[39].



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

Underlying decay modes

$\overline{B}^0 \to D^{*+}\ \bar{\nu}_\ell\ \ell^-$
$\quad\hookrightarrow D^0\ \pi^+$
$\quad\quad\hookrightarrow X\ K^-$

$\overline{B}^0 \to D^+\ \pi^-\ (K^-)$
$\quad\hookrightarrow \overline{K}^0\ \nu_\ell\ \ell^+$

$\overline{B}^0 \to \Lambda_c^+\ X^-$
$\quad\hookrightarrow \Lambda\ \pi^+$
$\quad\quad\hookrightarrow p\ \pi^-$

Figure 3.3: Schematic structure for Flavor Tagger. Rearranged plot in literature[38].

3.2.3 Control channel and data samples

We validated our analysis procedure with the charged version of our target (sub)decay of $B^{\pm} \to \eta' K^{\pm}$ since the kinematics and vertex situation of the decay is similar to our target one. Also, a three times larger branching ratio than the target channel enables us to validate with statistical significance. In this dissertation, we call the neutral version of our target decay as "Main channel", and the charged version as "Control channel".

For this analysis, we used various samples from the Belle II collaboration. Firstly, the basic methodology of the analysis is established using Monte Carlo simulation (MC) samples generated with event generators[40, 41, 42, 43] for B-meson events or background events, the detector simulator using the Geant4 framework[44], and digitization and event reconstruction software developed by the Belle II collaboration. There are two types of MC samples as follows:

- Generic MC: This sample contains reconstruction results from the Belle II official simulations from the neutral and charged B-meson pair events and events by quark

pairs lighter than b quarks, as a result of e^+e^- collision, including not only the signal events but also hadronic background events from the experimental environment with our models.

- Signal MC: This sample includes only the reconstruction results for the target signal decay for our analysis.

Also, we used the full experimental dataset (called "Data") collected until starting LS1. The cosmic MC and Data, events by cosmic muons, are also used for auxiliary studies.

Chapter 4

Event Reconstruction, Selection and Signal Extraction

We reconstructed $B^0 \to \eta' K^0_S$ events using sub-channel of $K^0_S \to \pi^+\pi^-$, $\eta' \to \eta\pi^+\pi^-$, and $\eta \to \pi^+\pi^-\pi^0$ and elaborate on the procedure to reconstruct, select, and estimate the yield of our signal decay in this chapter.

4.1 Reconstruction

We reconstructed the target signal and control channel with optimized criteria. Firstly, we selected all the tracks with the following criteria for selection:

PID selection A ratio of Particle Identification (PID) likelihood that the given track is a charged pion (or kaon to pick $K^\pm$ for the control channel) > 0.1

IP selection $|dr| < 0.5\,\text{cm}$ `and` $|dz| < 2\,\text{cm}$, where dr is the track parameter d_0 relative to the IP, and dz is a similar one for track parameter of z_0.

Track quality selection Number of hits of track in CDC should be greater than 20, and it should be in the acceptance of CDC.

In addition, the covariance matrix of helix parameters for all tracks is scaled to correct an underestimation of the uncertainties of the parameters[45].

For the neutral pions, we reconstructed two gammas and selected the pairs with the invariant mass criteria of $0.120\,\text{GeV}/c^2 < M_{\gamma\gamma} < 0.145\,\text{GeV}/c^2$ and momentum criteria of $p_{\pi^0} > 0.204$. Requirements to gammas for π^0 reconstruction is $E^{BARREL}_\gamma > 0.03\,\text{GeV}$ for ones detected at barrel, $E^{FWD}_\gamma > 0.08\,\text{GeV}$ at the forward endcap, and $E^{BWD}_\gamma > 0.06\,\text{GeV}$ at the backward endcap.

We reconstructed K^0_S by pairing the oppositely charged pions with special treatments called "V0Finder"[46] for K^0_S due to its long lifetime ($c\tau \approx 2.7$ cm) and selected K^0_S

candidates with an invariant mass selection of $0.482\,\mathrm{GeV}/c^2 < M_{K_S^0} < 0.510\,\mathrm{GeV}/c^2$. Also, we applied selection criteria with the special BDT discriminator which is designed to pick the signal K_S^0 from the IP[45, 47].

The selection criteria for π^0 momentum, the invariant mass of K_S^0, and the BDT discriminator from K_S^0 selector are optimized to maximize FoM with the MVA technique, and the detailed procedure is explained in App. E.

Then, we combined π^0 and two oppositely charged pions to reconstruct η candidates. We selected these candidates by using their invariant mass of them with $0.52\,\mathrm{GeV}/c^2 < M_\eta < 0.57\,\mathrm{GeV}/c^2$. With these η candidates and two oppositely charged pions, we constructed η' candidates. Since the invariant mass of η and η' is highly correlated, we utilized the difference of them, which $\Delta M_{\eta'-\eta} = M'_\eta - M_\eta$ to select η' candidates with $0.40\,\mathrm{GeV}/c^2 < \Delta M_{\eta'-\eta} < 0.42\,\mathrm{GeV}/c^2$. The 1D distribution of M_η and $\Delta M_{\eta'-\eta}$ is shown in Fig. 4.1 and 2D distribution of the two variables is shown in Fig. 4.2.

For the B^0 selection, we used two variables: M_{bc} and ΔE, which can be calculated using the reconstruction result and the beam parameter. M_{bc} is a beam-constraint mass and ΔE is the difference of energy in the center-of-mass (CMS) frame between reconstructed B^0 and nominal energy from beam parameters. The definition of two variables is given in the following equations:

$$M_{bc} \equiv \sqrt{(E^*_{beam}/c^2)^2 - (p^*_B/c)^2} \tag{4.1}$$

$$\Delta E \equiv E^*_B - E^*_{beam} \tag{4.2}$$

where E^*_{beam} is a beam energy, p^*_B and E^*_B is the momentum and energy of reconstructed B^0, respectively, and an asterisk on the superscript indicates the quantities are measured in the CMS frame. We applied a loose selection of $M_{bc} > 5.17\,\mathrm{GeV}/c^2$ and performed vertex fitting using the treeFit algorithm, a vertex fitting algorithm discussed in the previous chapter. Then, we applied tight criteria of $M_{bc} > 5.2\,\mathrm{GeV}/c^2 \wedge |\Delta E| < 0.25\,\mathrm{GeV}$. For the tight criteria, M_{bc} and ΔE are calculated with information before the vertex fitting. This treatment prevents distortion of the shape of M_{bc} and ΔE distribution around the edge of the selection criteria due to the vertex fitting, as shown in Fig. 4.3.

If an event has multiple candidates for B^0_{CP} reconstruction, we selected B^0_{CP} having the lowest χ^2 value from the *CP*-side vertex fitting among the reconstructed candidates. After the best candidate selection, we additionally required at least two charged pions among four pions for η' reconstruction to have PXD hits and selection criteria of $\left(\frac{\chi^2}{ndf}\right)^{CP} < 10 \wedge \sigma_\ell^{CP} < 100\,\mu\mathrm{m}$ to the reconstructed B^0, where $\left(\frac{\chi^2}{ndf}\right)^{CP}$ is the reduced χ^2 and σ_ℓ^{CP} is the uncertainty of the fitted vertex of B_{CP}.

Information on tag-side particles is important for the vertex fitting of B_{tag} and the flavor tagging. This information is built by collecting the rest of the information in an event after the B_{CP} reconstruction with the following conditions:

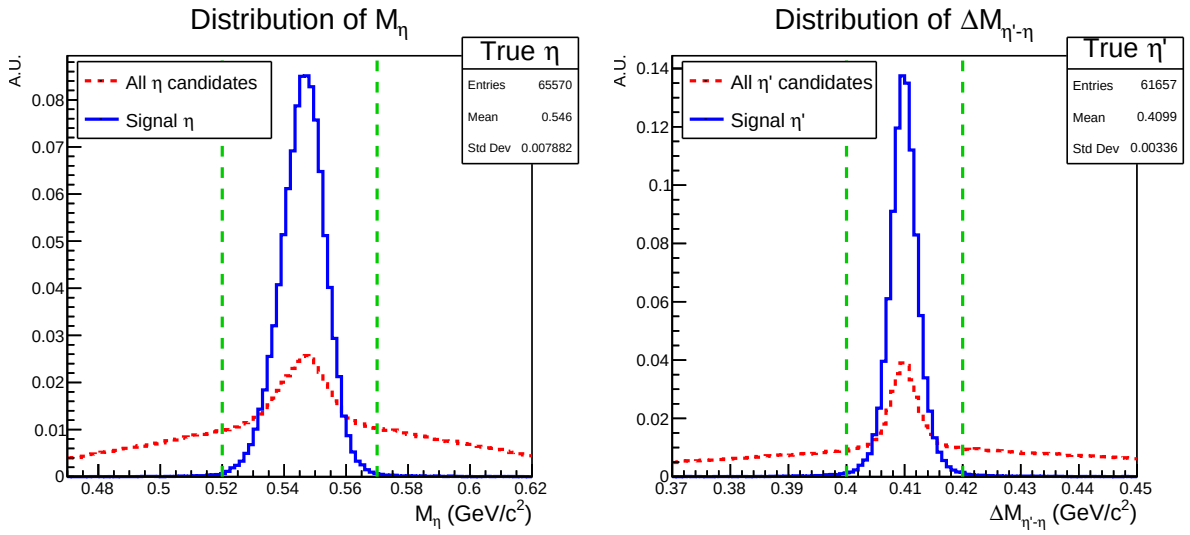


Figure 4.1: Distribution of invariant mass of η and difference between that of η' and η. The Green dashed lines indicate the threshold of selection for each quantity.

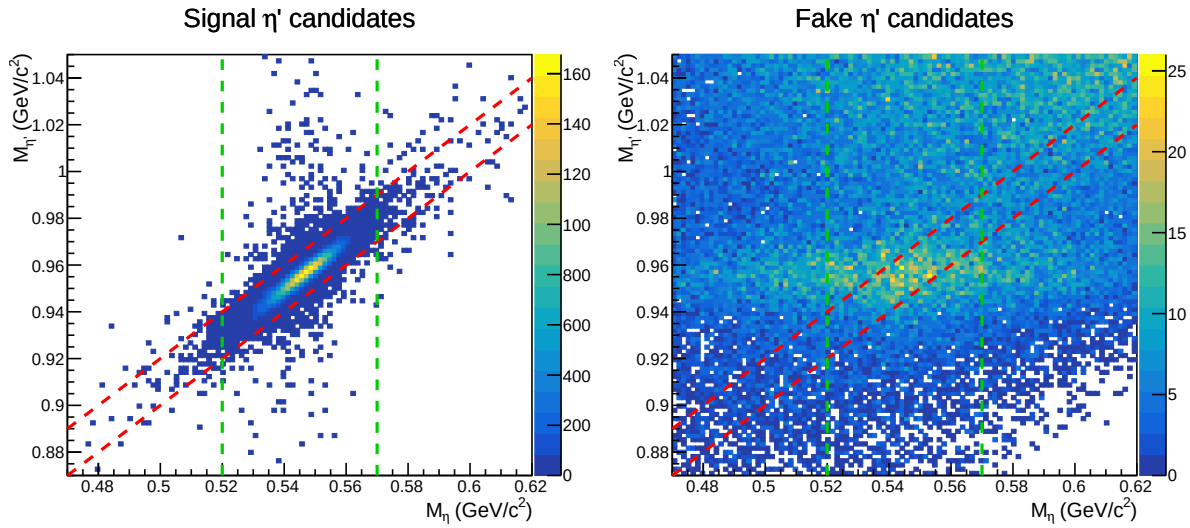


Figure 4.2: 2D scatter plot for M_η and M'_η. The areas enclosed with two red or green lines indicate the selection region for η' or η, respectively. Signal distribution is drawn with the signal MC samples, and the Fake one is plotted with the generic MC sample.

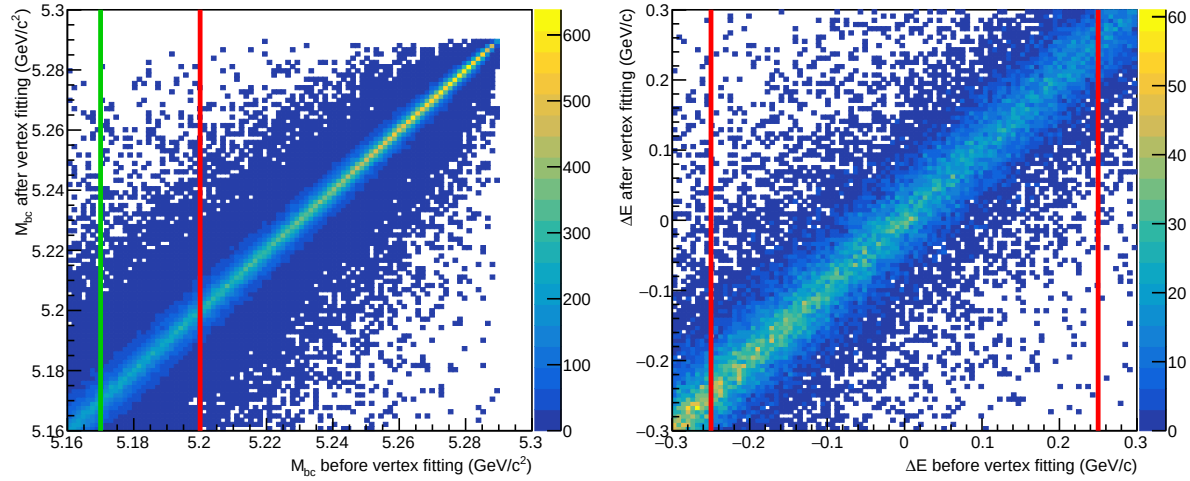


Figure 4.3: 2-D scatter plot of M_{bc} and ΔE between before and after vertex fitting. The green line shows thresholds of the loose selection for M_{bc}, and the red one indicates thresholds of tighter selection criteria.

- **Charged particles**: All charged tracks with the correction for the helix parameter uncertainty[45].
- **Photons**: All identified photons from ECL
- K_L^0: All K_L^0 candidates that identified by KLM

We performed the tag-side vertex fitting using the KFit algorithm, as discussed in the previous chapter, and we required additional requirements to input tracks from tag-side particles information as follows:

- should have one PXD hit at least,
- should not be a daughter of K_S^0,
- $d_0' < 500$ μm, and
- $\sigma_{z_0'} < 500$ μm

where d_0' and z_0' are the helix parameter d_0 and z_0 of tracks in the tag-side information that are relatively calculated for a decay vertex of B_{CP}^0, and $\sigma_{z_0'}$ means the uncertainty of z_0'. We determined whether a charged particle in the tag side is a daughter of K_S^0 or not with the invariant mass of a combination of two particles. If the invariant mass $m_{\pi^+\pi^-}$ is in the range of $M_{K_S^0} - 10\,\mathrm{MeV} < m_{\pi^+\pi^-} < M_{K_S^0} + 10\,\mathrm{MeV}$ where $M_{K_S^0}$ is the nominal mass of K_S^0, then we decided the particles came from K_S^0 and did not include them into the vertex fitting for tag-side. After the vertex fitting, we applied the selection criteria of $\sigma_\ell^{tag} < 100\mu$m, where σ_ℓ^{tag} is an uncertainty of the estimated vertex of B_{tag}.

We estimated parameters of w, Δw, ϵ, and μ that indicate the performance of the flavor tagger. For the MC sample, we calculated the parameters using the true tag-side flavor that can be determined using MC generation information. On the other hand, for experimental data, the parameters are estimated by analyzing flavor-specific control channels, such as $B^0 \to D^{(*)-}\pi^+$. The parameters from MC and calibrated one using Data are given in Tab. 4.1, and their comparison is shown in Fig. 4.4. Among these four parameters, we only consider three parameters of $w, \delta w$, and ϵ for the *CP* fitting. μ, a neglected parameter as it is zero-consistent in most r-bins and gives limited effects to the *CP* fitting, is considered in the systematic uncertainties as discussed in Sec. (1).

Dilution factor	ϵ (%)		μ (%)		w (%)		Δw (%)	
	Signal MC	Data	Signal MC	Data	Signal MC	Data	Signal MC	Data
0 - 0.1	15.42 ± 0.089	15.8 ± 0.13	-0.67 ± 0.57	-2.4 ± 1.2	48.00 ± 0.31	48.04 ± 0.54	0.49 ± 0.62	-0.69 ± 1.1
0.1 - 0.25	14.85 ± 0.087	15.53 ± 0.13	0.42 ± 0.59	1.4 ± 1.2	41.10 ± 0.31	42.4 ± 0.54	1.6 ± 0.62	3.8 ± 1.1
0.25 - 0.45	15.77 ± 0.089	16.52 ± 0.14	-0.65 ± 0.57	-1.2 ± 1.2	32.25 ± 0.29	34.10 ± 0.51	-0.79 ± 0.58	-1.9 ± 1
0.45 - 0.6	13.66 ± 0.084	13.92 ± 0.13	0.19 ± 0.62	0.88 ± 1.3	23.28 ± 0.28	23.62 ± 0.53	-0.31 ± 0.56	-0.70 ± 1.1
0.6 - 0.725	11.72 ± 0.079	11.61 ± 0.12	1.1 ± 0.67	3.9 ± 1.3	16.30 ± 0.26	16.75 ± 0.53	-0.27 ± 0.53	2.0 ± 1.1
0.725 - 0.875	11.48 ± 0.078	11.45 ± 0.12	0.72 ± 0.68	-2.0 ± 1.3	9.611 ± 0.21	10.73 ± 0.48	-0.93 ± 0.43	-0.054 ± 0.96
0.875 - 1	17.10 ± 0.092	15.17 ± 0.13	-0.52 ± 0.54	-1.2 ± 1.1	1.978 ± 0.083	2.737 ± 0.3	0.085 ± 0.17	0.24 ± 0.6
$\epsilon_{eff} = \sum_i(\epsilon_i \cdot (1-2w_i)^2)$	34.97 ± 0.2	31.68 ± 0.38						

Table 4.1: Parameters for Flavor Tagger from Signal MC and Data. Parameters for Data are calibrated using another channel.

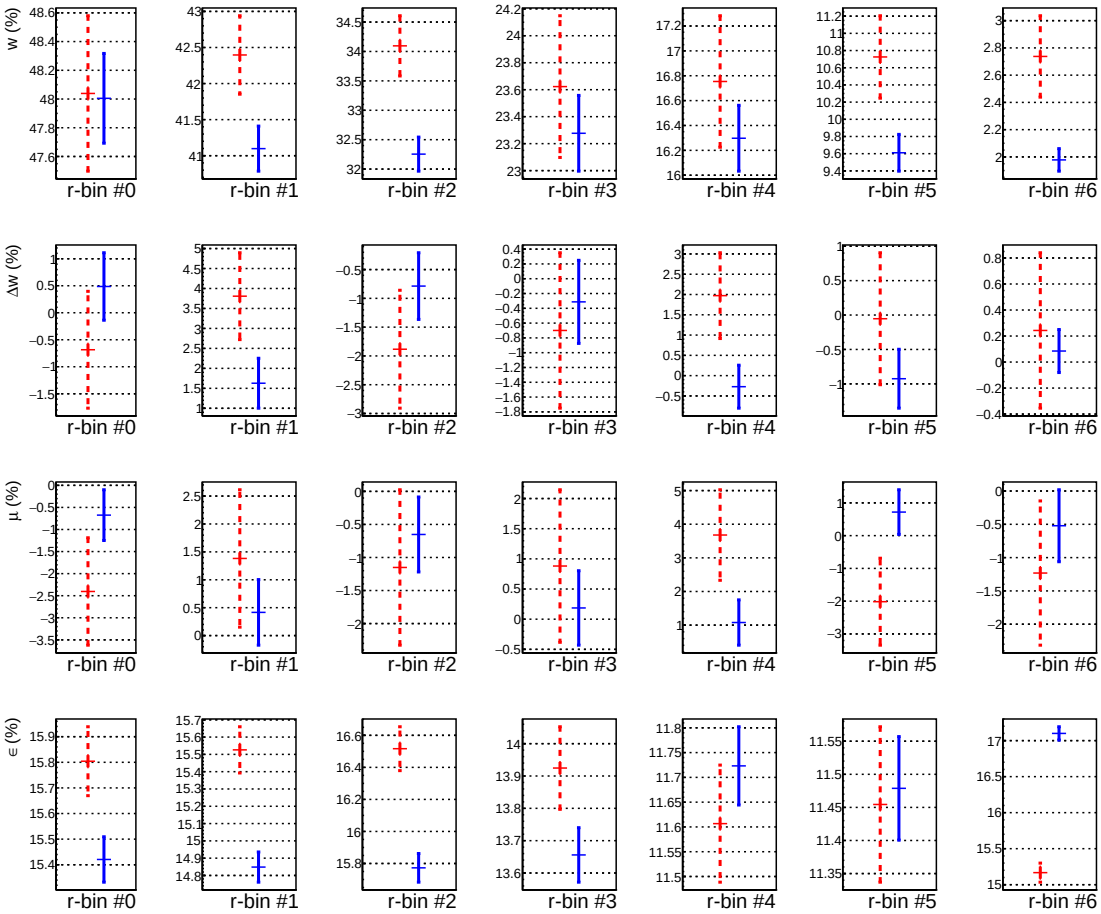


Figure 4.4: Comparison of parameters for flavor tagger between our signal MC (Blue solid) and calibrated results (Red dashed). Parameters from signal MC are calculated using signal and π exchanged SxF events. (See Sec. 4.3.1)

4.2 Continuum Suppression

The main source background of this analysis is $q\bar{q}$ event ($q = u, d, c, s$) from productions of a pair of quarks lighter than b quarks. Since the center of mass energy of the e^+e^- beam is tuned to about the same as the mass of a pair of $B\overline{B}$, the final state particles from $q\bar{q}$ are boosted by the extra energy. This difference between $B\overline{B}$ and $q\bar{q}$ makes the difference in the shape of events, which indicates the momentum distribution of particles in an event, and the situation is illustrated in Fig. 4.5. We can reject the $q\bar{q}$ events by utilizing this feature.

We used the following variables that indicate the event shapes[10] to distinguish $B\overline{B}$ and $q\bar{q}$ events:

- 15 modified Fox-Wolfram moments, called KSFW: The Fox-Wolfram moments[48] is a coefficient of the expansion of the momentum distribution in an event by the

spherical harmonics. Those variables are usually used to discriminate $b\bar{b}$ and $q\bar{q}$ events as they have different shapes of the momentum distribution.

- 9 CLEO Cone variables[49]: A summation of the momentum for particles that passes a region around the signal thrust axis divided into a cone-shaped section at nine polar angle intervals with 10° slicing.
- `thrustOm`: The magnitude of the thrust of tag-side particles, where the thrust is a 3-D vector that maximizes the summation of the projections of all the momenta of tag-side particles onto it.
- `cosTBTO`: Cosine value of the angle between the thrust axis of particles used to reconstruct the *CP*-side B meson and the thrust axis of tag-side particles.

The distribution of these variables for signal and background events is shown in App. H. We trained a boosted decision tree using the FastBDT library[50] to calculate a comprehensive index that considers all of the above 26 variables. For particles for the input to the training, we applied requirements for charged particles and photons in the tag-side information as follows:

- **Charged particles**: Should have CDC hits and $p^* <= 3.2\,\mathrm{GeV}/c$ with π hypothesis
- **Photons**: $p >= 0.05\,\mathrm{GeV}/c$ and $p^* <= 3.2\,\mathrm{GeV}/c$

where p^* is the momentum in the CMS frame.

The output value $\mathcal{O}'_{CS}$ from the boosted decision tree is a real number within a range of $(0, 1)$. We transformed the value to $\mathcal{O}_{CS}$ with the following equation:

$$\mathcal{O}_{CS} = \log\left(\frac{\mathcal{O}'_{CS}}{1-\mathcal{O}'_{CS}}\right) \tag{4.3}$$

Distribution of $\mathcal{O}_{CS}$ from MC samples of (Signal + SxF) and Background events and their comparison are shown in Fig. 4.6. The SxF category is the Self-Crossfeed event where we partially failed the reconstruction of our target decay, which will be discussed in the next section.

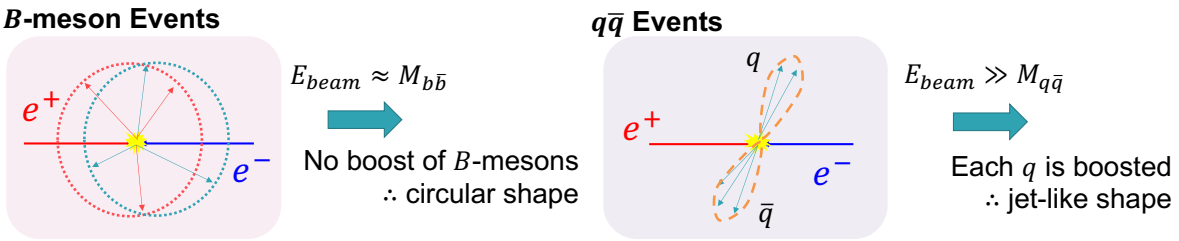


Figure 4.5: Comparison of event topologies between events from B-mesons and $q\bar{q}\,(q = u, d, s, c)$.

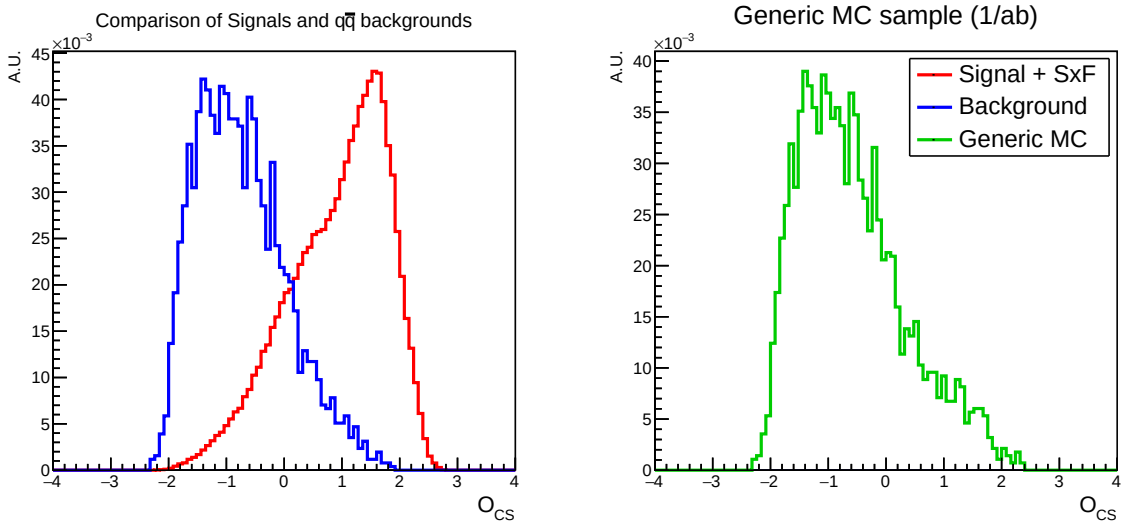


Figure 4.6: Comparison of $\mathcal{O}_{CS}$ between (Signal + SxF) and Background events (left) and distribution of $\mathcal{O}_{CS}$ from the 1/ab generic MC sample (right). The definition of (Signal + SxF) and Background events are discussed later. Each distribution is normalized to the unit area to compare the shapes.

4.3 Classification of reconstruction results

We classified reconstruction results into two categories: (Signal + SxF) and Background category. The reconstruction results in the Signal category are successful results in reconstructing our target decay, and the Self-Crossfeed (SxF) category contains results that partially failed in the reconstruction, even though our target decay exists in the event. We confirmed the two categories can be treated in the same footprint for the extraction of *CP* asymmetry, which will be discussed in the following subsection; we merged two categories into the (Signal + SxF) category. The Background category consists of all reconstruction results from events where our signal decay does not exist, of which the main source is $q\bar{q}$ events.[1] We consider events in the former category to be signal events that contain physically meaningful information for *CP* asymmetry measurement or lifetime determination and those in the latter category to be background events that we need to reject the effects from.

4.3.1 Self-Crossfeed (SxF) events

We studied the feasibility of using SxF events for *CP* fitting. The SxF events are events in which we have failed to reconstruct the signal decay from an event where the signal

[1] We could not find meaningful background events other than $q\bar{q}$ events in the MC study. (The number of events is $\mathcal{O}(1)$ from the whole available generic MC samples.) Thus we solely focused on the study of $q\bar{q}$ events for the background events.

decay exists. There are two types of SxF events: π-exchanged and Combinatorial.

π-exchanged SxF event is an event judged as backgrounds in the Belle II analysis software due to the swapping of charged pions between from η' and from η. Let us suppose a decay of $\eta' \to \eta\pi_1^+\pi_2^-$ and $\eta \to \pi^0\pi_3^+\pi_4^-$ where $\pi_i^\pm (i = 1, 2, 3, 4)$ is each instance of charged pions. If we reconstruct B^0 candidate with true-K_S^0 and η' of $\eta' \to \eta\pi_3^+\pi_2^-$ and $\eta \to \pi^0\pi_1^+\pi_4^-$ with true-π^0 in this situation, then the software framework will classify such reconstruct result as a fake B^0 candidate. However, we cannot distinguish the decay vertices of η' and η with our vertex resolution, so we can assume that four instances of charged pions are emitted from the same vertex, and the exchanging cannot affect Δt determination at all. Also, all particles for the reconstruction originated from B^0_{CP}, so M_{bc} and ΔE will be the same as that from signal events. Thus, we can treat signal B^0 events and the π-exchanged SxF events in the same footprint. We compared the distribution of M_{bc}, ΔE, and Δt from signal events and π-exchanged SxF events in the first subplot in Figs. 4.7, 4.8, and 4.9. There are no significant differences in the distributions between the two categories. We used these SxF events the same way as signal events for further analysis, such as resolution parameter estimation or determining pre-fit parameters.

The remaining SxF events are called the Combinatorial SxF events. The tag side or fake particles contaminate ingredients for B^0_{CP} reconstruction in this category, so M_{bc}, ΔE and Δt distribution are different from that in signal events. We classified Combinatorial SxF events into five categories. The distribution and the number of events from each category are shown from the second to sixth subplots in Fig. 4.7, 4.8, or 4.9 and Tab. 4.2. Almost 60% of events in this category are only-π^0-fake events. We can expect Δt distribution might not change many from signal events because π^0 does not participate in the vertex fitting. The second and third subplots of Fig. 4.9 show our expectation holds.

Since the *CP* fitting estimates *CP* asymmetry from an asymmetry of Δt distribution between $q_{tag} = \pm 1$, confirming the asymmetry of SxF events is important; We investigated the asymmetry of reconstructed Δt using the true flavor q_{tag}^{true} of $B_{\rm tag}$ using MC generation information. We calculated the uncertainty using the Clopper-Pearson confidence interval and fitted the asymmetry plot with a sine curve. The distribution of the asymmetries and fitted curve and parameters are shown in Fig. 4.10. It is worth noting that even SxF events having contamination by charged particles from the tag side or fake tracks show evident *CP* asymmetries in the plot, although the Δt distribution of the events has distorted from signal events. This suggests we can similarly treat all SxF categories as signal events with respect to *CP* fitting.

We compared parameters for the flavor tagger, such as wrong tagging fraction or tagging efficiency per a bin of the dilution factor between signal and two categories for SxF events. (Figure 4.11) There are no significant differences between the three categories except for efficiencies. We expect the difference in efficiency per each bin of dilution factor comes from the contamination on the tag side. The effects of this difference on the *CP* fitting will be taken into account in the systematic study.

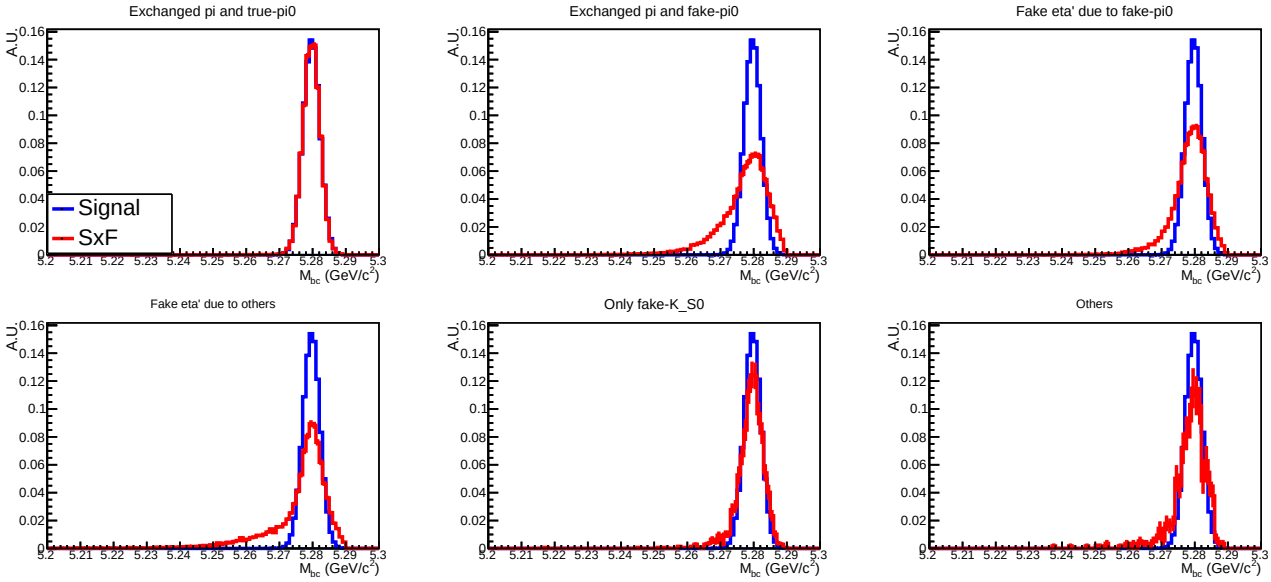


Figure 4.7: Comparison of M_{bc} between Signal and Combinatorial SxF events. The blue line indicates a distribution from signal events, and the red line shows that from SxF events of each category

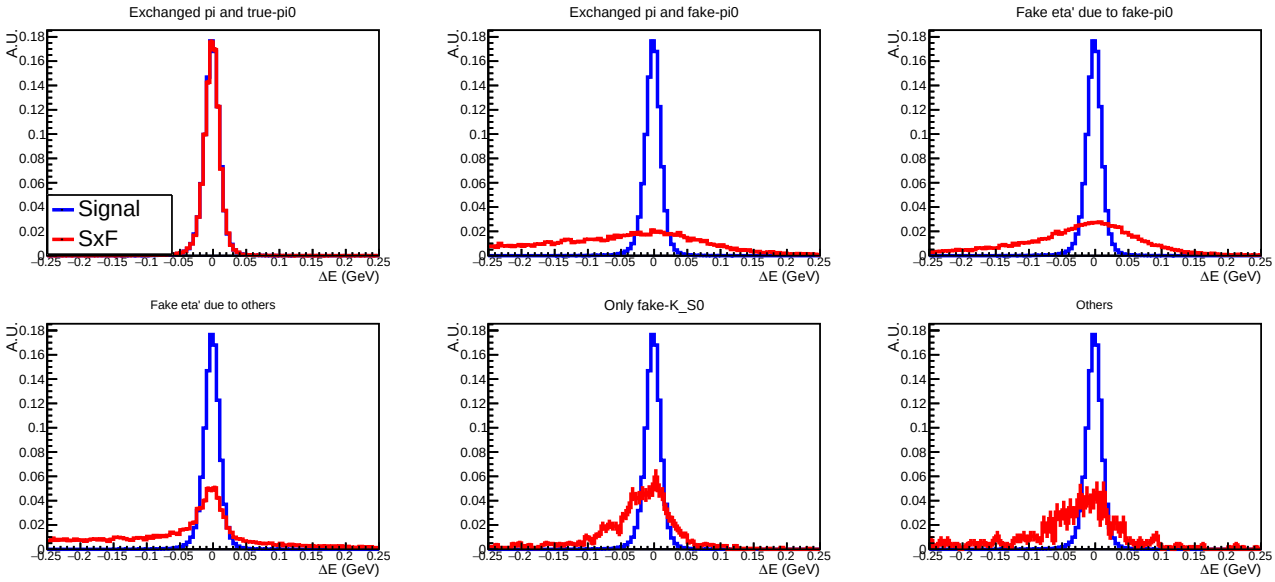


Figure 4.8: Comparison of ΔE between Signal and Combinatorial SxF events. The blue line indicates a distribution from signal events, and the red line shows that from SxF events of each category

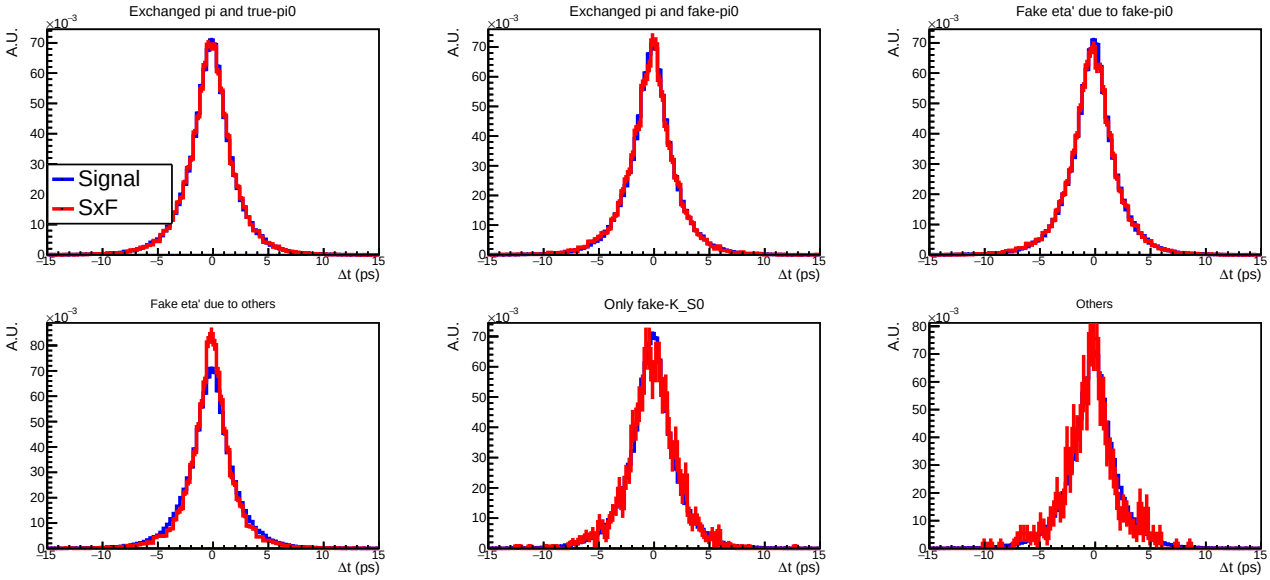


Figure 4.9: Comparison of Δt between Signal and Combinatorial SxF events. The blue line indicates a distribution from signal events, and the red line shows that from SxF events of each category

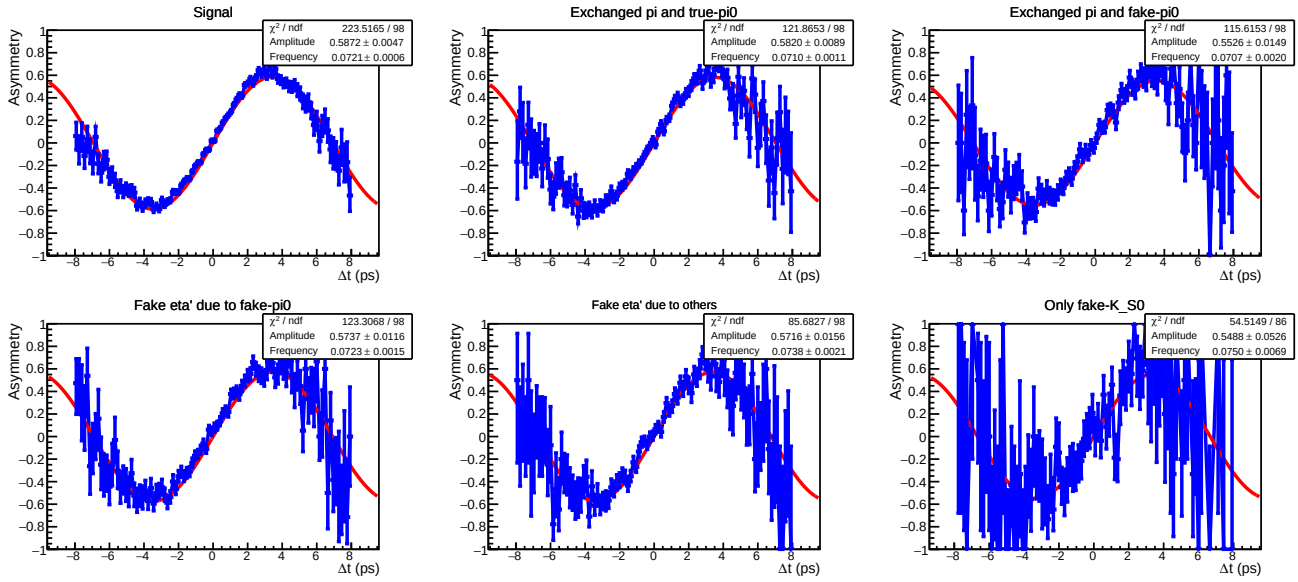


Figure 4.10: Asymmetry plot for signal and SxF events. We determined q_{tag} with MC truth information and calculated uncertainties with the Clopper-Pearson confidence interval.

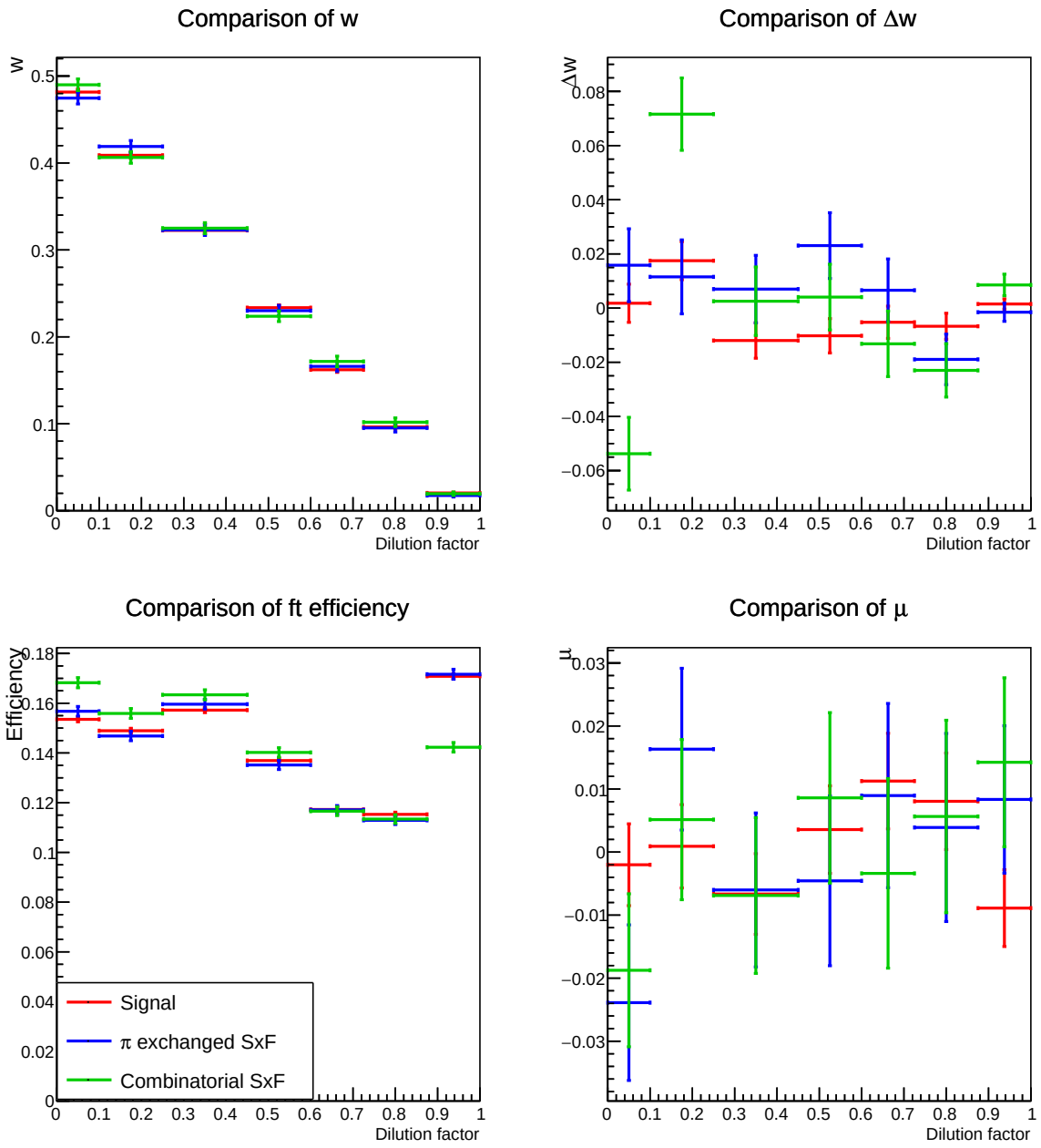


Figure 4.11: Comparison of flavor tagger parameters between Signal and SxF events.

Finally, we compared $\mathcal{O}_{CS}$ distributions between Signal and two types for SxF events, and Fig. 4.12 shows the comparison results. The shape of $\mathcal{O}_{CS}$ distribution between Signal and π-exchanged SxF is consistent, but that from Combinatorial SxF is slightly different from the other two. We consider that the contamination on the tag side made such a difference. Based on these studies, we include all the SxF events in the *CP* fitting and used the same configuration, such as resolution parameters.

Type of Events	Count	Fraction [%]
Signal event	130965	60.37
π-exchanged SxF event	35493	16.36
π-exchanged SxF event with fake-π^0	13655	6.29
η'-fake SxF event due to fake-π^0	20692	9.54
η'-fake SxF event due to other reasons	14198	6.55
Only K^0_S-fake SxF event	1366	0.63
Other SxF event	557	0.26

Table 4.2: Number of events of signal and SxF events of each category. Those values were counted using MC samples and the selection criteria stated Sec. 4.1.

4.4 Selection efficiencies

We calculated the efficiency for selection in the following summary.

- PID selection for charged π or K
- IP and track quality selection
- π^0 selection criteria
- $0.52\,\mathrm{GeV}/c^2 < M_\eta < 0.57\,\mathrm{GeV}/c^2$
- $0.40\,\mathrm{GeV}/c^2 < \Delta M_{\eta'-\eta} < 0.42\,\mathrm{GeV}/c^2$
- $M_{bc} > 5.17\,\mathrm{GeV}/c^2$ before the vertex fitting for B^0_{CP}, and $M_{bc} > 5.2\,\mathrm{GeV}/c^2 \wedge$

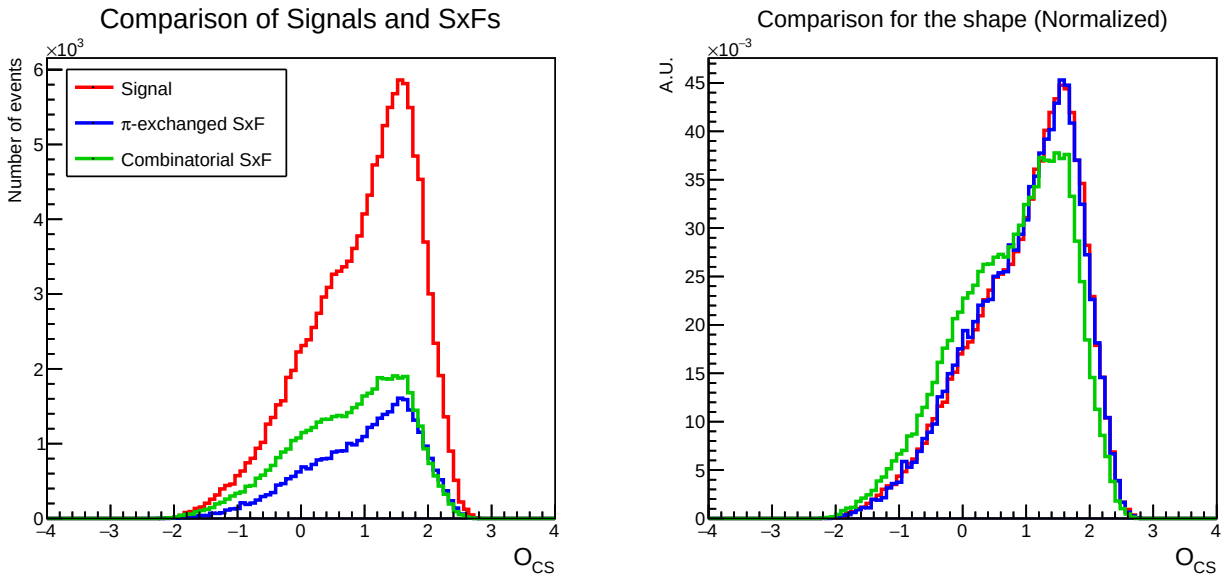


Figure 4.12: Comparison of $\mathcal{O}_{CS}$ distribution between Signal and two types for SxF events. The left plot shows the raw distribution, and the right one is the normalized plot of which the area is the unit area.

$|\Delta E| < 0.25\,\mathrm{GeV}$ after the vertex fitting. Note that M_{bc} and ΔE values are calculated with information before the vertex fitting.

- The flavor tagger and vertex fitting for tag-side succeed.
- The best candidate selection with a p-value of the vertex fitting for the CP side.
- Two pions at least should have PXD hits
- $p_{\pi^0} > 0.204\,\mathrm{GeV}/c$
- $0.482\,\mathrm{GeV}/c^2 < M_{K^0_S} < 0.510\,\mathrm{GeV}/c^2$ and `KsSelector > 0.8`
- $\sigma_z^{CP} < 50\mu\mathrm{m}$ and $\left(\frac{\chi^2}{ndf}\right)^{CP} < 10$
- $\sigma_z^{tag} < 100\mu\mathrm{m}$ and $\left(\frac{\chi^2}{ndf}\right)^{tag} < 20$

Table 4.3 shows the cumulative efficiencies for selecting B^0. We included all SxF events in the calculation for the table and used all signal and generic (1/ab) samples. The equivalent table for the control channel is given in App. B.

Selection	Signal + SxF events	Background events
η', η Selection, $M_{bc}, \Delta E$ and the best candidate selection	15.38% (15.38 %)	– (100 %)
$p_{\pi^0} > 0.204$ GeV/c	92.06% (14.16%)	84.01% (84.01%)
$0.482 < M_{K^0_S} < 0.510$, and `KsSelector > 0.8`	93.64% (13.25%)	16.39% (13.77%)
Selection for PXD hits	99.79% (13.23%)	99.67% (13.73%)
$\left(\frac{\chi^2}{ndf}\right)^{CP} < 10$	88.43% (11.7%)	68.15% (9.354%)
$\sigma_\ell^{CP} < 50\mu$m	97.17% (11.36%)	94.67% (8.855%)
Tag-side vertex fitting succeeded	98.3% (11.17%)	97.37% (8.622%)
$\sigma_\ell^{tag} < 100\mu$m	97.35% (10.88%)	97.41% (8.399%)
Flavor tagging succeeded	100% (10.88%)	100% (8.399%)

Table 4.3: The efficiency for the reconstruction of $B^0 \to \eta' K^0_S$. The percentage in parentheses is a cumulative efficiency, and the value for the Background category is the relative one for efficiency after the best candidate selection for B^0_{CP}.

4.5 Overview of the signal extraction

We estimated the yield of events in the (Signal + SxF) and Backgrounds categories and used three variables of $(M_{bc}, \Delta E, \mathcal{O}_{CS})$ to extract the yields. We added two variables of $\cos(\theta_B^{Boost}), |q_{tag} \cdot r|$ to calculate the signal probability f_{sig} on an event-by-event basis.

$\cos(\theta_B^{Boost})$ is the cosine value for the angle between the momentum of B^0_{CP} and the z-axis in the CMS frame and $|q_{tag} \cdot r|$ is a dilution factor from the flavor tagger. Since the decay of $\Upsilon(4S) \to B^0\overline{B}^0$ is a $V \to PP$ decay, $\cos(\theta_B^{Boost})$ are given by

$$\frac{3}{4}(1 - \cos^2(\theta_B^*)), \tag{4.4}$$

theoretically. Meanwhile, $\cos(\theta_B^{Boost})$ value of $q\bar{q}$ event does not have such constraint; its distribution will be flat. The distributions of $|q_{tag} \cdot r|$ values for (Signal + SxF) and

Background events are also different because the flavor tagger is tuned to pick the CKM-flavored decay of b quarks. We use $\cos(\theta_B^{Boost})$ and $|q_{tag} \cdot r|$ as conditional variables for the *CP* fitting, and according to Punzi[51], we should take into account such differences for the conditional variables to avoid bias and wrong error estimation in the fitting. f_{sig} is the probability that an event is a signal one with consideration of the effect, which is given as the following formula:

$$f_{sig}(\vec{x}) = \frac{F_{sig}P_{sig}(\vec{x})}{F_{sig}P_{sig}(\vec{x}) + (1 - F_{sig})P_{bckg}(\vec{x})} \tag{4.5}$$

where F_{sig} is the overall fraction of (Signal + SxF) category in the fitting region that is determined in the signal extraction fit,

$$\vec{x} = (M_{bc}, \Delta E, \mathcal{O}_{CS}, E^*_{beam}, \cos(\theta_B^{Boost}), |q_{tag} \cdot r|), \tag{4.6}$$

$$P_{sig}(\vec{x}) = P_{sig}^{M_{bc}}(M_{bc}) \cdot P_{sig}^{\Delta E}(\Delta E) \cdot P_{sig}^{\mathcal{O}_{CS}}(\mathcal{O}_{CS}) \cdot P_{sig}^{|q_{tag} \cdot r|}(|q_{tag} \cdot r|) \cdot P_{sig}^{\cos(\theta_B^{Boost})}(\cos(\theta_B^{Boost})), \text{ and} \tag{4.7}$$

$$P_{bckg}(\vec{x}) = P_{bckg}^{M_{bc}}(M_{bc}|E^*_{beam}) \cdot P_{bckg}^{\Delta E}(\Delta E) \cdot P_{bckg}^{\mathcal{O}_{CS}}(\mathcal{O}_{CS}) \cdot P_{bckg}^{|q_{tag} \cdot r|}(|q_{tag} \cdot r|) \cdot P_{bckg}^{\cos(\theta_B^{Boost})}(\cos(\theta_B^{Boost})). \tag{4.8}$$

Detailed procedures for signal extraction, the definition of each function, and a description of each variable will be discussed in this chapter. We used `RooFit` library[52] included in the ROOT framework[53] as a statistical analysis software and `MINUIT` as the backend library for minimization[54].

4.5.1 Fitting Procedure

We described a distribution of $(M_{bc}, \Delta E, \mathcal{O}_{CS})$ using the following PDF:

$$\begin{aligned} P(M_{bc}, \Delta E, \mathcal{O}_{CS}|E^*_{beam}) =& F_{sig}P_{sig}^{M_{bc}}(M_{bc})P_{sig}^{\Delta E}(\Delta E)P_{sig}^{\mathcal{O}_{CS}}(\mathcal{O}_{CS}) + \\ & (1 - F_{sig})P_{bckg}^{M_{bc}}(M_{bc}|E^*_{beam})P_{bckg}^{\Delta E}(\Delta E)P_{bckg}^{\mathcal{O}_{CS}}(\mathcal{O}_{CS}) \end{aligned} \tag{4.9}$$

where $P^x_{sig,bckg}(x)$ is the corresponding function for variable x for (Signal + SxF) and Background events, and E^*_{beam} is the measured collision energy which changes per each run. With the PDF above, we estimated the yield of (Signal + SxF) and Background category and the shape of PDF for the Background category by maximizing the following extended likelihood function with a floating parameter of F_{sig} in an unbinned fashion as follows:

$$\mathcal{L}(M_{bc}, \Delta E, \mathcal{O}_{CS}|E^*_{beam}) = \frac{e^{-N_{total}}}{n!} \cdot \prod_i^n N_{total} \cdot P(M_{bc}, \Delta E, \mathcal{O}_{CS}|E^*_{beam}) \tag{4.10}$$

where n is the number of events in the dataset to be fitted, and N_{total} is the estimated number of total events within the fitting region given in Eq. 4.11.

We assumed no correlation between the five variables in the signal extraction. However, we confirmed that there was a weak correlation between some variables. We will take into account the effects of ignoring such correlation on the *CP* parameter estimation in the study of systematic uncertainties. Detail of the correlation is discussed in App. B.

The signal extraction fitting has been performed within the following region.

$$(5.2 < M_{bc} < 5.3)\,\mathrm{GeV}/c^2 \wedge |\Delta E| < 0.25\,\mathrm{GeV} \wedge -4 < \mathcal{O}_{CS} < 4. \tag{4.11}$$

We call this region a fitting region (FR).

We divided the fitting region into the signal region (SR) and the sideband region (SB). The signal region is defined as

$$(5.27 < M_{bc} < 5.29)\,\mathrm{GeV}/c^2 \wedge |\Delta E| < 0.1\,\mathrm{GeV} \tag{4.12}$$

and we used events only in this region for the fitting using Δt and q_{tag} information such as lifetime determination or *CP* parameter extraction.

The other one is the sideband (SB) region. There are two types of SB regions which are M_{bc} and ΔE as follows:

- M_{bc}-type: $5.2 < M_{bc} < 5.26\,\mathrm{GeV}/c^2$
- ΔE-type: $|\Delta E| > 0.1\,\mathrm{GeV}$

each type of sideband region is used to determine the parameters of corresponding background PDFs, which will be discussed later.

We checked the homogeneity in the MC samples between sideband events and background events in the signal region using MC truth information to confirm that the sideband events can represent all background events used for *CP* fitting. Comparison results shown in Figs. 4.13, 4.14, and 4.15 indicate there are no significant differences between the two populations both for main ($B^0 \to \eta' K^0_S$) and control ($B^\pm \to \eta' K^\pm$) channel. In this comparison, we did not apply selection criteria to K^0_S and π^0 shown in Tab. E.1 to obtain enough statistics for the comparison.

Thus, the procedures for the signal extraction are,

1. determine the parameters of pre-fit PDFs for the (Signal + SxF) events using signal MC,

2. determine the parameters of pre-fit PDFs for the Background events using MC-truth background events for MC samples or the sideband events for the experimental data,

3. performs the unbinned maximum likelihood fitting to target MC sample(1/ab) or experimental data(362/fb) to extract N_{total} and F_{sig} in each dataset. In this step, we also set some parameters for the Background PDF of M_{bc}, ΔE, and $\mathcal{O}_{CS}$ to floating parameters. The result of the pre-fit for the Background PDF will be used as an initial value for this step.

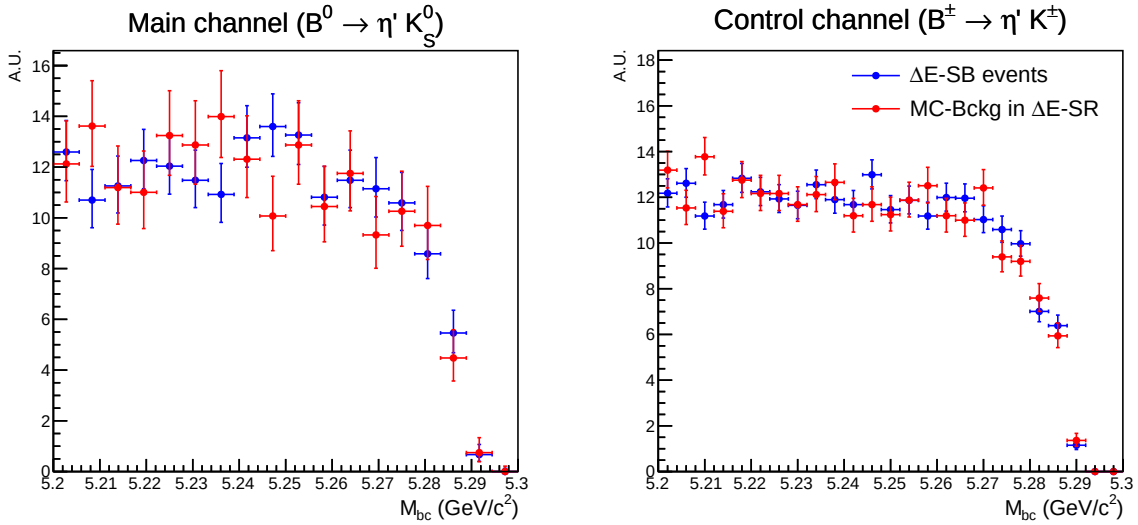


Figure 4.13: Comparison of M_{bc} distribution between events from the ΔE-sideband region(only $0.1 < |\Delta E| < 0.25\,\text{GeV}$ selection applied) and background events in the ΔE-signal region(only $|\Delta E| < 0.1\,\text{GeV}$ selection applied) from MC samples. "MC-Bckg" indicates the true-Background events picked using the MC information.

The procedure is performed independently in $B^0 \to \eta' K^0_S$ and for $B^\pm \to \eta' K^\pm$.

We used MIGRAD functionality to minimize negative-log-likelihood from the above PDF for signal extraction. We applied HESSE and MINOS techniques after MIGRAD to estimate the proper uncertainty of determined yields under the RooFit framework.

4.6 Pre-fit PDFs

We define the pre-fit PDFs, which describe the shape of variables for signal extraction in the following subsections. PDFs for the main channel and control channel are the same.

4.6.1 Signal + SxF events

The pre-fit PDFs for Signal + SxF events are defined as in Eq. 4.13, 4.14, and 4.15. Each function has 7, 6, and 6 parameters, respectively.

$$\begin{aligned} P_{sig}^{M_{bc}}(M_{bc}) =& f_{CB}^{sig} \cdot \text{CrystalBall}(M_{bc}; \mu_{CB}^{sig}, \sigma_{CB}^{sig}, \alpha_{CB}^{sig}, N_{CB}^{sig}) \\ &+ (1 - f_{CB}^{sig}) \cdot \text{Gaus}(M_{bc}; \mu_{G_{M_{bc}}}^{sig}, \sigma_{G_{M_{bc}}}^{sig}) \end{aligned} \tag{4.13}$$

$$\begin{aligned} P_{sig}^{\Delta E}(\Delta E) =& f_{Voigt}^{sig} \cdot \text{Voigt}(\Delta E; \mu_{Voigt}^{sig}, \sigma_{Voigt}^{sig}, w_{Voigt}^{sig}) \\ &+ (1 - f_{Voigt}^{sig}) \cdot \text{Gaus}(\Delta E; \mu_{G_{\Delta E}}^{sig}, \sigma_{G_{\Delta E}}^{sig}) \end{aligned} \tag{4.14}$$

$$\begin{aligned} P_{sig}^{\mathcal{O}_{CS}}(\mathcal{O}_{CS}) =& f_{BifG}^{sig} \cdot \text{BifGaus}(\mathcal{O}_{CS}; \mu_{BifG}^{sig}, \sigma_{Left}^{sig}, \sigma_{Right}^{sig}) \\ &+ (1 - f_{BifG}^{sig}) \cdot \text{Gaus}(\mathcal{O}_{CS}; \mu_{G_{\mathcal{O}_{CS}}}^{sig}, \sigma_{G_{\mathcal{O}_{CS}}}^{sig}) \end{aligned} \tag{4.15}$$

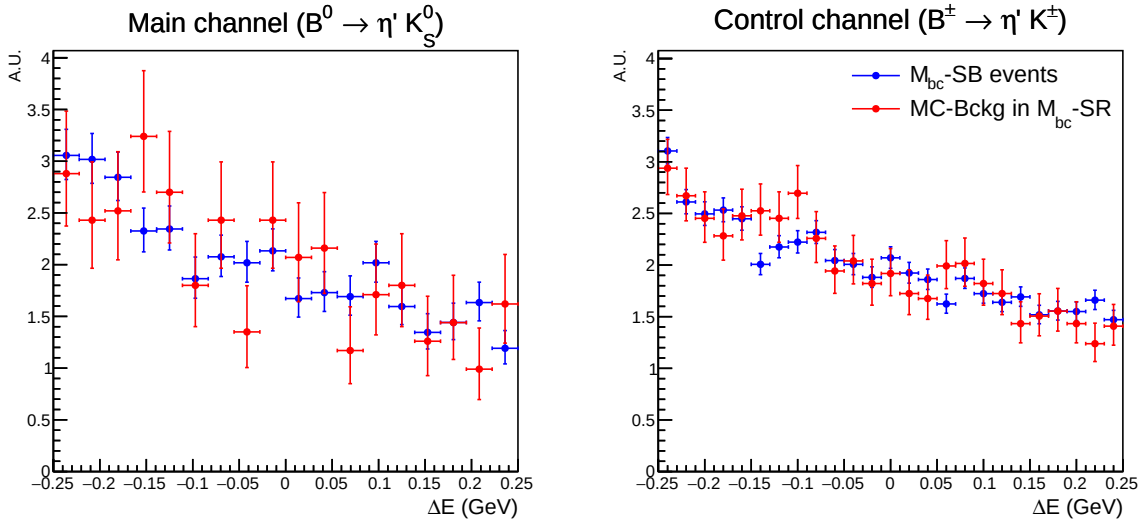


Figure 4.14: Comparison of M_{bc} distribution between events from the M_{bc}-sideband region(only $5.2 < M_{bc} < 5.26\,\mathrm{GeV}/c^2$ selection applied) and background events in the M_{bc}-signal region(only $5.27 < M_{bc} < 5.29\,\mathrm{GeV}/c^2$ selection applied) from MC samples. "MC-Bckg" indicated the true-Background events picked using the MC information.

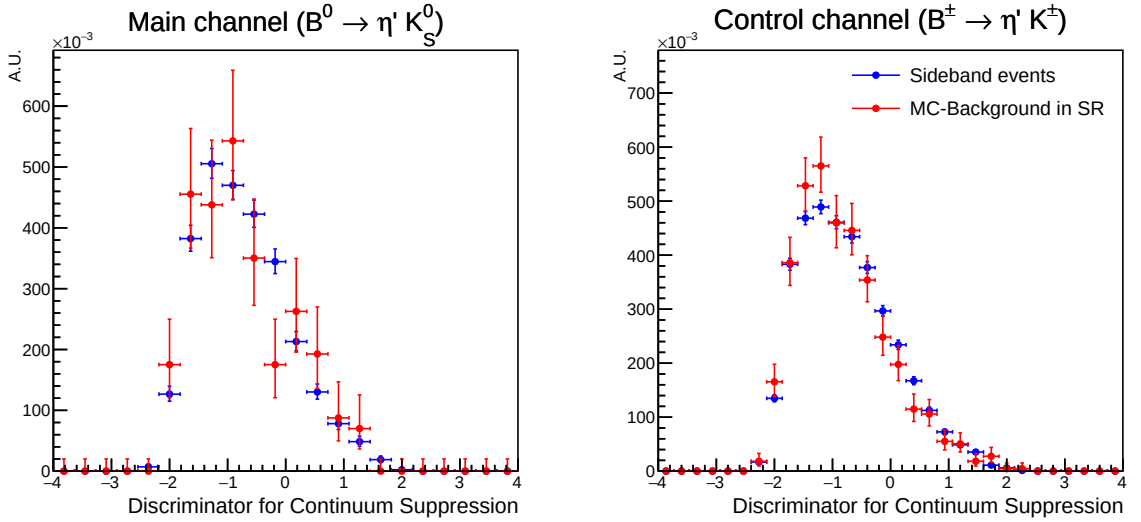


Figure 4.15: Comparison of $\mathcal{O}_{CS}$ distribution between events from the sideband region and background events in the signal region from MC samples. "MC-Background" indicates the true-Background events picked using the MC information.

and the definition of special functions that were used for this parameterization is as follows:

$$\text{Gaus}(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4.16}$$

$$\text{Voigt}(x;\mu,\sigma,w) = \int_{-\infty}^{\infty}\text{Gaus}(x';\sigma)\cdot\frac{w}{\pi((x-x')^2+w^2)}dx' \tag{4.17}$$

$$\text{BifGaus}(x;\mu,\sigma_L,\sigma_R) = \mathcal{N}\cdot\begin{cases}\text{Gaus}(x;\mu,\sigma_L) & \text{if} \quad x<\mu\\ \text{Gaus}(x;\mu,\sigma_R) & \text{if} \quad x\geq\mu\end{cases} \tag{4.18}$$

$$\text{CrystalBall}(x;\mu,\sigma,\alpha,N) = \mathcal{N}\cdot\begin{cases}\left[\left(\frac{N}{|\alpha|}\right)\left(\frac{N-|\alpha|^2}{|\alpha|}+\frac{x-\mu}{\sigma}\right)\right]^N e^{-\frac{|\alpha|^2}{2}} & \text{if } \frac{x-\mu}{\sigma}<-\alpha\\ e^{-\frac{(x-\mu)^2}{2\sigma^2}} & \text{if } \frac{x-\mu}{\sigma}\geq-\alpha\end{cases} \tag{4.19}$$

where $\mathcal{N}$ means the normalization constant for each distribution.

We determined all parameters of pre-fit PDFs for the (Signal + SxF) category $P^x_{sig}(x)$ with the unbinned ML fit to the entire signal samples. Determined parameters from the signal sample are shown in Tab. 4.4, and curves for M_{bc}, ΔE, $\mathcal{O}_{CS}$, and $\cos(\theta_B^{Boost})$ are shown in Figs. 4.17, 4.20, 4.23, and 4.26, respectively, which shows our modeling can describe MC points from signal MC samples well. Also, the integration of $P_{sig}(M_{bc},\Delta E,\mathcal{O}_{CS}) = P^{M_{bc}}_{sig}(M_{bc})P^{\Delta E}_{sig}(\Delta E)P^{\mathcal{O}_{CS}}_{sig}(\mathcal{O}_{CS})$ over our signal region defined in Eq. 4.12 is shown in "Signal PDF" row in Tab. 4.5.

For the $\cos(\theta_B^{Boost})$ variable, we determined the pre-fit PDF using the signal samples rather than using the theoretical PDF given in Eq. 4.4. We are using the angle between the momentum of B^0 and the boost axis, which slightly differs from the z-axis, and there are SxF events that have contamination from the tag side; the distribution of $\cos(\theta_B^{Boost})$ changes from the theoretical one. Therefore, we did not use the theoretical PDF as is; we used a PDF of $P^{\cos(\theta_B^{Boost})}_{sig}(\cos(\theta_B^{Boost})) = \mathcal{N}\cdot(1-\alpha^{sig}\cdot(\cos(\theta_B^{Boost}))^2)$ and determined α with the same fashion as that for $M_{bc}, \Delta E$ and $\mathcal{O}_{CS}$.

For the dilution factor, we used a histogram-based PDF to calculate the signal probability per event. Blue colored distribution in Fig. 4.16 shows the PDF for the (Signal + SxF) category. This PDF has been determined by counting the number of events in each bin using the signal MC samples.

4.6.2 Background events

The pre-fit PDFs for Background events are defined as in Eq. 4.20, 4.21, and 4.22. Each function has 1, 1, and 3 parameters, respectively.

$$P^{M_{bc}}_{bckg}(M_{bc}|E^*_{beam}) = \text{ARGUS}(M_{bc};E^*_{beam}/2,c^{bckg},p^{bckg}) \tag{4.20}$$

$$P^{\Delta E}_{bckg}(\Delta E) = \text{Pol}^1(\Delta E;a^{bckg}) \tag{4.21}$$

$$P^{\mathcal{O}_{CS}}_{bckg}(\mathcal{O}_{CS}) = \text{BifGaus}(\mathcal{O}_{CS};\mu^{bckg}_{BifG},\sigma^{bckg}_{Left},\sigma^{bckg}_{Right}) \tag{4.22}$$

Function	Parameter	Value for main channel	Value for control channel
$P_{sig}^{M_{bc}}(M_{bc})$	f_{CB}^{sig}	$0.294^{+0.013}_{-0.013}$	$0.321^{+0.014}_{-0.013}$
	μ_{CB}^{sig}	$5.28122^{+0.00009}_{-0.00009}$	$5.28070^{+0.00009}_{-0.00009}$
	σ_{CB}^{sig}	$0.00297^{+0.00004}_{-0.00004}$	$0.00307^{+0.00005}_{-0.00004}$
	α_{CB}^{sig}	$0.633^{+0.029}_{-0.027}$	$0.710^{+0.029}_{-0.027}$
	N_{CB}^{sig}	$7.50^{+0.67}_{-0.57}$	$6.35^{+0.42}_{-0.38}$
	$\mu_{G_{M_{bc}}}^{sig}$	$5.27943^{+0.00002}_{-0.00003}$	$5.27912^{+0.00003}_{-0.00003}$
	$\sigma_{G_{M_{bc}}}^{sig}$	$0.00246^{+0.00002}_{-0.00002}$	$0.00250^{+0.00002}_{-0.00002}$
$P_{sig}^{\Delta E}(\Delta E)$	f_{Voigt}^{sig}	$0.8880^{+0.0018}_{-0.0018}$	$0.8702^{+0.0019}_{-0.0020}$
	μ_{Voigt}^{sig}	$-0.00168^{+0.00003}_{-0.00003}$	$-0.00167^{+0.00003}_{-0.00003}$
	$\sigma Voigt^{sig}$	$0.00793^{+0.00007}_{-0.00007}$	$0.00988^{+0.00008}_{-0.00008}$
	w_{Voigt}^{sig}	$0.01035^{+0.00015}_{-0.00015}$	$0.00972^{+0.00018}_{-0.00018}$
	$\mu_{G_{\Delta E}}^{sig}$	$-0.0950^{+0.0025}_{-0.0026}$	$-0.0916^{+0.0022}_{-0.0023}$
	$\sigma_{G_{\Delta E}}^{sig}$	$0.1354^{+0.0017}_{-0.0017}$	$0.1346^{+0.0015}_{-0.0014}$
$P_{sig}^{\mathcal{O}_{CS}}(\mathcal{O}_{CS})$	f_{BifG}^{sig}	$0.5800^{+0.0087}_{-0.0090}$	$0.5857^{+0.0082}_{-0.0086}$
	μ_{BifG}^{sig}	$0.768^{+0.016}_{-0.016}$	$0.806^{+0.016}_{-0.016}$
	σ_{Left}^{sig}	$0.9444^{+0.0067}_{-0.0066}$	$0.9795^{+0.0068}_{-0.0066}$
	σ_{Right}^{sig}	$0.498^{+0.015}_{-0.016}$	$0.486^{+0.015}_{-0.016}$
	$\mu_{G_{\mathcal{O}_{CS}}}^{sig}$	$1.6404^{+0.0053}_{-0.0054}$	$1.6649^{+0.0054}_{-0.0054}$
	$\sigma_{G_{\mathcal{O}_{CS}}}^{sig}$	$0.3606^{+0.0024}_{-0.0024}$	$0.3578^{+0.0024}_{-0.0024}$
$P_{sig}^{\cos(\theta_B^{Boost})}(\cos(\theta_B^{Boost}))$	α^{sig}	$-0.9469^{+0.0020}_{-0.0019}$	$-0.9452^{+0.0019}_{-0.0018}$

Table 4.4: List of parameters of pre-fit PDFs for (Signal + SxF) events.

and the definition of special functions that were used for this parameterization is as follows (Definition of BifGaus is in Eq. 4.18.):

$$\mathrm{ARGUS}(x; x_0, c) = \mathcal{N} \cdot x \cdot \sqrt{1 - \left(\frac{x}{x_0}\right)^2} \cdot e^{c \cdot \left(1 - (x/x_0)^2\right)} \tag{4.23}$$

$$\mathrm{Pol}^1(x; a) = \mathcal{N} \cdot (1 + a \cdot x) \tag{4.24}$$

where $\mathcal{N}$ means the normalization constant for each distribution.

We determined all five parameters of pre-fit PDFs for the Background category similarly to that for the (Signal + SxF) category. These parameters will be set to a floating parameter in the next step (The signal extraction fit) so that the determined values for parameters will be used as an initial value. In detail, the initial value for the parameters in the MC study is determined by fitting P_{bckg} to true-Background category events using MC truth events. For the experimental data study, sideband events are used in the fitting of P_{bckg}. To illustrate, the shape of $P_{bckg}^{M_{bc}}$ is determined by using ΔE-type sideband and vice versa for $P_{bckg}^{\Delta E}$. For $P_{bckg}^{\mathcal{O}_{CS}}$, we used side-band events both of two types. As a result, fitted curves and MC (or Data) points are illustrated in Fig. 4.18, 4.21, and 4.24

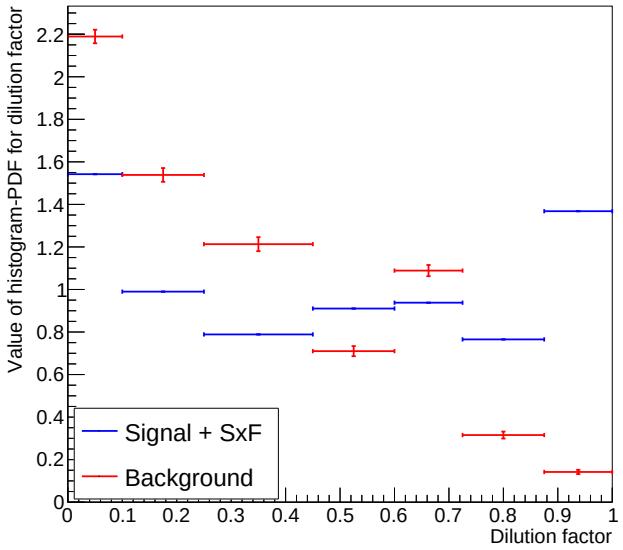


Figure 4.16: Histogram-based PDF $P_{sig,bckg}^{|q_{tag}\cdot r|}(|q_{tag}\cdot r|)$ for (Signal + SxF) and Background categories.

for MC samples and Fig. 4.19, 4.22, and 4.25 for the experimental data, which show good agreement with MC (Data) points, and our modeling can describe the distributions of M_{bc}, ΔE, and $\mathcal{O}_{CS}$ variables well. In addition, we calculated the integration of $P_{bckg}(M_{bc}, \Delta E, \mathcal{O}_{CS})$ over the signal region, which is given in Tab. 4.5. The integration of Background PDF is smaller than that of Signal PDF, which is due to the shape difference between the two PDFs and enables us to determine the amount of each category in the fitting.

We used a constant PDF of $P_{bckg}^{\cos(\theta_B^{Boost})}(\cos(\theta_B^{Boost})) = 0.5$ to model the distribution of $\cos(\theta_B^{Boost})$ for the Background category to avoid the Punzi effect, and Figs. 4.27 and 4.28 show the PDF can model the distribution well. Also, we consider the dilution factor by a histogram-based PDF shown in the red-colored histogram of Fig. 4.16. This distribution was determined using the information on sideband events of experimental data.

PDF & Parameter set	Main	Control
Signal PDF	90.02%	89.15%
Background PDF for MC study	6.11%	6.06%
Background PDF for Data	5.58%	5.30%

Table 4.5: Result of the integral over the signal region for PDFs for each category. Also, parameters of Background PDFs after the signal extraction are used for the integration.

4.7 Signal extraction fit

We performed a 3D unbinned ML fit for signal extraction using PDF shown in Eq. 4.9. Before extracting the yield of events in the (Signal + SxF) category, we performed the ensemble test using the Toy MC generator to check the validity of the fitting procedure. The result indicates no problem in the signal extraction procedure. In addition, we investigated bias from correlation by generating $M_{bc} - \Delta E$ correlated toy samples. Detailed results for the ensemble test and correlations are discussed in App. B. There is no significant bias due to the correlation, so we will consider the effects of this correlation on the *CP* parameter estimation in systematic error studies.

Then, we used a 1/ab dataset from MC15 generic sample for the MC study, and 362/fb experimental data was used for the yield estimation. We determined seven parameters in this step. N_{total}, F_{sig} are the number of total candidates from the reconstruction and the overall fraction of events in the (Signal + SxF) category within the fitting region, respectively. Remaining 5 parameters are $c^{bckg}, a^{bckg}, \mu^{bckg}_{BifG}, \sigma^{bckg}_{Left}, \sigma^{bckg}_{Right}$ which are the shape parameters for the background PDFs. Determined parameters are shown in Tab. 4.6(MC) and 4.7(Data). Fitted curves and MC or Data points are given in Fig. 4.29, 4.30, and 4.31 from the MC study, and Fig. 4.32, 4.33, and 4.34 for experimental data.

Parameter	Value for main channel	Value for control channel
N_{total}	2820^{+53}_{-53}	14007^{+119}_{-118}
F_{sig}	$0.0907^{+0.0060}_{-0.0058}$	$0.0623^{+0.0023}_{-0.0023}$
c^{bckg}	$-25.8^{+2.2}_{-2.2}$	$-25.74^{+0.98}_{-0.98}$
a^{bckg}	$-0.377^{+0.032}_{-0.032}$	$-0.325^{+0.014}_{-0.014}$
μ^{bckg}_{BifG}	$-1.419^{+0.036}_{-0.037}$	$-1.493^{+0.016}_{-0.016}$
σ^{bckg}_{Left}	$0.322^{+0.022}_{-0.022}$	$0.2945^{+0.0095}_{-0.0093}$
σ^{bckg}_{Right}	$1.159^{+0.028}_{-0.027}$	$1.208^{+0.012}_{-0.012}$

Table 4.6: List of parameters from signal extraction for generic MC 1/ab sample.

Parameter	Value for main channel	Value for control channel
N_{total}	816 ± 29	4053 ± 64
F_{sig}	0.075 ± 0.010	0.0610 ± 0.0043
c^{bckg}	-25.5 ± 4.2	-21.8 ± 1.9
a^{bckg}	-0.307 ± 0.061	-0.361 ± 0.027
μ^{bckg}_{BifG}	-1.141 ± 0.071	-1.221 ± 0.032
σ^{bckg}_{Left}	0.478 ± 0.043	0.454 ± 0.019
σ^{bckg}_{Right}	1.058 ± 0.050	1.107 ± 0.023

Table 4.7: List of parameters from signal extraction for experimental data. We didn't apply MINOS in this estimation

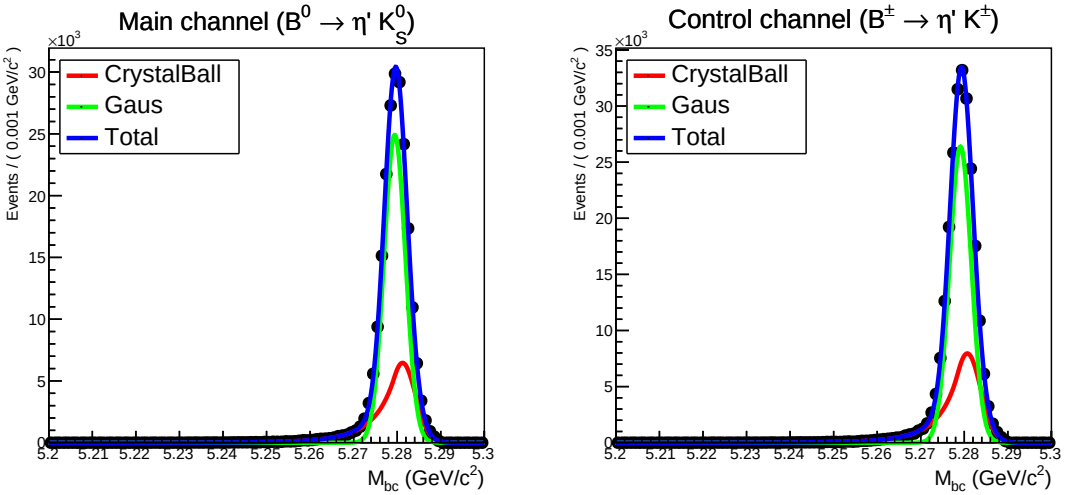


Figure 4.17: Pre-fit result of M_{bc} distribution for (Signal + SxF) category from signal MC. The blue curve represents $P_{sig}^{M_{bc}}$, and the green and red one shows Gaussian and Crystal Ball components.

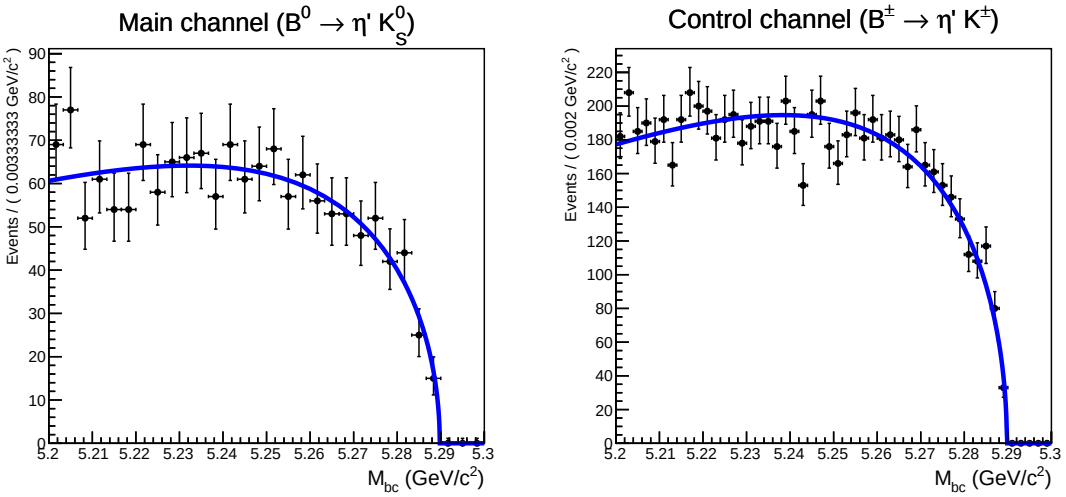


Figure 4.18: Pre-fit Result of M_{bc} distribution for Background category from generic MC samples. MC-truth information was used to pick Background events.

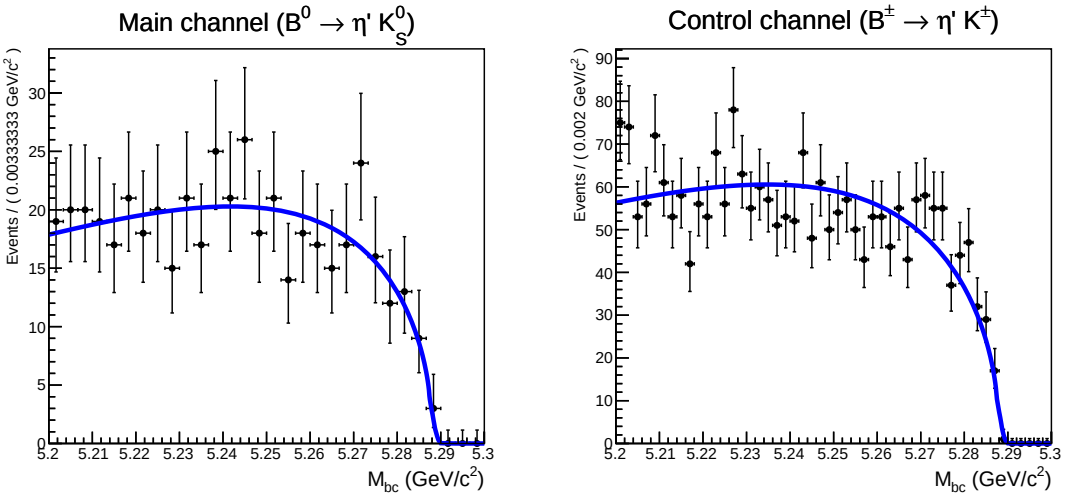


Figure 4.19: Pre-fit Result of M_{bc} distribution for Background category from ΔE-type SB events.

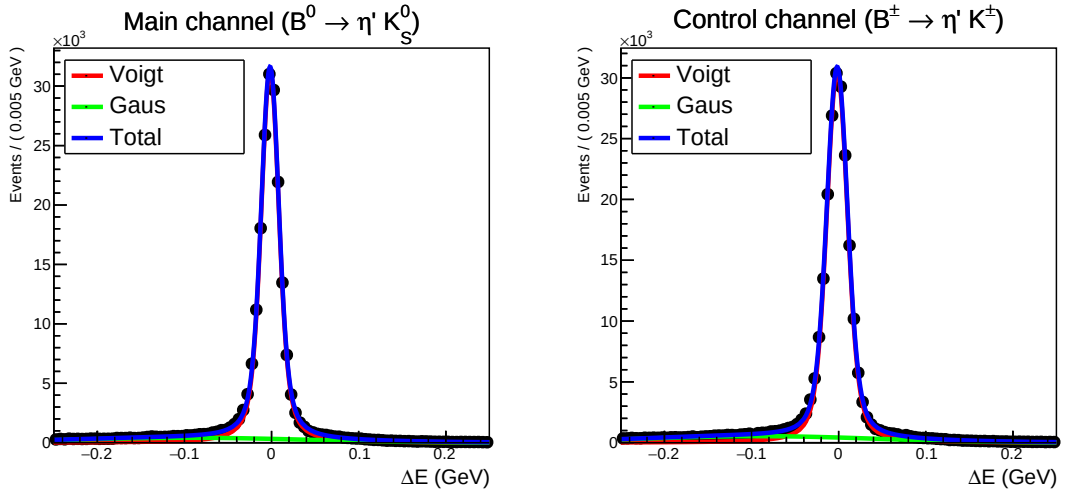


Figure 4.20: Pre-fit result of ΔE distribution for (Signal + SxF) category from signal MC. The blue curve represents $P_{sig}^{\Delta E}$, and the green and red one shows Gaussian and Voigt profile components.

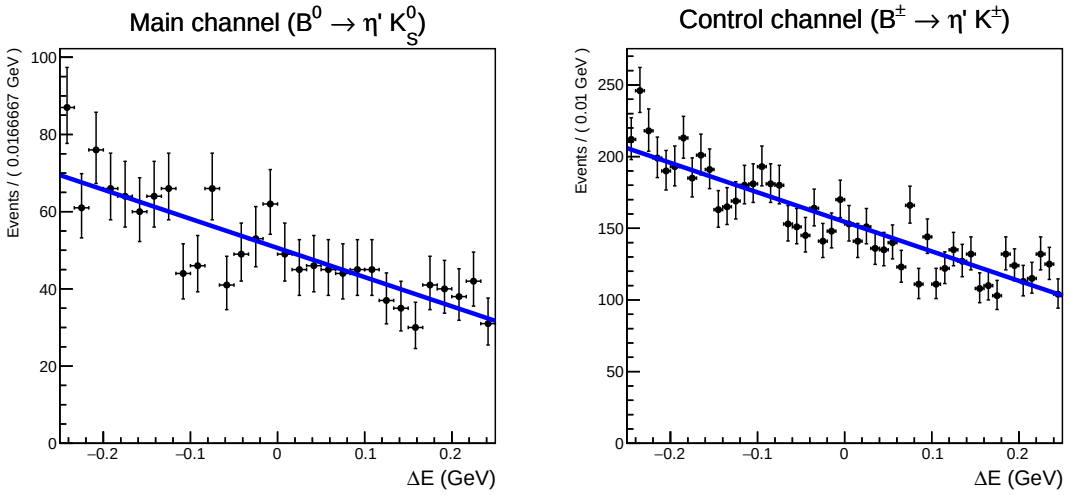


Figure 4.21: Pre-fit Result of ΔE distribution for Background category from generic MC samples. MC-truth information was used to pick Background events.

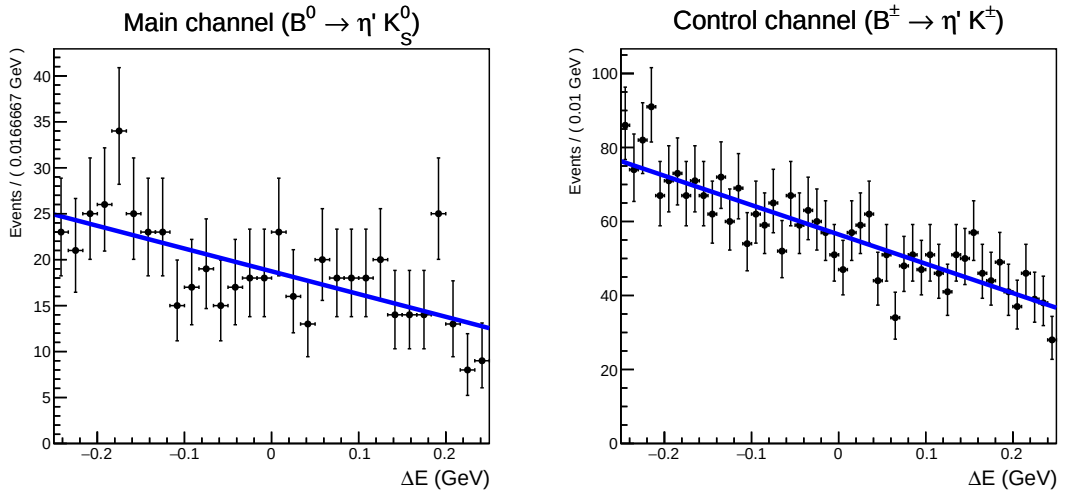


Figure 4.22: Pre-fit Result of ΔE distribution for Background category from M_{bc}-type SB events.

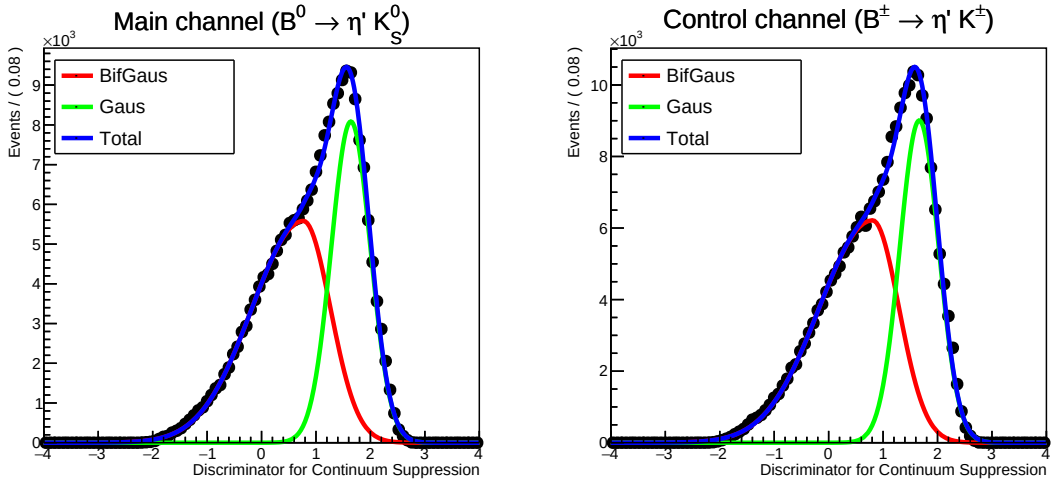


Figure 4.23: Pre-fit result of $\mathcal{O}_{CS}$ distribution for (Signal + SxF) category from signal MC. The blue curve represents $P_{sig}^{\mathcal{O}_{CS}}$, and the green and red one shows Gaussian and Bifurcated Gaussian.

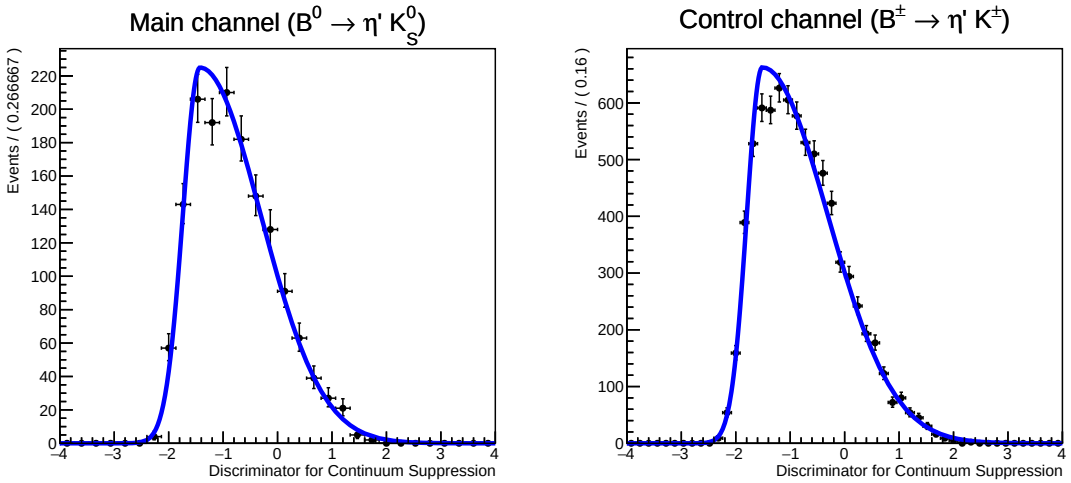


Figure 4.24: Pre-fit Result of $\mathcal{O}_{CS}$ distribution for Background category from generic MC samples. MC-truth information was used to pick Background events.

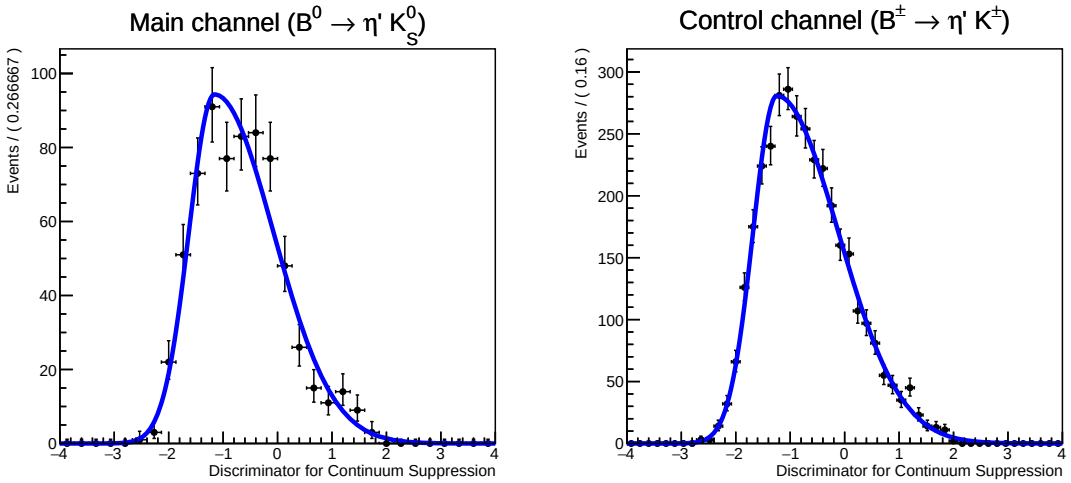


Figure 4.25: Pre-fit result of $\mathcal{O}_{CS}$ distribution for Background category from SB events.

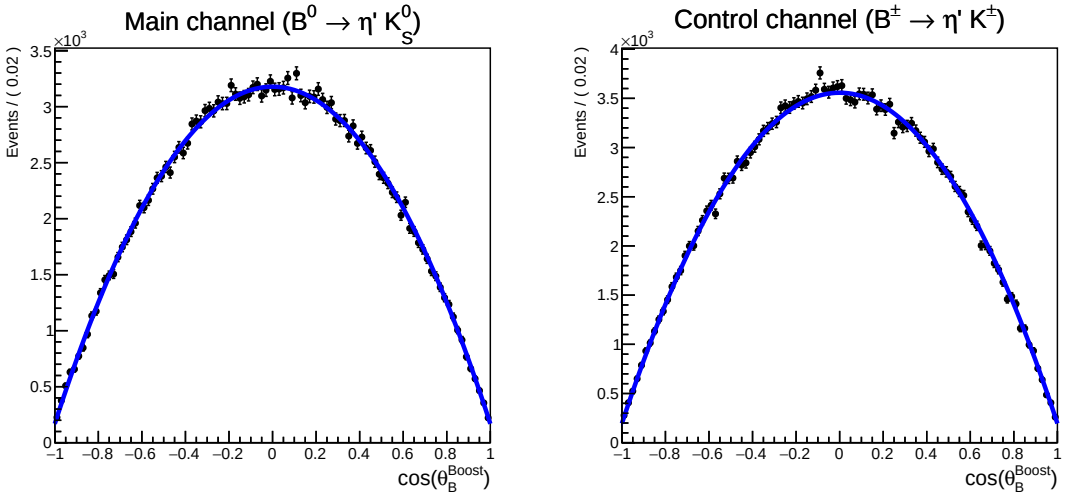


Figure 4.26: Pre-fit result of $\cos(\theta_B^{Boost})$ distribution for (Signal + SxF) category from signal MC. The blue curve represents $P_{sig}^{\cos(\theta_B^{Boost})}$.

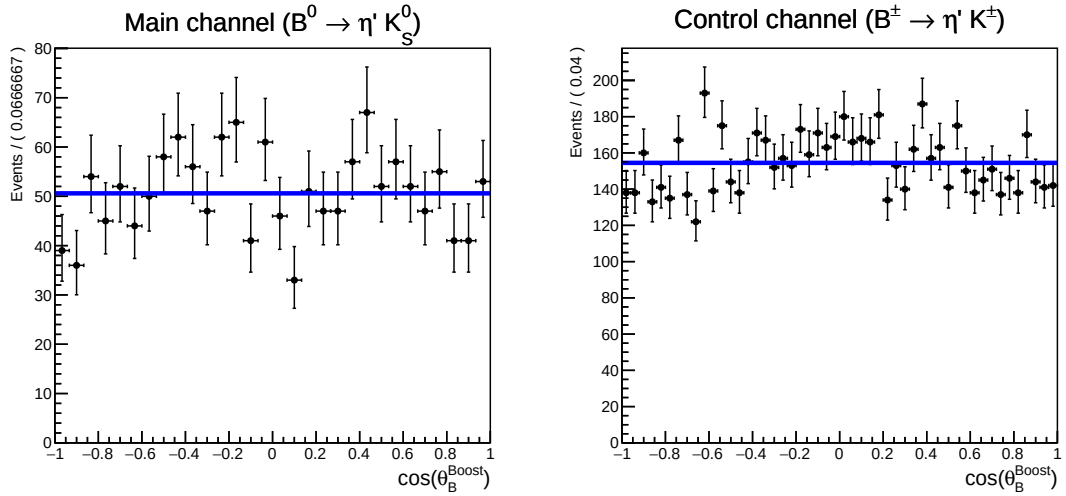


Figure 4.27: Comparison between theoretical PDF (constant) and MC points for Background category from generic MC sample. MC-truth information was used to pick Background events.

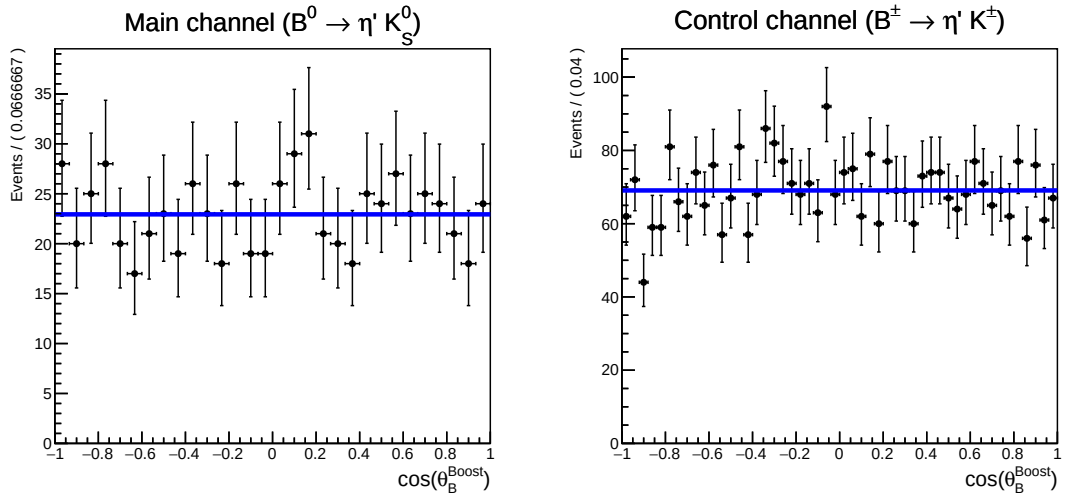


Figure 4.28: Comparison between theoretical PDF and Data points from SB events.

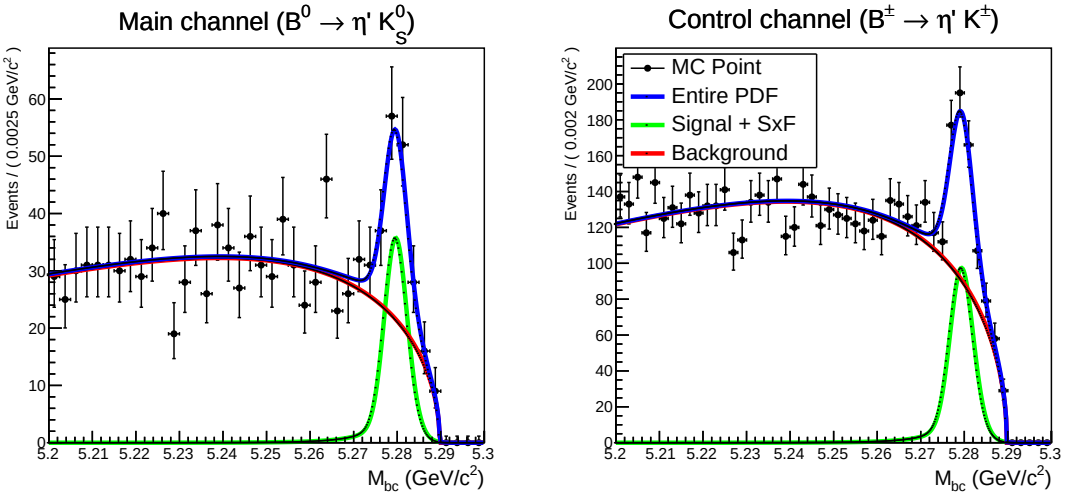


Figure 4.29: Projection on M_{bc} axis of the fitted result from the signal extraction fit for MC sample

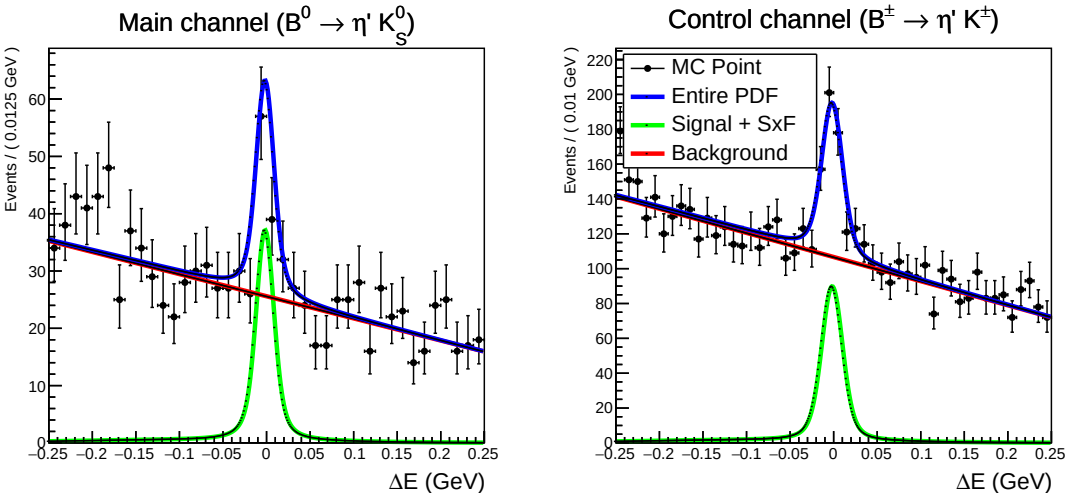


Figure 4.30: Projection on ΔE axis of the fitted result from the signal extraction fit for MC sample

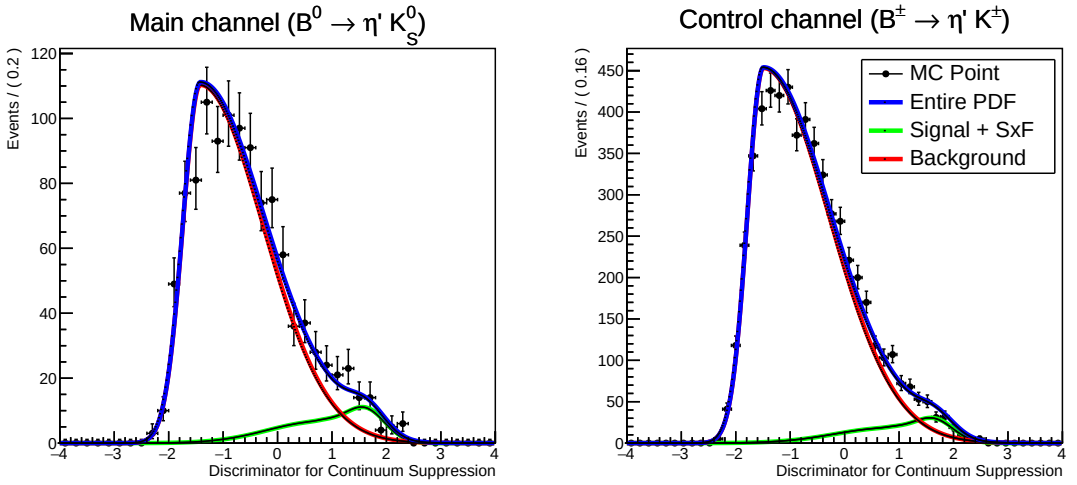


Figure 4.31: Projection on $\mathcal{O}_{CS}$ axis of the fitted result from the signal extraction fit for MC sample

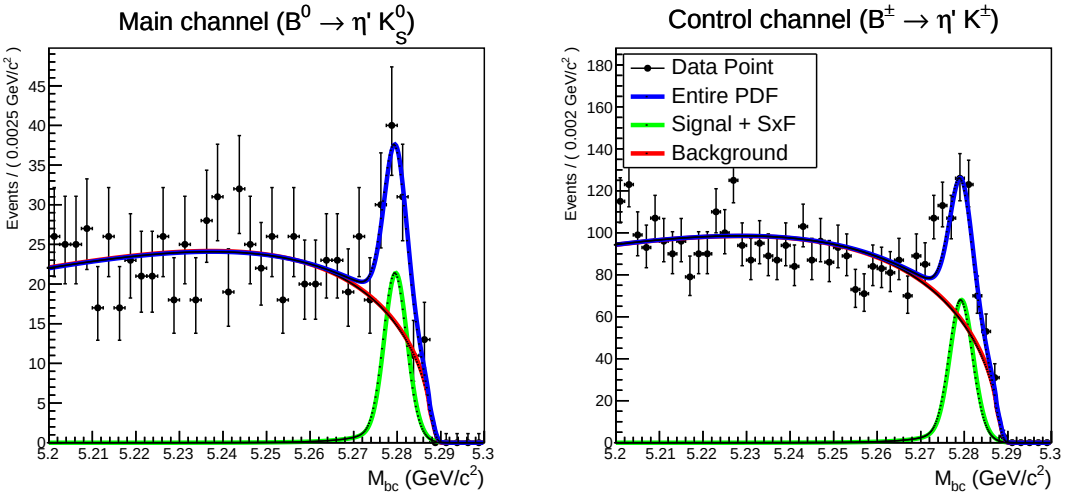


Figure 4.32: Projection on M_{bc} axis of the fitted result from the signal extraction fit for the experimental data

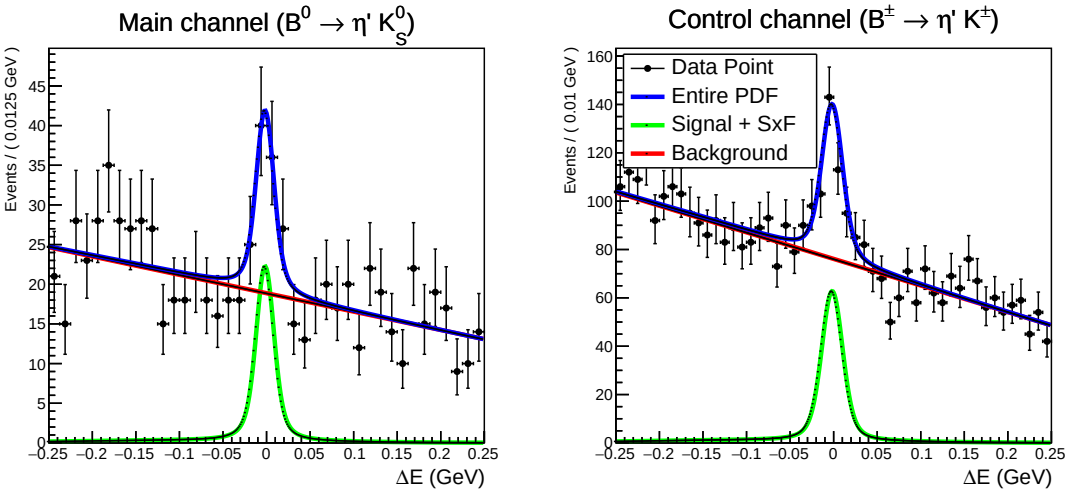


Figure 4.33: Projection on ΔE axis of the fitted result from the signal extraction fit for the experimental data

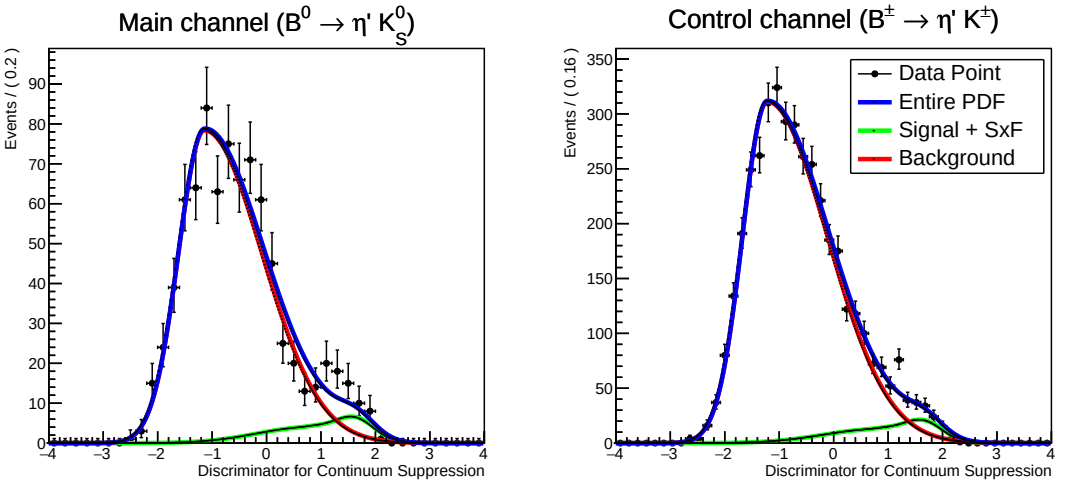


Figure 4.34: Projection on $\mathcal{O}_{CS}$ axis of the fitted result from the signal extraction fit for the experimental data

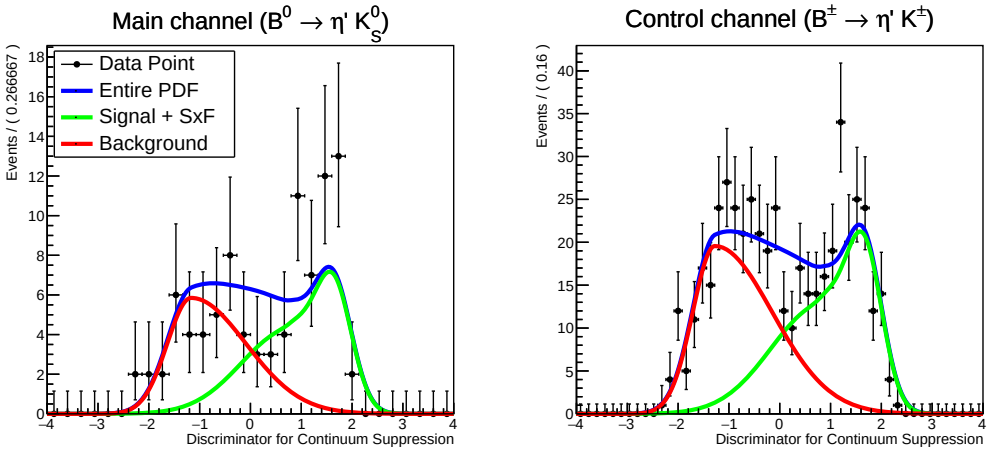


Figure 4.35: Projection on $\mathcal{O}_{CS}$ axis of the fitted result within the signal region. (Experimental data)

4.8 Results for the signal extraction fit

In this section, we compare the actual yield of the (Signal + SxF) category and expectation values calculated using known information to check the validity of our procedure for the signal extraction fit. First, we convert N_{total} and F_{sig} to the signal yield N_{sig}. We confirmed there is no correlation between N_{total} and F_{sig} (< 0.0001) from the result of the signal extraction fitting shown in Tab. 4.7 so that we can calculate the yield of signal events by calculating a product of two parameters and the value is $N_{sig} = 61.2 \pm 8.6$ for the main channel and 247 ± 18 for the control channel.

Next, we calculate the expectation yield for the $B^0 \to \eta' K^0_S$ channel with our subchannel from the number of $B\overline{B}$ pair, the partial branching ratio, reconstruction efficiency, and correction factor for the difference of the reconstruction efficiency between MC and experimental data. The number of $B\overline{B}$ pair from the full dataset is $(387 \pm 6) \times 10^6$, and the branching ratio can be calculated as[11]

- $\mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0) = (4.86 \pm 0.06) \times 10^{-1}$,
- $\mathcal{B}(B^0 \to \eta' K^0) = (6.6 \pm 0.4) \times 10^{-5}$,
- $\mathcal{B}(K^0 \to K^0_S) = 0.5$,
- $\mathcal{B}(K^0_S \to \pi^+\pi^-) = (6.920 \pm 0.005) \times 10^{-1}$,
- $\mathcal{B}(\eta' \to \eta\pi^+\pi^-) = (4.25 \pm 0.05) \times 10^{-1}$,
- $\mathcal{B}(\eta \to \pi^+\pi^-\pi^0) = (2.302 \pm 0.025) \times 10^{-1}$, and
- $\mathcal{B}(\pi^0 \to \gamma\gamma) = (9.8823 \pm 0.0034) \times 10^{-1}$.
- For the control channel, $\mathcal{B}(\Upsilon(4S) \to B^+B^-) = (5.14 \pm 0.06) \times 10^{-1}$, and
- $\mathcal{B}(B^\pm \to \eta' K^\pm) = (7.04 \pm 0.25) \times 10^{-5}$.
- For both channels, B-mesons are produced in pairs, so a factor of 2 is multiplied.

Therefore, the branching ratio for the main and control channel is $(2.15 \pm 0.17) \times 10^{-6}$ and $(7.00 \pm 0.29) \times 10^{-6}$, respectively. Then, we should consider the signal efficiency for the

main and control channels, which is 10.88% and 12.02%. We can get the expected yield by multiplying 0.8436 ± 0.0488, which is the ratio of reconstruction efficiency between MC and Data for the main channel, and for the control channel, 0.8570 ± 0.0691 was multiplied, which is described in App. D. The computation results and comparison with the fitting results are shown in Tab. 4.8, which indicates the result are insignificantly smaller than the expectation by 1.31σ for the main channel and 1.01σ for the control channel. We did not dig into the reason for this discrepancy since our target is *CP* asymmetry studies, which does not be affected much by the signal efficiencies, and the systematic uncertainty by PDFs for the signal extraction is not large. (See Sec. 6.3)

Channel	Fit result	Expectation
Main	61.2 ± 8.6	76.2 ± 7.5
Control	247 ± 18	279 ± 26

Table 4.8: Expected and estimated signal yield for the main and control channel.

Chapter 5

Δt Functions for the *CP* Fitting

This chapter discusses the components for *CP* fitting using Δt distribution in the main channel. We need two components for Δt PDF in the *CP* fitting, a probability density function to extract the *CP* asymmetry from the $(q_{tag}, \Delta t)$ distribution of reconstructed events: One is the resolution function, which considers the incompleteness of Δt determination and is convoluted to PDFs with regard to the Δt variables, such as the PDF for *CP* fitting or that for the lifetime determination, and the other is the background Δt PDF, which takes into account the effect of Δt distribution from events in the Background category. The same components for the control channel are also required for the validation of our analysis procedure, which is described in App. J.

5.1 Resolution function

Resolution functions consider the difference between Δt and $\Delta \tau$ (the true value for Δt) values. We define the resolution function R as the convolution of three functions, which is inspired by the Belle-style resolution function described in literature[36].

$$R = R_k \otimes R_{CP} \otimes R_{tag} \tag{5.1}$$

The resolution function R has 16 parameters in total, and these are determined by fitting the difference between $\Delta \tau$ and reconstructed-Δt values to the resolution function using the signal MC samples. Determined parameters using the MC sample are calibrated with the cosmic samples. We describe the detailed procedure for the calibration in the next section. Each part of Eq. 5.1 describes following components in Δt determination:

- R_k: Correction to the effects on Δt calculation by the motion of B^0 in the CMS frame of $\Upsilon(4S)$.
- R_{CP}: Resolution function of the detector in determining the vertex of B^0_{CP}.
- R_{tag}: Resolution function for the determination of the vertex of B^0_{tag}. This component includes the detector resolution and smearing from the non-primary vertex.

5.1.1 Kinematical approximation

When we calculate the proper time difference between B^0_{CP} and B^0_{tag}, we should consider the relativistic effect and the motion of each B^0 in the rest frame of $\Upsilon(4S)$. We will consider the effects using the kinematical approximation and discuss it in this subsection.

Let's assume B^0_{CP} is moving with the momentum $p^*_{B^0}$ in the rest frame of $\Upsilon(4S)$, and this $\Upsilon(4S)$ is boosted with $\beta\gamma$ in the laboratory frame. Then, we can calculate $\beta\gamma$ of B^0_{CP} in the lab frame with the following (inverse) Lorentz transformation. Here we use natural units.

$$\begin{pmatrix} E_{B^0} \\ p_{Boost} \end{pmatrix} = \begin{pmatrix} E_{CP} \\ m_{B^0}(\beta\gamma)_{CP} \end{pmatrix} = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} E^*_{beam}/2 \\ p^*_{B^0}\cos(\theta^{Boost}_B) \end{pmatrix} \tag{5.2}$$

where p_{Boost} is the component of p_{B^0} along the direction of the boost vector of $\Upsilon(4S)$, E_{B^0} is the kinetic energy of B^0_{CP}, θ^{Boost}_B means the angle between the momentum of B^0_{CP} and the axis of $\Upsilon(4S)$ boosting, E_{beam} is the beam energy, and an asterisk on superscript means the quantity for the CMS frame of $\Upsilon(4S)$. If we solve the above equation for $(\beta\gamma)_{CP}$, we can get the following equations.

$$(\beta\gamma)_{CP} = \beta\gamma\frac{E^*_{beam}}{2m_{B^0}} + \gamma\frac{p^*_{B^0}\cos(\theta^{Boost}_B)}{m_{B^0}} = (a_k + c_k)\beta\gamma \tag{5.3}$$

$$(\beta\gamma)_{tag} = \beta\gamma\frac{E^*_{beam}}{2m_{B^0}} - \gamma\frac{p^*_{B^0}\cos(\theta^{Boost}_B)}{m_{B^0}} = (a_k - c_k)\beta\gamma \tag{5.4}$$

where $a_k = E^*_{beam}/(2m_{B^0})$ and $c_k = (p^*_{B^0}\cos(\theta^{Boost}_B))/(\beta m_{B^0})$. Typical value in Belle II for $a_k \sim 1.002$ and $c_k \sim 0.22\cos(\theta^{Boost}_B)$. Equation 5.4 is calculated using the momentum relation between B^0_{CP} and B^0_{tag} in the back-to-back decay situation. Thus, we can calculate Eq. 3.9 with natural units in a kinematically-smeared way as follows:

$$\begin{aligned} \Delta t &= \frac{1}{\beta\gamma}(\ell_{CP} - \ell_{tag}) = \frac{1}{\beta\gamma}(t_{CP}(\beta\gamma)_{CP} - t_{tag}(\beta\gamma)_{tag}) \\ &= (a_k + c_k)t_{CP} - (a_k - c_k)t_{tag} = a_k(t_{CP} - t_{tag}) + c_k(t_{CP} + t_{tag}) \\ &= a_k\Delta t_{true} + c_k T \end{aligned} \tag{5.5}$$

where T is $t_{CP} + t_{tag}$ and Δt_{true} is a true value of the decay time difference. In this formulation, a_k stands for the relativistic time dilation, and c_k is a change of $(\beta\gamma)_{CP,tag}$ due to the motion of $B^0_{CP,tag}$ in the CMS frame of $\Upsilon(4S)$. The kinematically smeared PDFs can be calculated with the following convolution of P_{true} and $R_k = \delta(\Delta t - (a_k\Delta t_{true} + c_k T))$ as follows:

$$P(\Delta t) = \int_{-\infty}^{+\infty} d(\Delta t_{true}) \int_{|\Delta t_{true}|}^{+\infty} dT P_{true}(\Delta t_{true}, T)\delta(\Delta t - (a_k\Delta t_{true} + c_k T)) \tag{5.6}$$

where δ is the Dirac delta function.

For example, in the case of decays of two B-mesons, we can define a test function P_{true} in Eq. 5.6 as $P_{decay}(t_{CP})P_{decay}(t_{tag}) = (1/\tau_{B^0}^2)\exp\left(-\frac{t_{CP}+t_{tag}}{\tau_{B^0}}\right)$. Then, Eq. 5.6 becomes

$$\begin{aligned} E_k(\Delta t|\tau) &= \frac{1}{2\tau_{B^0}^2}\int_{-\infty}^{+\infty} d(\Delta t_{true}) \int_{|\Delta t_{true}|}^{+\infty} dT e^{-\frac{T}{\tau_{B^0}}} \delta(\Delta t - (a_k \Delta t_{true} + c_k T)) \\ &= \frac{1}{2a_k\tau_{B^0}} e^{-\frac{|\Delta t|}{(a_k \pm c_k)\tau_{B^0}}}. \ \ (+) \text{ for } \Delta t \geq 0, (-) \text{ for otherwise.} \end{aligned} \tag{5.7}$$

where a factor of 1/2 originates from a Jacobian of a transformation given as $(t_{CP}, t_{tag}) \to (T, \Delta t_{true}) : t_{CP} = 1/2(T + \Delta t_{true}),\ t_{tag} = 1/2(T - \Delta t_{true})$.

The equation in the above example is used for the lifetime determination of B^0 and $B^\pm$ in our analysis. a_k, c_k, and kinematical approximation effects are calculated in an event-by-event fashion from the beam parameter and from the reconstruction result of B^0_{CP}, respectively.

5.1.2 *CP*-side resolution function

We represent the detector resolution R_{CP} of the Belle II detector for B^0_{CP} as follows:

$$\begin{aligned} R_{CP}\left(\delta t^{CP}|\sigma_t^{CP}, \left(\frac{\chi^2}{ndf}\right)^{CP}\right) = \\ f_{core}^{CP}\text{Gaus}\left(\delta t^{CP}|\mu = 0, \sigma = \left(c_{core}^{CP} + s_{core}^{CP}\left(\frac{\chi^2}{ndf}\right)^{CP}\right)\sigma_t^{CP}\right) \\ +(1 - f_{core}^{CP})\text{Gaus}\left(\delta t^{CP}|\mu = 0, \sigma = \left(c_{tail}^{CP} + s_{tail}^{CP}\left(\frac{\chi^2}{ndf}\right)^{CP}\right)\sigma_t^{CP}\right) \end{aligned} \tag{5.8}$$

where δt^{CP} is the difference between the true decay vertex and the reconstructed one of B^0_{CP} in time unit, $\left(\frac{\chi^2}{ndf}\right)^{CP}$ is reduced χ^2 from the vertex fitting of B^0_{CP}, and σ_t^{CP} is the uncertainty of ℓ_{CP} converted to units of time with a formula of $\sigma_t^{CP} = \sigma_\ell^{CP}/(\beta\gamma c)$. The resolution strongly depends on conditional variables used in the above equation, so we are taking the variables into account in the calculation of R_{CP}.

All the five parameters of R_{CP} are determined using all reconstruction results from the signal sample of our sub-channel. Table 5.1 shows determined parameters using signal MC samples, and Figs. 5.1, 5.2 show our resolution model can describe the residual distribution as well as conditional variables nicely.

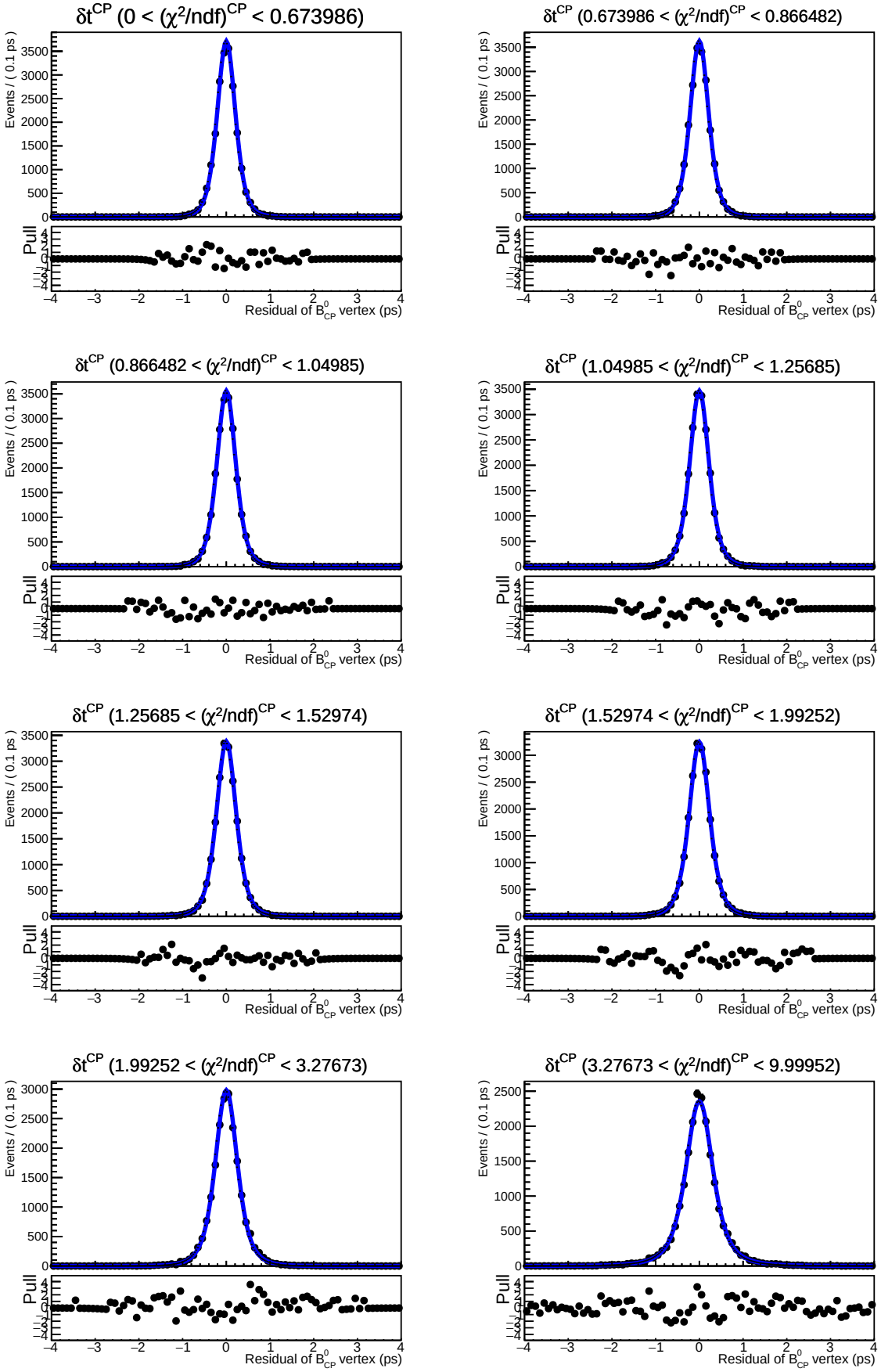


Figure 5.1: Residual distribution for B^0_{CP} vertex and a curve of R_{CP} with slicing by reduced χ^2 of CP-side vertex fitting. Each slicing contains 12.5% of total events.

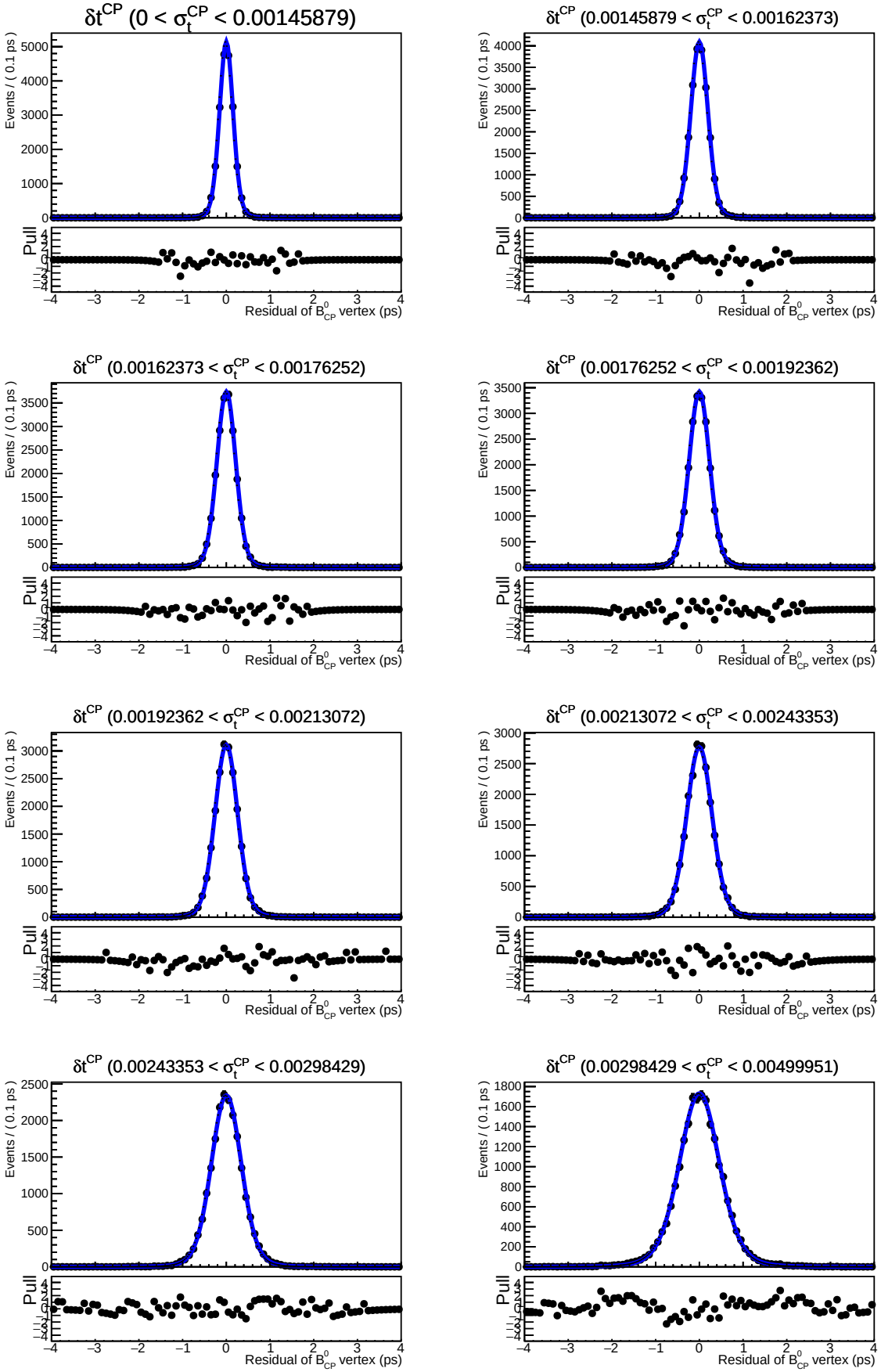


Figure 5.2: Residual distribution for B^0_{CP} vertex and a curve of R_{CP} with slicing by uncertainty of CP-side vertex fitting. The unit of σ_t^{CP} is cm, and each slicing contains 12.5% of total events.

Parameters	Fitted value for main channel	Fitted value for control channel
f_{core}^{CP}	$0.831^{+0.011}_{-0.012}$	$0.8283^{+0.0034}_{-0.0034}$
c_{core}^{CP}	$0.9414^{+0.0050}_{-0.0048}$	
s_{core}^{CP}	$0.0792^{+0.0030}_{-0.0032}$	
c_{tail}^{CP}	$0.939^{+0.019}_{-0.020}$	
s_{tail}^{CP}	$0.417^{+0.015}_{-0.014}$	

Table 5.1: Parameters of R_{CP} determined using Signal MC

5.1.3 Tag-side resolution function

We model the residual for the vertex of B_{tag} with two parts: Primary and non-Primary. Primary part is for a vertex reconstructed using primary particles from B_{tag}. It stands for detector resolution for primary-vertex, which has one Gaussian function to model the detector resolution as given in Eq. 5.9.

$$R_\delta\left(\delta t^{tag}|\sigma_t^{tag},\left(\frac{\chi^2}{ndf}\right)^{tag}\right) = \text{Gaus}\left(\delta t^{tag}|\mu=0,\sigma=\left(c_\delta^{tag}+s_\delta^{tag}\left(\frac{\chi^2}{ndf}\right)^{tag}\right)\sigma_t^{tag}\right) \tag{5.9}$$

where δt^{tag} is the residual of vertex determination for B_{tag}^0, $\left(\frac{\chi^2}{ndf}\right)^{tag}$ is reduced χ^2 of the vertex fitting, and σ_t^{tag} is the uncertainty.

The non-Primary part stands for smearing by the long-lived intermediate state that includes a charm quark. The intermediate states make a bias on the vertex determination of B_{tag}. Thus, this part consists of a convolution of two functions of Gaussian and exponential as stated in Eq. 5.10. The Gaussian component describes the detector resolution, and the exponential component models the decay of the intermediate state.

$$R_{np}\left(\delta t^{tag}|\sigma_t^{tag},\left(\frac{\chi^2}{ndf}\right)^{tag}\right) = (f_n \exp_n(\delta t^{tag}/\tau) + (1-f_n)\exp_p(-\delta t^{tag}/\tau)) \otimes \text{Gaus}\left(\delta t^{tag}|\mu=0,\sigma=\left(c_{np}^{tag}+s_{np}^{tag}\left(\frac{\chi^2}{ndf}\right)^{tag}\right)\sigma_t^{tag}\right) \tag{5.10}$$

where $\exp_{n,p}(x)$ is a one-sided exponential function that is an exponential function if $x<0$ or $x\geq 0$, respectively, and zero for otherwise, and τ and f_n is as follows:

$$f_n = \max\left(0,\min\left(1,f_{n,const}^{tag}+f_{n,slope}^{tag}\sigma_\ell^{tag}\right)\right) \tag{5.11}$$

$$\tau = \min\left(\tau_{ceiling}^{tag},\tau_{const}^{tag}+\tau_{slope}^{tag}\left(\frac{\chi^2}{ndf}\right)^{tag}\right) \tag{5.12}$$

where σ_ℓ^{tag} is uncertainty of the tag-side vertex fitting. Finally, R_{tag} is combined as

$$R_{tag} = f_\delta R_\delta + (1 - f_\delta) R_{np} \tag{5.13}$$

where f_δ is a fraction of the Primary part given by

$$f_\delta = \max\left(0, \min\left(1, f_{\delta,const}^{tag} + f_{\delta,slope}^{tag}\left(\frac{\chi^2}{ndf}\right)^{tag}\right)\right). \tag{5.14}$$

Similarly to R_{CP}, the residual distribution strongly correlates with reduced χ^2 or uncertainty of the tag-side vertex fitting. We determine 11 parameters by fitting R_{np} to the residual distribution of all events from signal MC samples. Figures 5.3 and 5.4 shows the residual distribution and fitted curve of R_{tag} in bins of σ_t^{tag} and $\left(\frac{\chi^2}{ndf}\right)^{tag}$, and Tab. 5.2 is a list of determined parameters.

Parameters	Fitted value for main channel	Fitted value for control channel
$\tau_{\mathrm{exp},const}^{tag}$	1.604 ± 0.022	1.395 ± 0.022
$\tau_{\mathrm{exp},slope}^{tag}$	0.1999 ± 0.0042	0.1813 ± 0.0042
$\tau_{\mathrm{exp},ceil}^{tag}$	9.4 ± 8.4	
$f_{\delta,const}^{tag}$	0.8384 ± 0.0060	
$f_{\delta,slope}^{tag}$	-0.1277 ± 0.0053	
c_δ^{tag}	1.0375 ± 0.0080	
s_δ^{tag}	0.1503 ± 0.0084	
$f_{n,const}^{tag}$	0.0943 ± 0.0062	
$f_{n,slope}^{tag}$	0.0024 ± 0.00019	
c_{np}^{tag}	0.412 ± 0.028	
s_{np}^{tag}	0.1852 ± 0.0056	

Table 5.2: Parameters of R_{tag} determined using Signal MC

5.1.4 Data-MC comparison of conditional variables for $R_{CP,tag}$

We compare the shape of conditional variables for resolution functions between generic MC and the full dataset of experimental data. The plots of the comparison are given in Figs. 5.5, 5.7, 5.6, 5.8, 5.9, and 5.10. As shown in the figures, the difference of the vertex quality variables between MC and Data is insignificant in *CP*-side ones, but that in tag-side shows notable differences. Appendix F contains the MC and Data comparison of vertex quality variables by separating the (Signal + SxF) and Background categories using *sPlot* techniques[55]. In the section, we confirm that the difference mainly comes from the Background category, which can be explained by imperfect $q\bar{q}$ simulation. Also, the differences can originate from differences in the shape of the variables between (Signal + SxF) and Background categories, as shown in App. F, and a different ratio of two

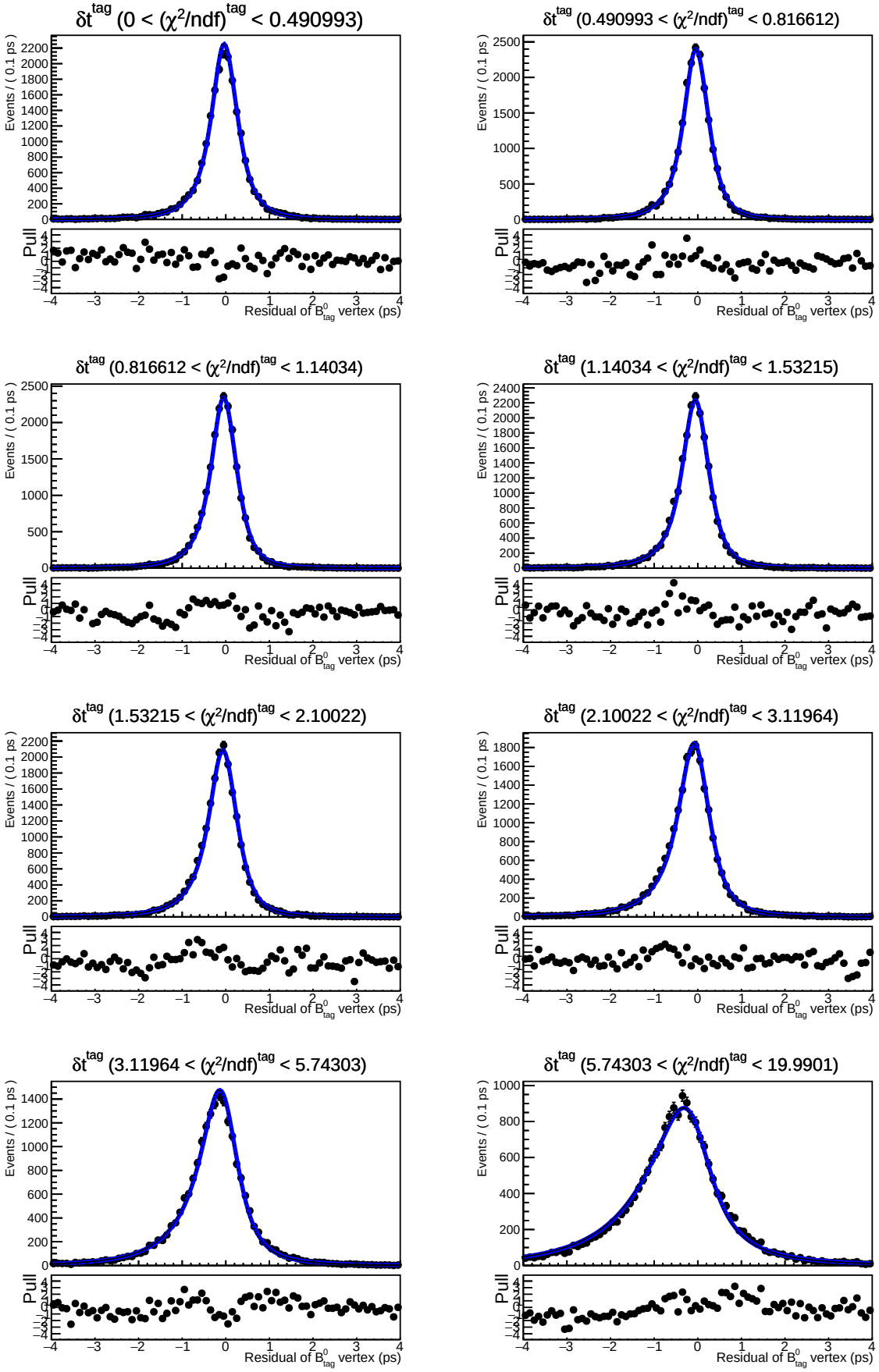


Figure 5.3: Residual distribution for B^0_{tag} vertex and a curve of R_{tag} with slicing by reduced χ^2 of tag-side vertex fitting. Each slicing contains 12.5% of total events.

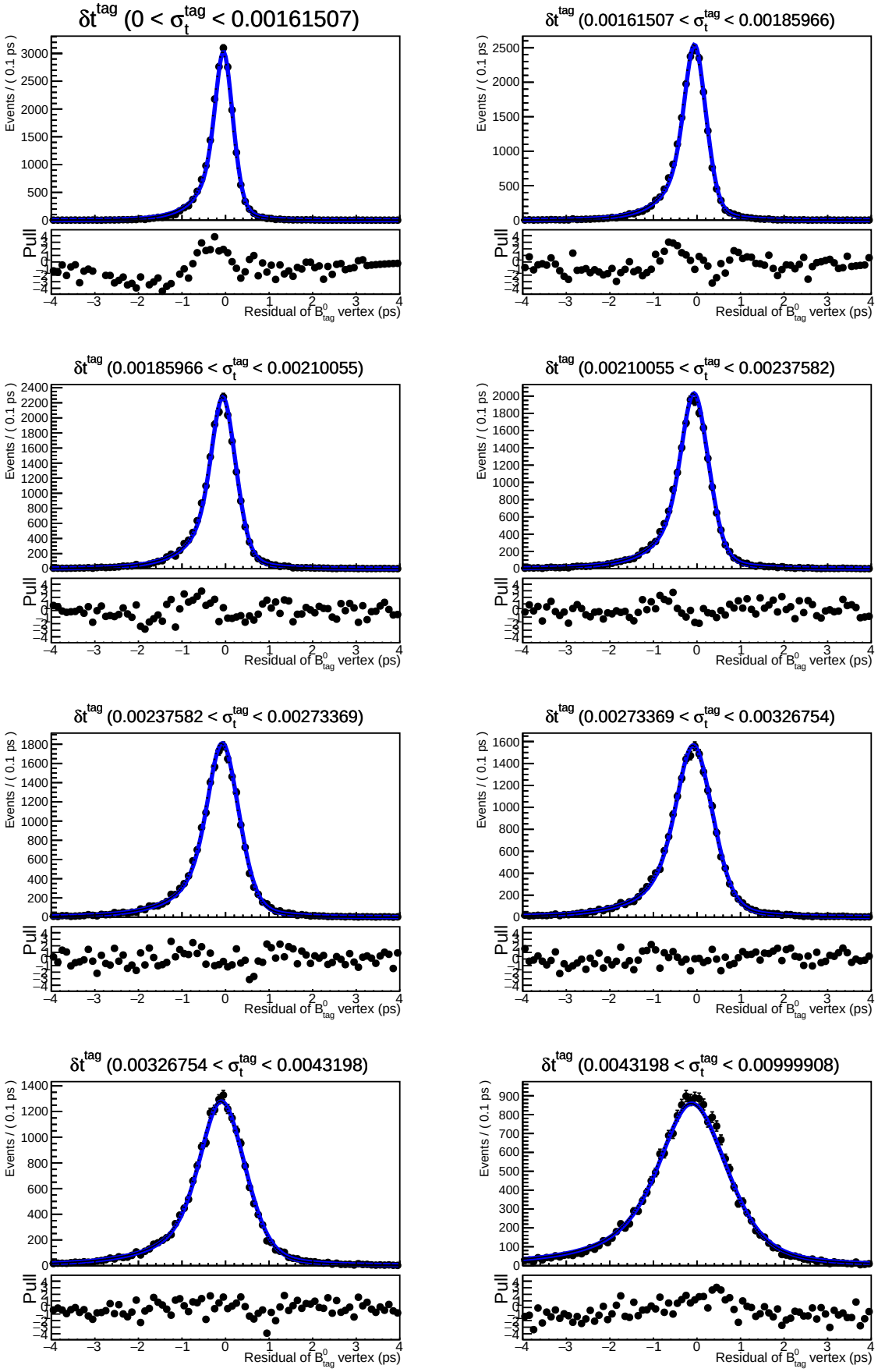


Figure 5.4: Residual distribution for B^0_{tag} vertex and a curve of R_{tag} with slicing by uncertainty of tag-side vertex fitting. The unit of σ_t^{tag} is cm, and each slicing contains 12.5% of total events.

categories (F_{sig}) between the generic MC and the experimental data. We confirmed some differences in the vertex quality, and this will be dealt with the systematic uncertainties and the calibration of the resolution parameters.

5.2 Calibration of parameters for resolution functions

The resolution function and its parameter are determined based on the studies on signal MC samples. However, the Belle II simulation is not perfect due to backgrounds that the simulation cannot consider, inaccurate B-field mapping, inaccurate geometries or limitations of the physics model of the Geant4 framework for the detector simulation, or any other reasons. For this reason, we should check the legitimacy of the resolution function and consider correction (calibration) to the parameters of the function. The calibration determines the difference in the detector resolution between MC and experimental data regarding the vertex fitting for B^0 using the cosmic samples.

One of the widely used methods to calibrate the parameters is setting some of the parameters to floating ones in the *CP* fitting so that the fitting program can determine both *CP* asymmetries and the parameters simultaneously in the experimental data. However, this method has some limitations, as the following list:

- If the size of a dataset of the experimental data is small, this method is not reliable,
- Cannot calibrate the entire parameters,
- If there is a correlation between the parameters, the calibration is not reliable and affects the result of measurement, and
- The dependency of the resolution by the decay mode is not trivial, so the validation per each decay mode is necessary.

Those problems can be addressed by calibrating the parameters independently from the *CP* fitting. We devised such a method to examine differences in the tracking resolution of the Belle II detector between the simulation and reality, which is the main input for the *CP* and tag side vertex resolution using cosmic samples. The reason for using the cosmic samples is that we can reproduce the pull distribution of track parameters without the MC generator information. We elaborate on the detailed procedure for the parameter calibration in the following subsections.

5.2.1 Origin of χ^2 dependency of the resolution function

We study the origin of the χ^2 dependency of the resolution function with a simple MC simulation first. We expect the pull distribution of a vertex fitting to be flat over the reduced-χ^2 value, but it isn't. For example, R_{CP} given in Eq. 5.8 consists of two Gaussians with the σ value that depends on reduced-χ^2 from *CP*-side vertex fitting. We find the reason is related to the structure of inputs for the vertex fitting. To investigate it in detail, we performed a simple MC simulation that mimics the vertex fitting.

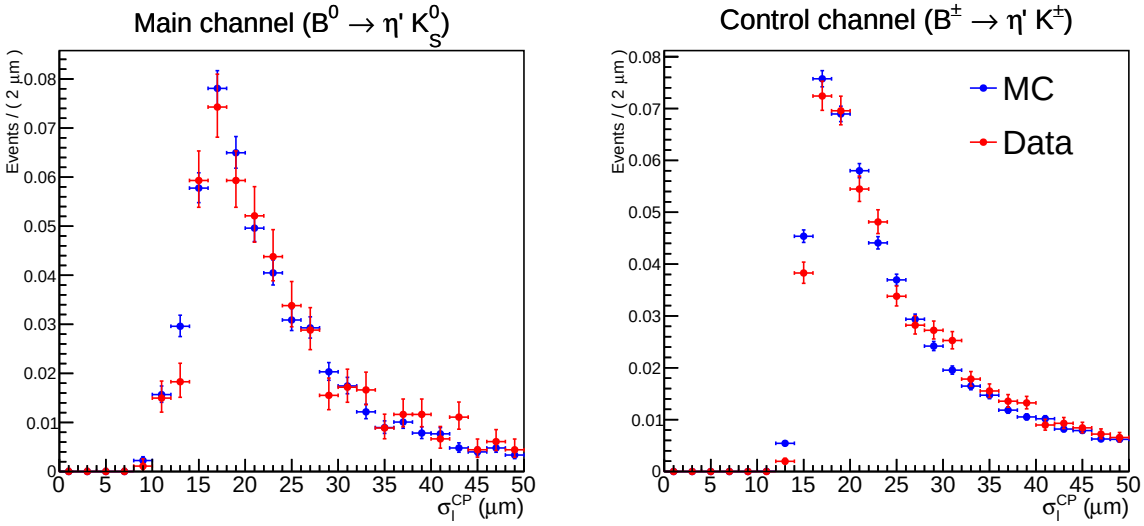


Figure 5.5: Comparison of σ_ℓ^{CP} between MC and Data.

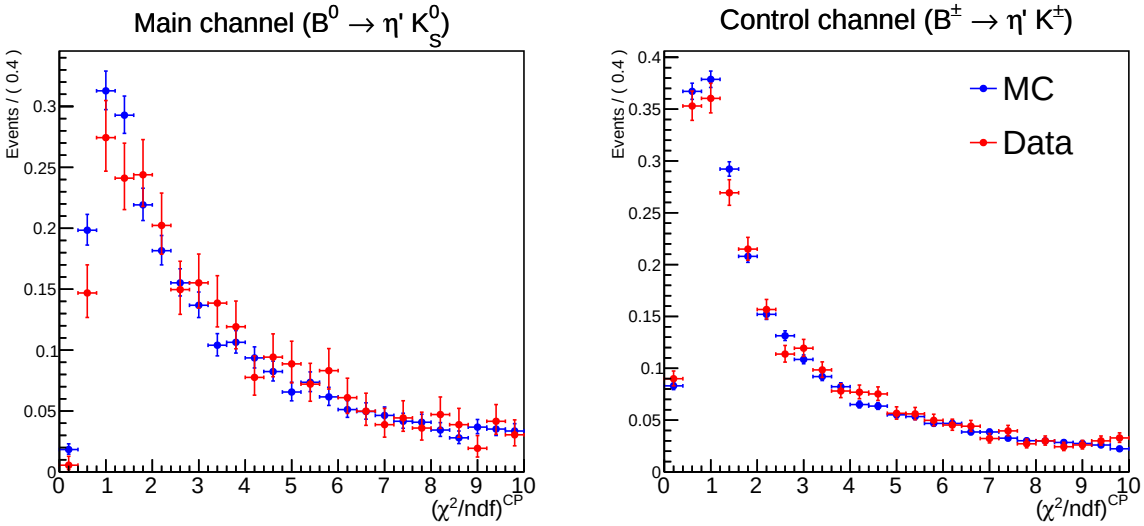


Figure 5.6: Comparison of $(\chi^2/ndf)^{CP}$ between MC and Data.

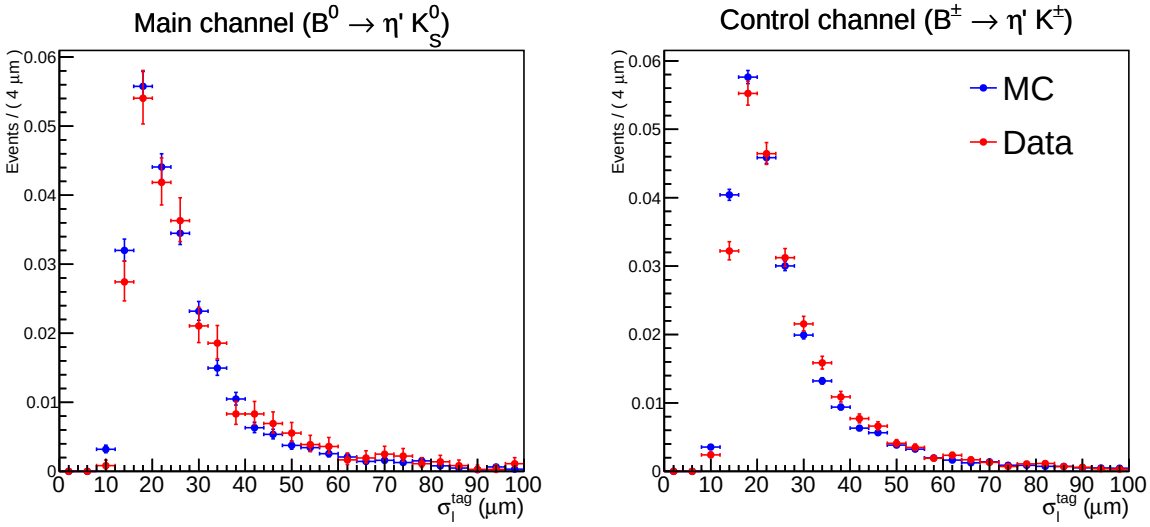


Figure 5.7: Comparison of σ_ℓ^{tag} between MC and Data.

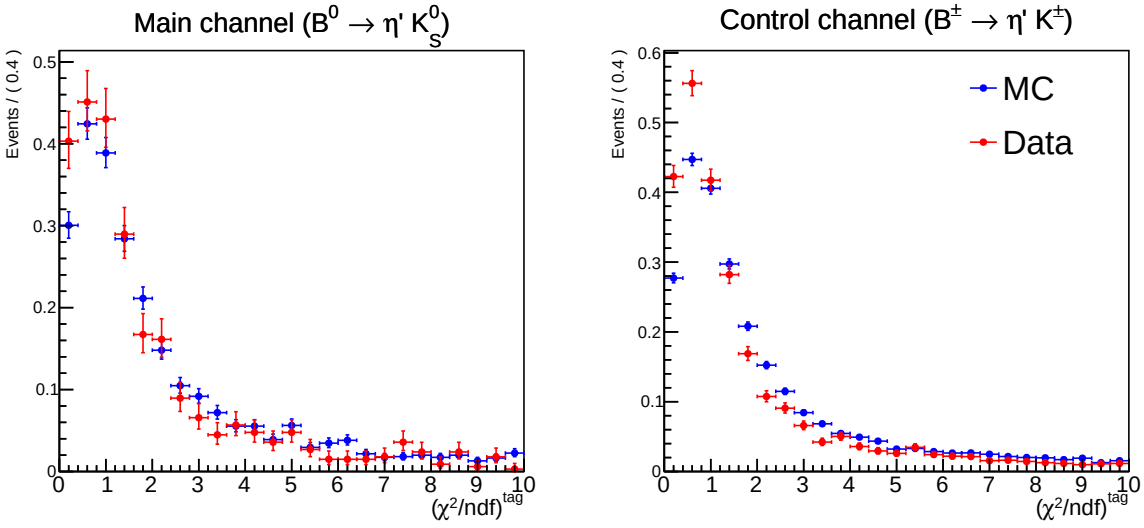


Figure 5.8: Comparison of $(\chi^2/ndf)^{tag}$ between MC and Data.

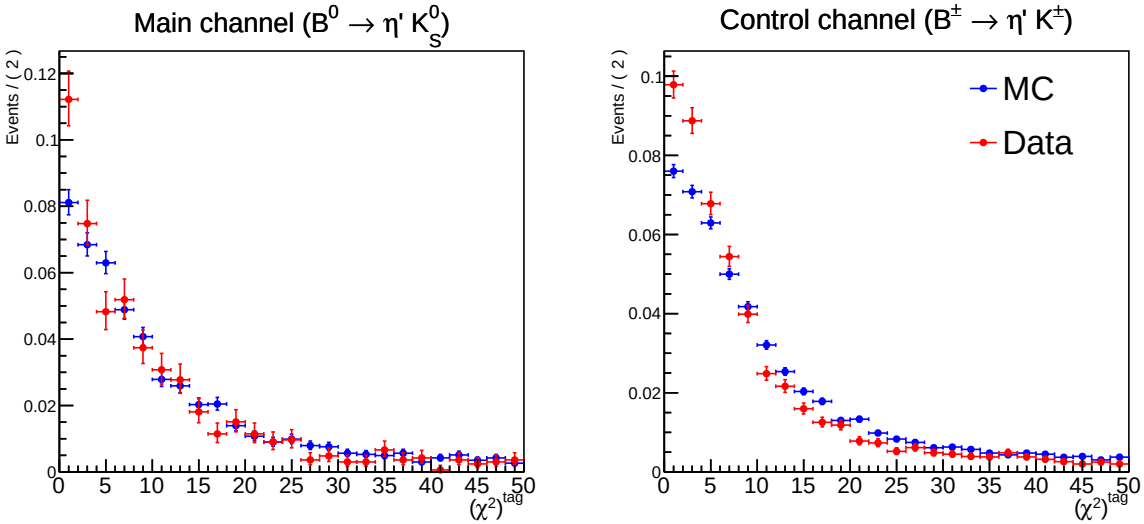


Figure 5.9: Comparison of $(\chi^2)^{tag}$ between MC and Data.

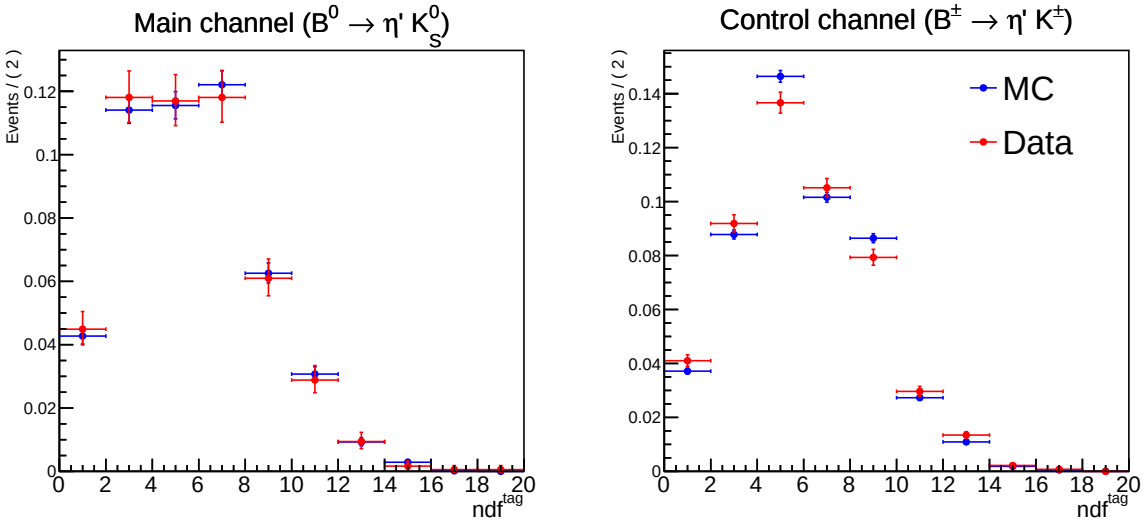


Figure 5.10: Comparison of $(ndf)^{tag}$ between MC and Data.

Firstly, we generate six (which is the number of $\pi^{\pm}$ from η' and K_S^0) x with a single Gaussian that has $\mu = 0, \sigma = \sigma_\ell^{CP}$. The σ_ℓ^{CP} values for each x value are obtained from the *CP*-side vertex fitting results of the signal MC samples to mimic the real error distribution.

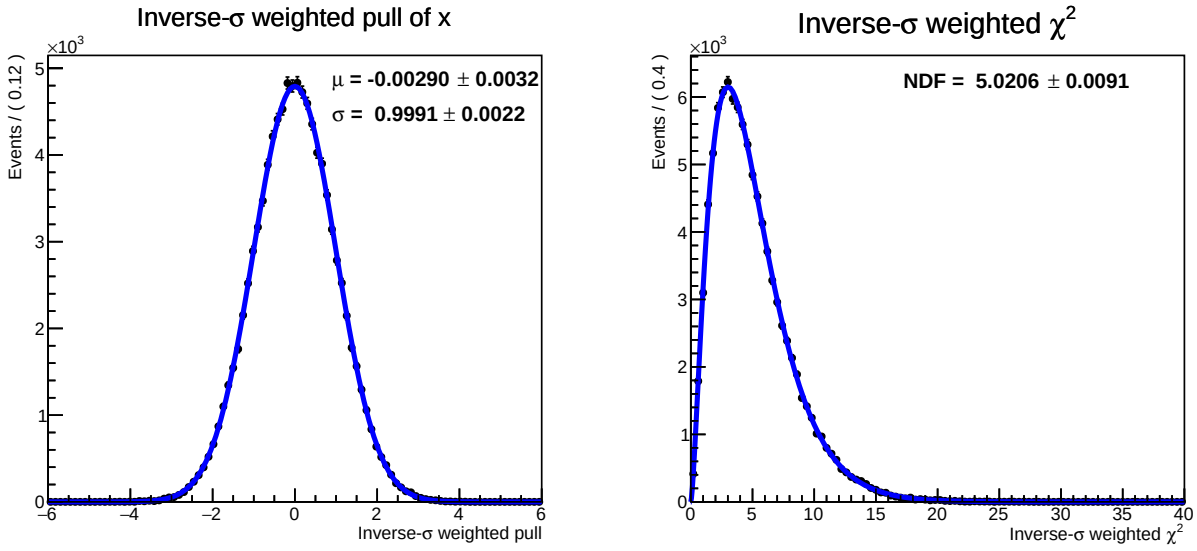


Figure 5.11: The fitted result for pull distribution (left) with a Gaussian and χ^2 distribution with a χ-squared PDF function.

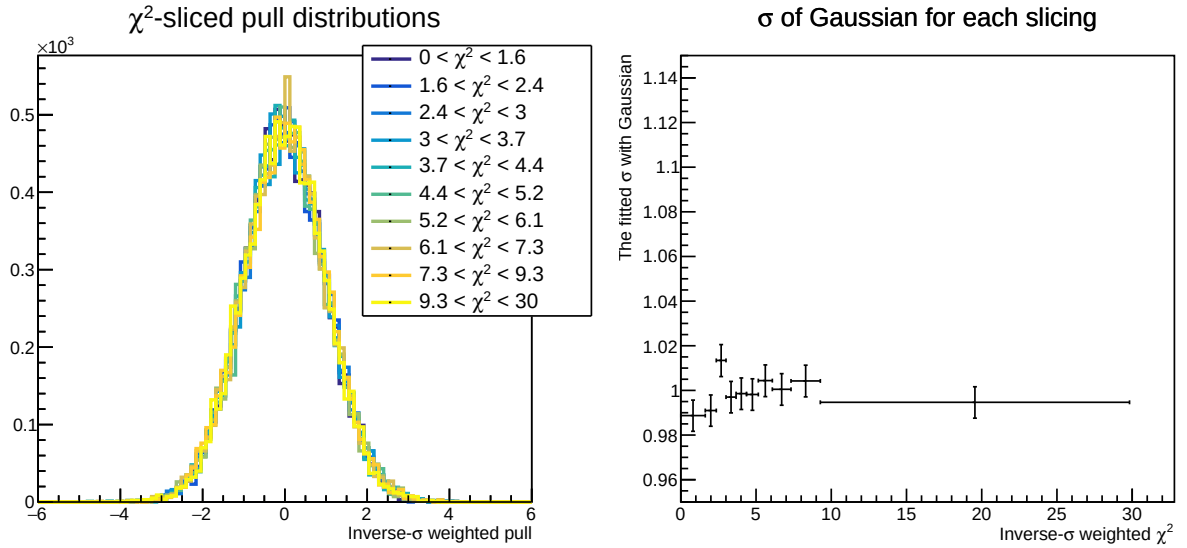


Figure 5.12: χ^2-sliced pull distributions and the σ values of fitted Gaussian for each slicing. All slices have equal statistics.

Then, we calculate inverse-σ weighted mean μ, pull p, and χ^2 values as follows:

$$\mu = \frac{\Sigma_{i=1}^{6}(x_i/\sigma_i^2)}{\Sigma_{i=1}^{6}(1/\sigma_i^2)} \tag{5.15}$$

$$p = \mu \cdot \sqrt{\Sigma_{i=1}^{6}(1/\sigma_i^2)} \tag{5.16}$$

$$\chi^2 = \Sigma_{i=1}^{6}\left[((x_i-\mu)/\sigma_i)^2\right] \tag{5.17}$$

The first step of generating x values represents the determination of track parameters, and the second step simulates the calculation procedure of the pull values from the vertex fitting. This procedure was repeated 100000 times, and we fit the distribution of pull values with a Gaussian and the distribution of χ^2 values with a χ-squared distribution. The functions fit well with the distributions, and the parameters are as we expected, as shown in Fig. 5.11. If we slice the pull distribution with χ^2 values and fit the distribution with a Gaussian, we can see the width of the distribution is consistent over all slicing as in Fig. 5.12.

Next, we generate six x values with a double Gaussian of $\mathrm{PDF}(x|\mu,\sigma) = 0.9 \cdot \mathrm{Gaus}(x|0,\sigma_\ell^{CP}) + 0.1 \cdot \mathrm{Gaus}(x|0,2\sigma_\ell^{CP})$ and calculate μ, p, and χ^2 values as stated above. Then, the pull distribution differs from the standard normal distribution, and we cannot model the χ^2 distribution with a simple χ-squared PDF, as shown in Fig. 5.13. Similar to what we did in the case with a single Gaussian, we slice the pull distribution by χ^2 and fit it with a Gaussian. At this time, we confirm that the width of Gaussian depends on χ^2, as shown in Fig. 5.14.

To sum up, we imitate the vertex fitting procedure from track parameters by generating six x values and calculating μ and χ^2 values with inverse-σ weighting. Then, we emulate the determination for parameters of resolution functions by calculating pull distribution and fitting it with a Gaussian. We confirm that there is no χ^2 dependency in the case of generating x values with a Gaussian, but the dependency emerges in the case of generating x values with a double Gaussian. This result suggests the structure of input data for the vertex fitting (track parameters of $\pi^\pm$ from η' and K_S^0 for our case) makes χ^2 dependency of the resolution functions.

We checked the pull distribution of track parameters from the signal MC samples and fitted a double Gaussian function to it. The example distribution of pull values for d_0 and z_0 track parameters, fitted curve, and parameters are shown in Fig. 5.15. We can confirm that the pull distributions are well described with the double Gaussian, and there is a sizable fraction ($\sim$10%) of minor(tail) components of the function. This result confirms that the χ^2 dependency of the resolution function originates from the tail component of the pull distribution for the track parameters.

5.2.2 Calibration procedure

In this subsection, we briefly discuss the calibration procedure of parameters of the resolution function using cosmic samples from MC simulation or experimental dataset. Basi-

cally, the calibration is a correction to a set of resolution parameters determined using the signal MC samples, which are given in Tabs. 5.1 and 5.2, and factors for the correction are determined from two sets of resolution parameters determined using the cosmic MC and data samples. The following contents describe how we determined the correction factor and how it is applied to the resolution parameters from the signal MC samples.

In the previous subsection, we confirmed that the pull distribution of track parameters significantly affects the resolution of vertex fitting. Thus, we examine the difference in resolution parameters due to the difference in pull distributions of track parameters using

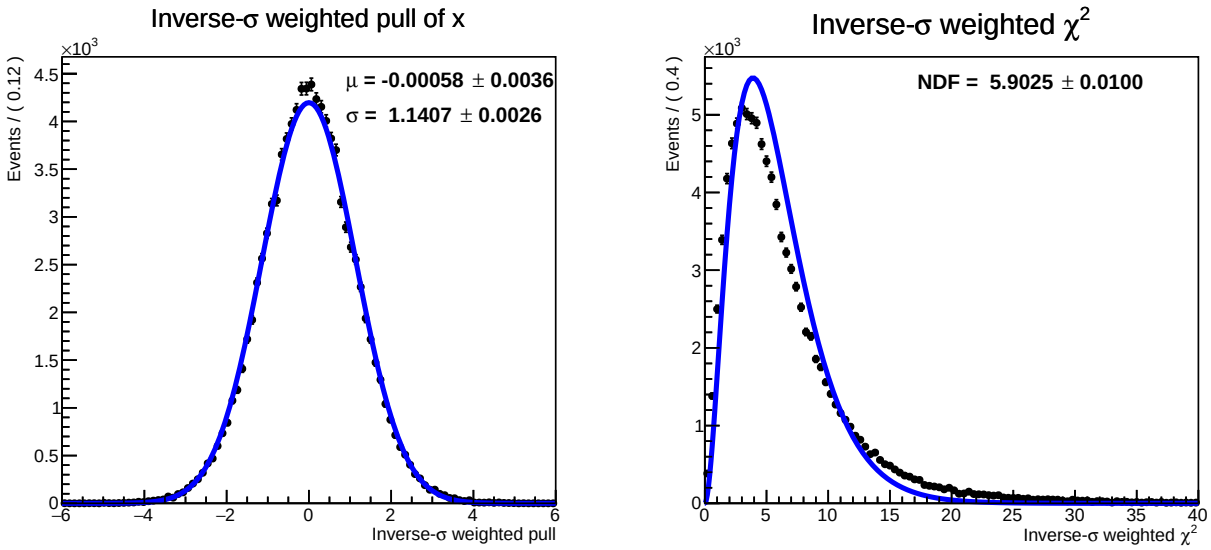


Figure 5.13: The fitted result for pull distribution (left) with a Gaussian and χ^2 distribution with a χ-squared PDF function. The x values are generated with double Gaussian.

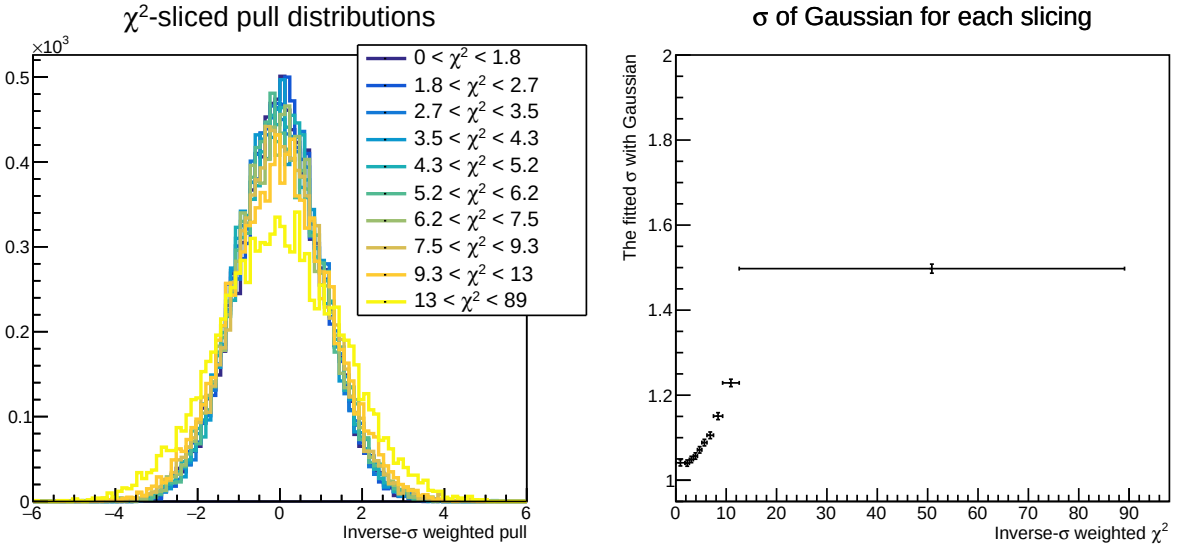


Figure 5.14: χ^2-sliced pull distributions from the generation with double Gaussian and the σ values of fitted Gaussian for each slicing. All slices have equal statistics.

cosmic samples. In detail, we can reproduce the pull distribution of track parameters using cosmic samples and examine the difference in the pull distribution between MC and experimental data samples. This can be done by exploiting the track reconstruction process of Belle II. In the process, a trajectory for cosmic muons is reconstructed by two tracks as given in Fig. 5.16.

Since the two tracks are independent measurements for a single particle (cosmic muon), we can extract the pull value p'_x of each track parameter x as follows:

$$p'_x = \frac{x^{upper} - x^{lower}}{\sqrt{(\sigma_x^{upper})^2 + (\sigma_x^{lower})^2}} \tag{5.18}$$

This enables us to calculate the pull distribution in both samples of MC and experimental data and to assess the difference in the pull value between the two types of samples. We model the pull distribution with a double Gaussian to reproduce the minor(tail) component of the pull distribution of track parameters. The comparison and fitting results for the pull distribution of d_0 and z_0 between cosmic MC and data samples are shown in Fig. 5.17. Also, we confirmed the difference in the property of the pull distribution of track parameters from the cosmic samples and the signal MC samples, and the difference is considered by a correction method explained in App. C.

The double Gaussian is embedded into the pull distribution of track parameters in signal MC samples to examine the MC and data difference in the aspect of resolution parameters. This procedure is called "Fast Simulation". Then, the differences are converted to the correction factor for resolution parameters, which is defined by two types: Bias and Scale factor. Each type of correction factor is applied to the resolution parameters given in Tabs. 5.1 and 5.2, which is determined using original signal MC samples, with

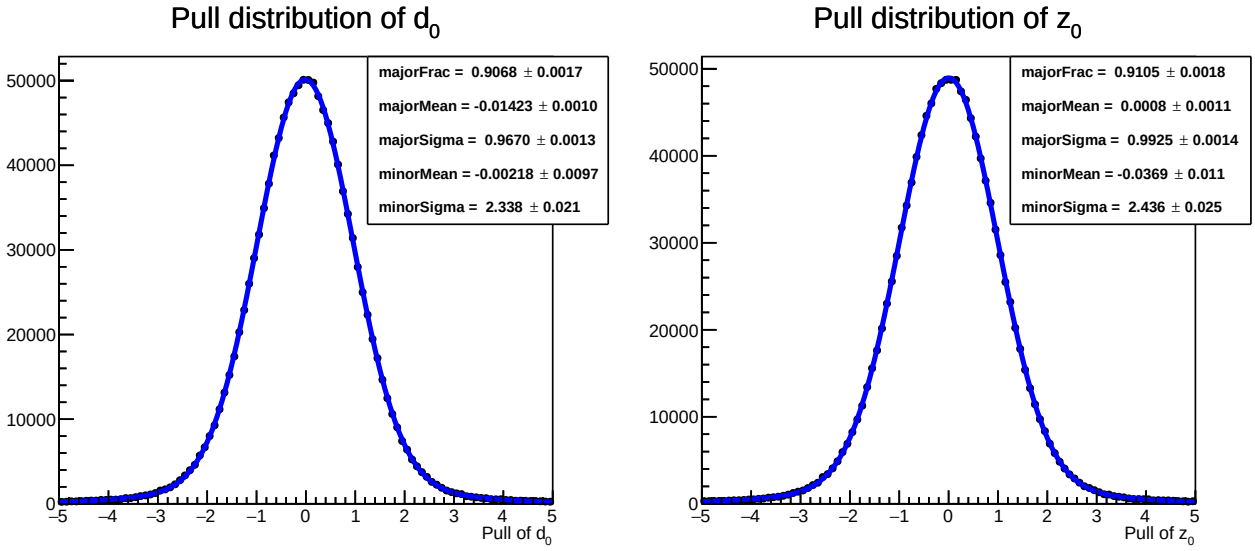


Figure 5.15: Example distributions of pull value for track parameters of d_0 and z_0 from the signal MC sample, and their fitting result to the double Gaussian with fitted parameters.

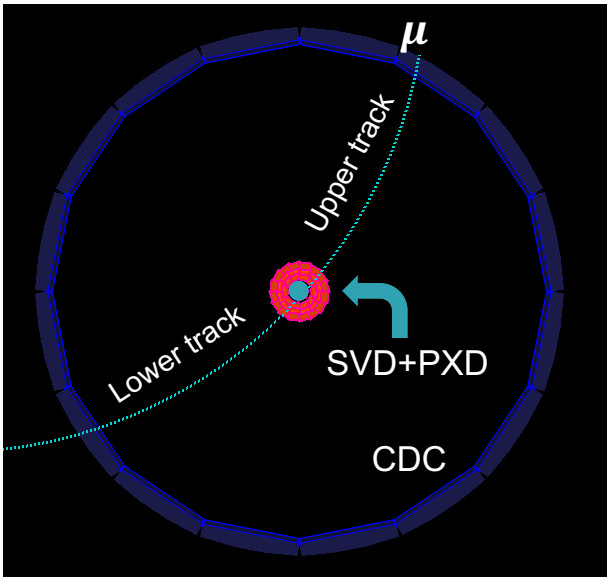


Figure 5.16: An example snapshot of a cosmic MC event where the muon passes around of IP. This snapshot is projected on the $(r-\phi)$ plane.

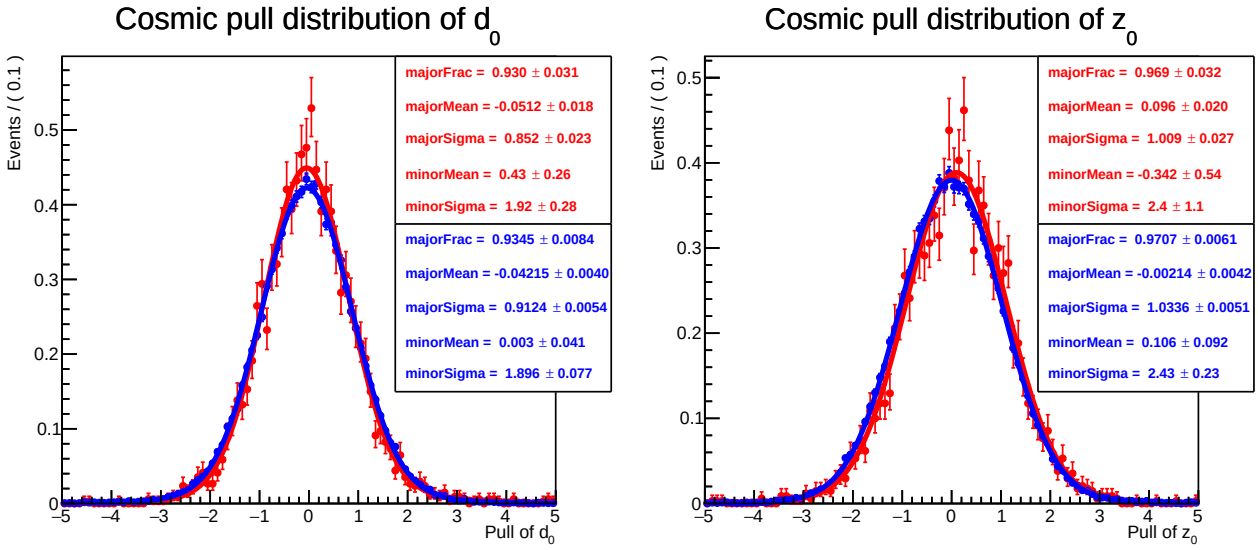


Figure 5.17: Comparison for the pull distribution of track parameters of d_0 and z_0 between cosmic MC(blue) and data(red). The solid line represents the fitted double Gaussian to each distribution, and parameters are shown.

the following way for a given parameter α of the resolution function.

$$\text{For Scale factor } \beta \quad : \quad \alpha \to \alpha_{scale}^{corrected} = \alpha \cdot \beta \tag{5.19}$$

$$\text{For Bias factor } \gamma \quad : \quad \alpha \to \alpha_{bias}^{corrected} = \alpha + \gamma \tag{5.20}$$

The correction factors for calibration of resolution parameters are given in Tab. 5.3,

and the comparison of the resolution function before and after the calibration is shown in Fig. 5.18, which indicates the resolution after the calibration is slightly worsened. We consider the reason for the deterioration of the resolution in the experimental data due to the imperfect MC simulation of Belle II, such as ignoring of electrical noise of detector signals in the simulation for digitization.

Parameters	Correction method	Correction factor
f_{core}^{CP}	Scale factor	0.794
c_{core}^{CP}	Bias factor	-0.108
s_{core}^{CP}	Scale factor	1.87
c_{tail}^{CP}	Bias factor	-0.0618
s_{tail}^{CP}	Scale factor	1.29
$\tau_{\mathrm{exp},const}^{tag}$	Bias factor	0.108
$\tau_{\mathrm{exp},slope}^{tag}$	Scale factor	0.888
$\tau_{\mathrm{exp},ceil}^{tag}$	Bias factor	4.13
c_{δ}^{tag}	Bias factor	-0.0347
s_{δ}^{tag}	Scale factor	1.38
c_{np}^{tag}	Bias factor	-0.00248
s_{np}^{tag}	Scale factor	0.913

Table 5.3: Correction factors for resolution parameters.

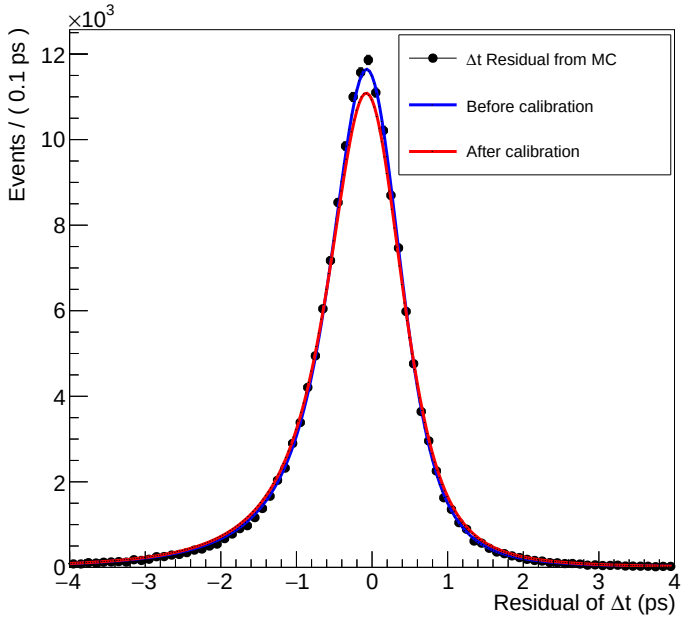


Figure 5.18: Comparison of the resolution function before (blue line) and after (red line) the calibration. The black dots are the distribution of residual values for the Δt determination.

The methodology for the calibration is summarized in the following list, and detailed treatments and procedures of the calibration for resolution parameters are explained in App. C.

1. Determine the resolution parameters using the signal MC samples.
2. Determine parameters for the pull distribution of track parameters from cosmic MC and data samples.
3. Perform Fast Simulation using track parameters from cosmic MC or data samples and determine the parameters for resolution functions R_{CP} and R_{tag} from each set of track parameters.
4. Calculate scaling or bias factor using the difference of resolution parameter between cosmic MC and data.
5. Apply factors from 4. to resolution parameters obtained using signal MC sample to get the parameters for experimental data.

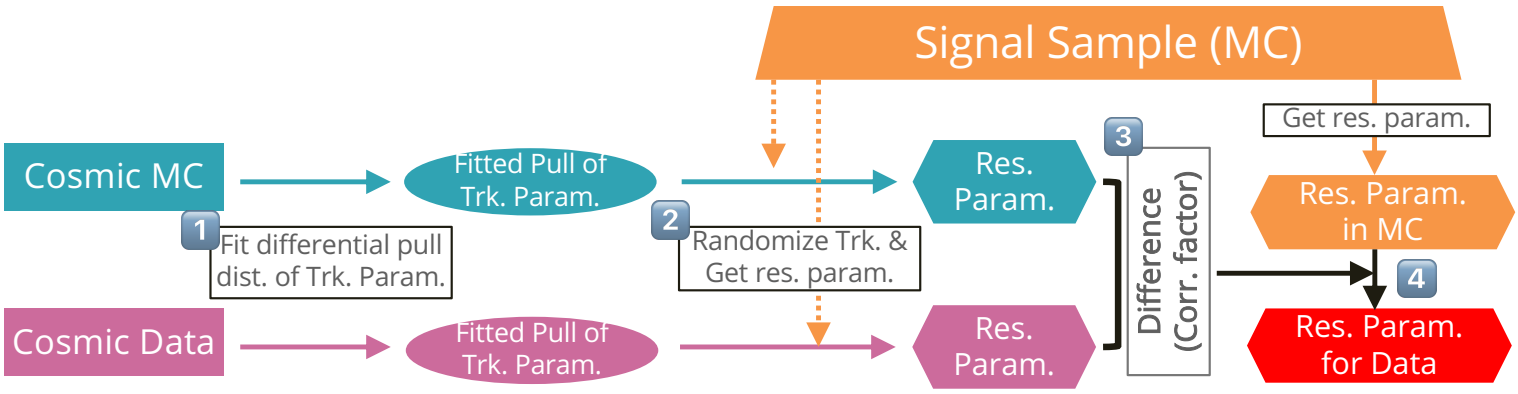


Figure 5.19: Procedure diagram for the resolution parameter calibration using cosmic samples.

Especially, if the modeling of the pull distribution of track parameters is ideal and Fast Simulation can reproduce resolution parameters from signal MC samples perfectly, we can skip the last step so that the entire procedure can be simplified. We leave this possible improvement to the next analyst since the current methodology is sufficient for the current statistics. This procedure is illustrated in Figure 5.19. The validation of resolution parameters calibrated using this methodology is discussed in the next chapter.

5.3 Background Δt function

The main background component for this analysis is $q\bar{q}$ events. These events decay at collision points, so the Belle II detector cannot distinguish the flight of $q\bar{q}$; thus, we can

model a distribution of Δt from the events with a convolution of the Dirac delta for the decay of $q\bar{q}$ and double Gaussian function for the detector resolution. The PDF for the modeling has seven parameters and is given as Eq. 5.21.

$$\begin{aligned} P_{bckg}\left(\Delta t|\Sigma_t, X^2\right) = & f_{core}^{bckg}\text{Gaus}\left(\Delta t|\mu=\mu_{core}^{bckg}, \sigma=(c_{core}^{bckg}+s_{core}^{bckg}X^2)\Sigma_t\right) \\ & + (1-f_{core}^{bckg})\text{Gaus}\left(\Delta t|\mu=\mu_{tail}^{bckg}, \sigma=(c_{tail}^{bckg}+s_{tail}^{bckg}X^2)\Sigma_t\right) \end{aligned} \tag{5.21}$$

where X^2 is the average value of reduced χ^2 of vertex fitting for *CP* and tag-side vertex fitting and Σ_t is the uncertainty of total Δt determination given as following equations:

$$X^2 = \frac{1}{2}\left(\left(\frac{\chi^2}{ndf}\right)^{CP} + \left(\frac{\chi^2}{ndf}\right)^{tag}\right) \tag{5.22}$$

$$\Sigma_t = \sqrt{\left(\sigma_t^{CP}\right)^2 + \left(\sigma_t^{tag}\right)^2} \tag{5.23}$$

We determine parameters for the background Δt PDF by fitting Δt distributions from background events using MC-truth information from the generic MC sample 400/fb for the MC study and sideband events (both types of M_{bc} and ΔE) for experimental data with a selection criterion of $\Delta t < 10$ps to reject a few outlier events. Determined parameters for the main channel are given in Tab. 5.4, and Δt distribution and fitted curves are shown in Figs. 5.20 for MC and 5.21 for experimental data. The background Δt PDF determined using MC backgrounds is used in *CP* fitting of MC samples, and the PDF determined using sideband events is used in *CP* fitting with the experimental data.

Channel	Parameters	Fitted value using	
		MC backgrounds	Sideband in experimental data
Main	f_{core}^{bckg}	$0.923^{+0.013}_{-0.015}$	$0.889^{+0.028}_{-0.034}$
	μ_{core}^{bckg}	$0.008^{+0.012}_{-0.012}$	$0.011^{+0.023}_{-0.023}$
	c_{core}^{bckg}	$0.852^{+0.043}_{-0.042}$	$0.733^{+0.089}_{-0.091}$
	s_{core}^{bckg}	$0.336^{+0.019}_{-0.019}$	$0.309^{+0.033}_{-0.031}$
	μ_{tail}^{bckg}	$0.19^{+0.16}_{-0.15}$	$0.11^{+0.23}_{-0.23}$
	c_{tail}^{bckg}	$2.84^{+0.71}_{-0.60}$	$2.99^{+0.80}_{-0.58}$
	s_{tail}^{bckg}	$1.03^{+0.24}_{-0.21}$	$0.79^{+0.26}_{-0.22}$

Table 5.4: List of parameters of Background Δt PDFs from MC and the Data.

To confirm events from the sideband can represent events that are actually used for the *CP* fitting, we checked the homogeneity of Δt distribution in the MC samples between sideband events and background events in the signal region using MC truth information as we did in Sec. 4.5.1. The result in Fig. 5.22 indicates no significant differences between the two populations.

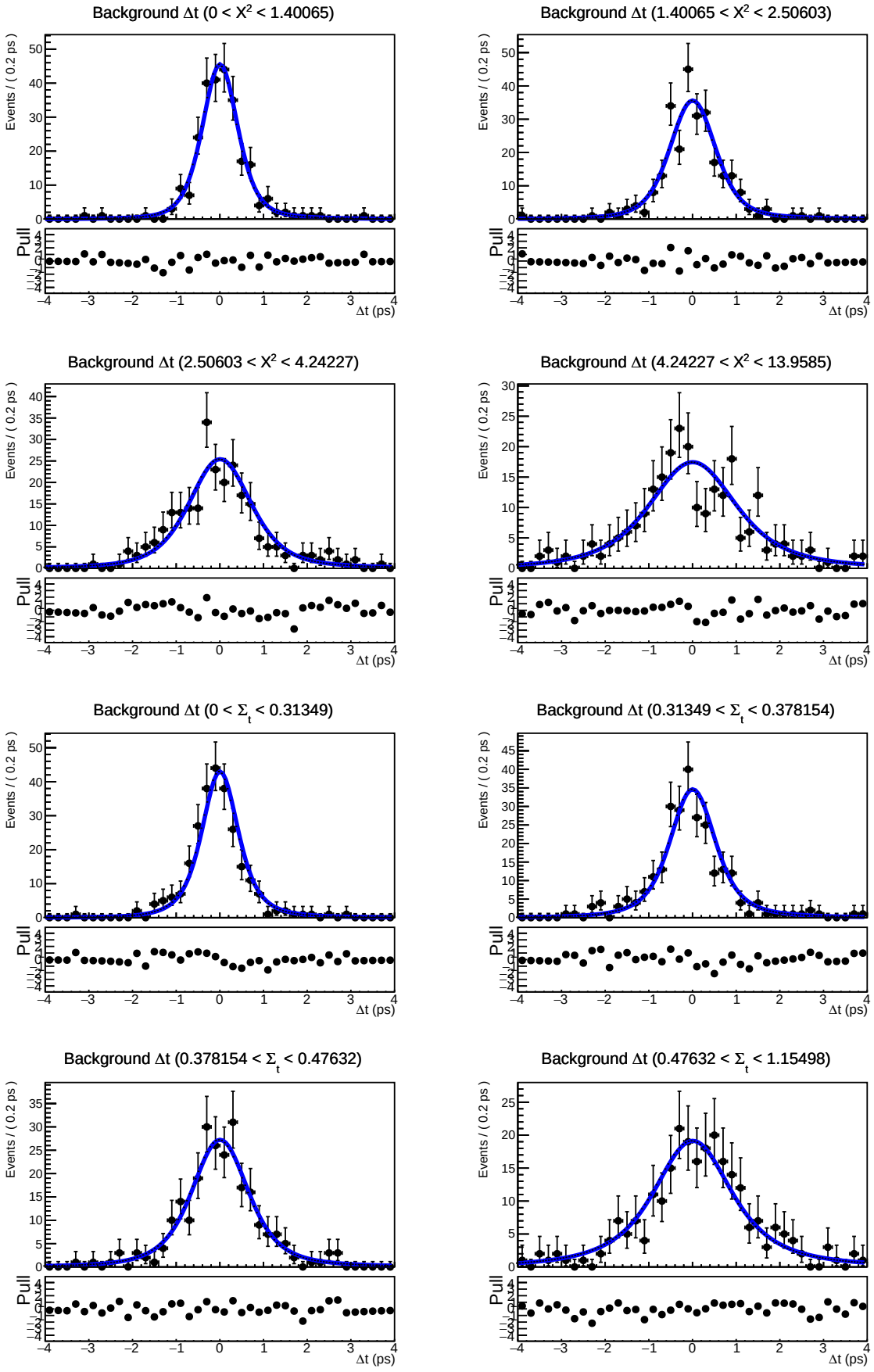


Figure 5.20: Δt distribution of background events from 400/fb Generic MC and fitted curve of Background Δt PDF for the main channel. The upper four plots are sliced by X, and Σ_t slices the below four. The unit of Σ_t is ps, and each slicing contains 25% of total events.

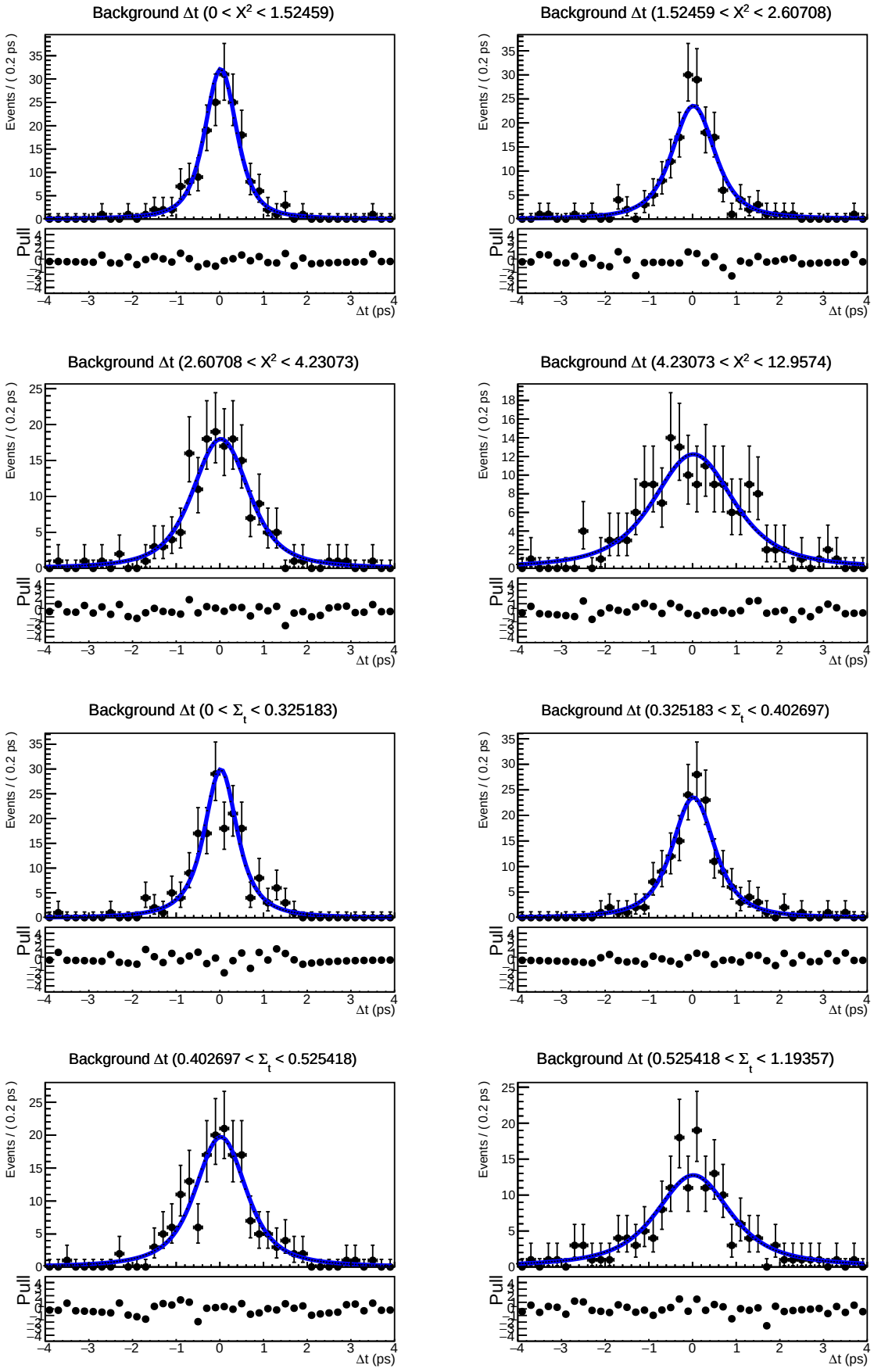


Figure 5.21: Δt distribution of sideband events from experimental data and fitted curve of Background Δt PDF for the main channel. The upper four plots are sliced by X, and Σ_t slices the below four. The unit of Σ_t is ps, and each slicing contains 25% of total events.

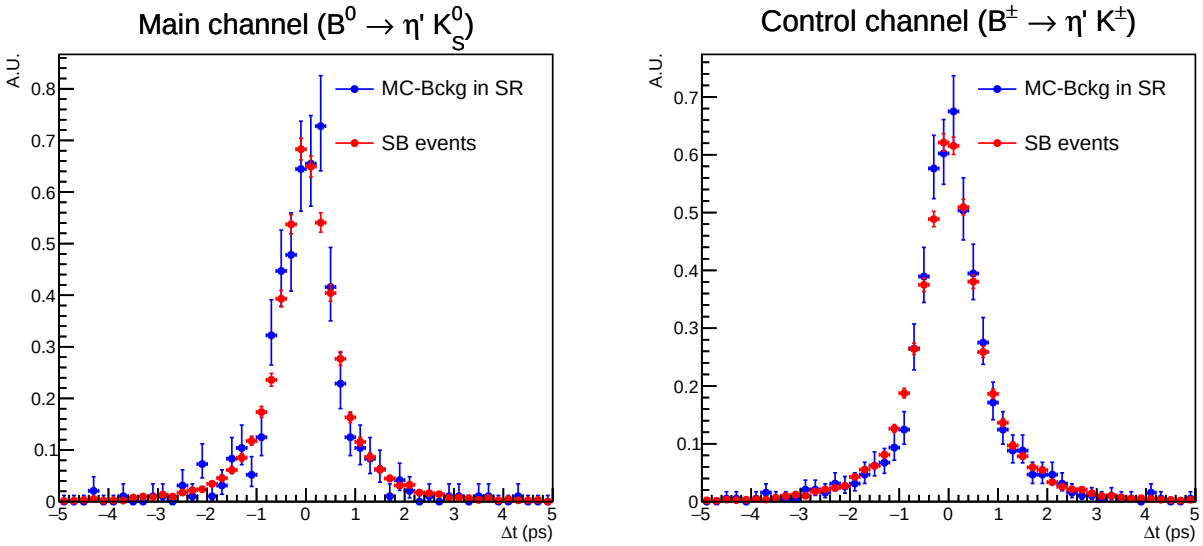


Figure 5.22: Comparison of Δt distribution between a sideband and MC-background events from the signal region.

Chapter 6

Extraction of *CP* Asymmetries

In this chapter, we will discuss the methodology for the *CP* fitting and its validation using various methods and finally present the result of *CP* asymmetries in our target decay using the Belle II experimental data.

To consider the incompleteness of the reconstruction method of our target decay, we used all components discussed in the previous chapters for the *CP* fitting and its relevant tests. The tests for the *CP* fitting are the lifetime determination, *CP* fitting with MC samples, linearity tests using specially generated signal MC samples, and ensemble tests using the Toy MC dataset by an unbinned Maximum-Likelihood fitting. We will explain all these tests first, then show the results of *CP* fitting with experimental data and discuss the systematic uncertainties of the *CP* asymmetry measurement.

6.1 Validation by determining lifetimes

We performed the lifetime fitting to check our analysis procedure. MC samples or experimental data were used for the lifetime fitting, and ensemble tests using Toy MC data were also performed to check the validation of the fitting. The following sections will describe the details.

6.1.1 Signal sample

We confirmed the validity of the resolution function and its parameters determined in the previous chapter by estimating the lifetime of signal samples. We used all reconstructed events of the (Signal + SxF) category from the signal sample that has the lifetime input of $\tau_{B^0} = 1.519\,\text{ps}$ and $\tau_{B^+} = 1.638\,\text{ps}$ for the Belle II simulation.

Two unbinned ML fittings are performed to determine the lifetime for two (main, control) channels with the following variables: PDFs for the lifetime fitting are defined as follows:

- `mcDeltaTau`($\Delta\tau$): A "true" value for the difference of decay times between B^0_{CP} and B^0_{tag} from MC event generator, which corresponds $\Delta\tau$ in Eq. 3.8. This fitting uses a PDF of $E_f(\Delta\tau|\tau_B) = (1/2\tau_B)\exp(-|\Delta\tau|/\tau_B)$ and determines the reference lifetime of B^0 using reconstructed B^0_{CP} events as population.
- `DeltaT` $\left(\Delta t = \frac{\ell_{CP}-\ell_{tag}}{\beta\gamma c}\right)$: The difference of proper decay times with all effects from kinematics and the Belle II detector, that we can obtain only this value from the experimental data. We used $(E_k \otimes R_{CP} \otimes R_{tag})(\Delta t|\tau_B)$ for PDF.

where E_k is a kinematically-smeared E_f which is given in Eq. 5.7.

The fitted lifetime with the $\Delta\tau$ variable shown in Tab. 6.1 shows good agreement with simulation inputs for both channels. However, for the results with the Δt variable from the table, the control channel shows a discrepancy of 2σ from the estimation result using $\Delta\tau$. The fitting result is consistent with the lifetime input for the simulation. However, a discrepancy between `mcDeltaTau` and `DeltaT` suggests there might be a bias in the lifetime due to the resolution function, thus we will examine the effects of it on *CP* fitting in the systematic studies. Figures 6.1 and 6.2 shows the distribution of each variable from the signal MC sample and fitted curves for both of main and control channel. The lifetime fitting with other MC variables that contain intermediate information is also performed, and its results are given in App. K.

Variables for fitting	Fitted τ_{B^0}	Fitted $\tau_{B^\pm}$
`mcDeltaTau`	1.520(3)ps	1.641(4)ps
`DeltaT`	1.520(4)ps	1.633(4)ps

Table 6.1: Determined lifetimes of B^0 and $B^\pm$ from $\Delta\tau$ and Δt

6.1.2 Generic MC samples

We determined the lifetime of B-mesons using generic MC samples that are equivalent to 1/ab amount of integrated luminosity to check the validity of f_{sig} calculation and background Δt modeling with the following PDF:

$$P_{\text{Lifetime}}(\Delta t|\tau_B) = f_{sig}(E_k \otimes R_{CP} \otimes R_{tag})(\Delta t|\tau_B) + (1 - f_{sig})P_{bckg}(\Delta t). \tag{6.1}$$

The fitting result for both channels is $\tau_{B^0} = (1.634^{+0.125}_{-0.115})$ ps and $\tau_{B^\pm} = (1.751^{+0.070}_{-0.067})$ ps, which is consistent with the input lifetime for the MC simulation within 2σ. Figures 6.3 and 6.4 show Δt distribution of generic MC and background-subtracted distribution using *sPlot* technique[55] with fitted curves. An ensemble test is also performed to check the validity of this lifetime fitting, and the test does not find any problems with the fitting procedure, which is discussed in App. K.

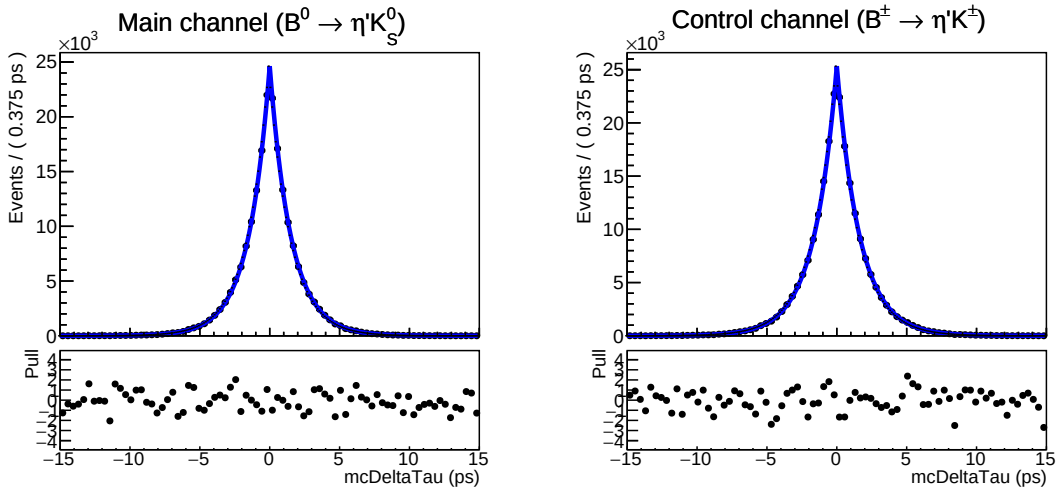


Figure 6.1: mcDeltaTau distribution and fitted curve of E_f for main and control channel.

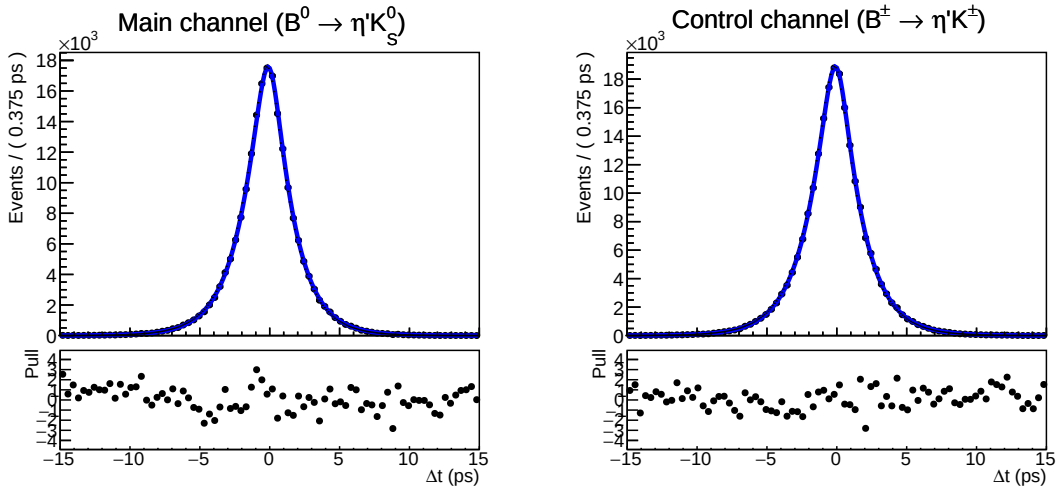


Figure 6.2: Δt distribution and fitted curve of $E_k \otimes R_{CP} \otimes R_{tag}$ for both channels.

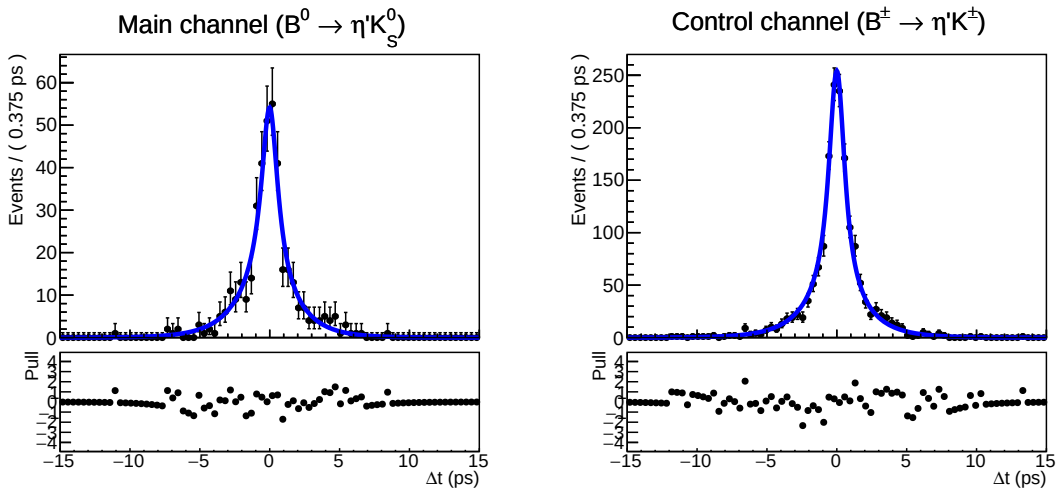


Figure 6.3: Δt distribution of generic MC sample and fitted curve of P_{Lifetime}.

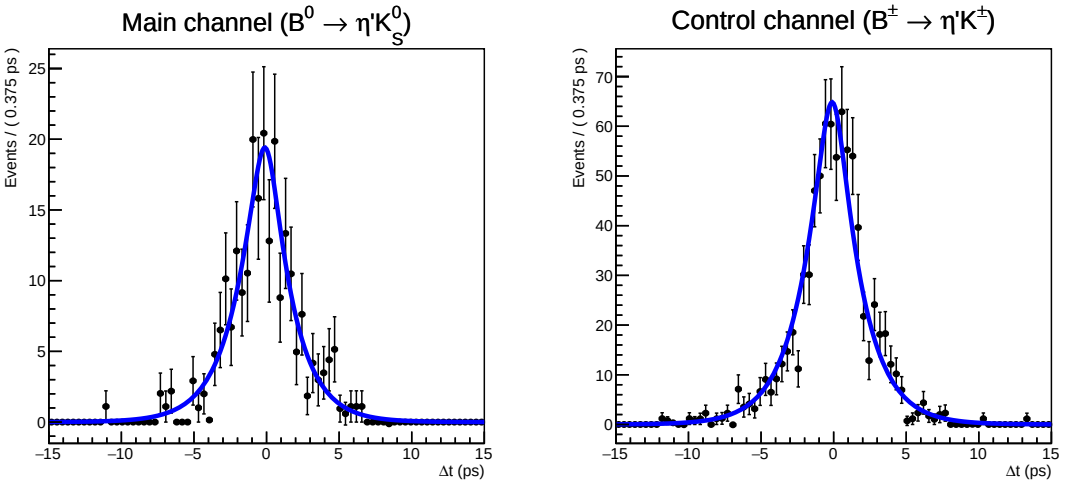


Figure 6.4: The background-subtracted Δt distribution of generic MC sample using *sPlot* and a curve of signal-only PDF with the fitted parameter τ_B.

6.1.3 Experimental data

We performed the lifetime estimation in the experimental data to check the validity of our analysis procedure. Parameters for the resolution functions are calibrated using a cosmic sample, as discussed in Ch. 5. The determined lifetime from experimental data is $\tau_{B^0} = 1.553 \pm 0.259\,\text{ps}$ and $\tau_{B^\pm} = 1.516 \pm 0.115\,\text{ps}$, and Δt distribution and fitted curve are shown in Figs. 6.5 and 6.6($s\mathcal{P}lot$). The resultant lifetime for the neutral and charged B-mesons is consistent with PDG average lifetime ($\tau_{B^0} = 1.519\,\text{ps}$ and $\tau_{B^\pm} = 1.638\,\text{ps}$)[11] and the curves fit well with data points.

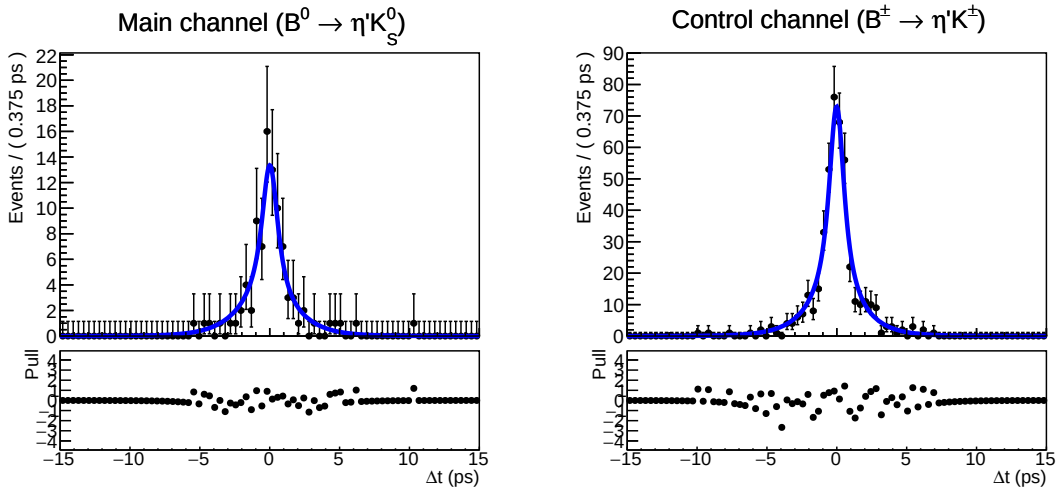


Figure 6.5: Δt distribution of experimental data of full dataset and fitted curve of P_{Lifetime}.

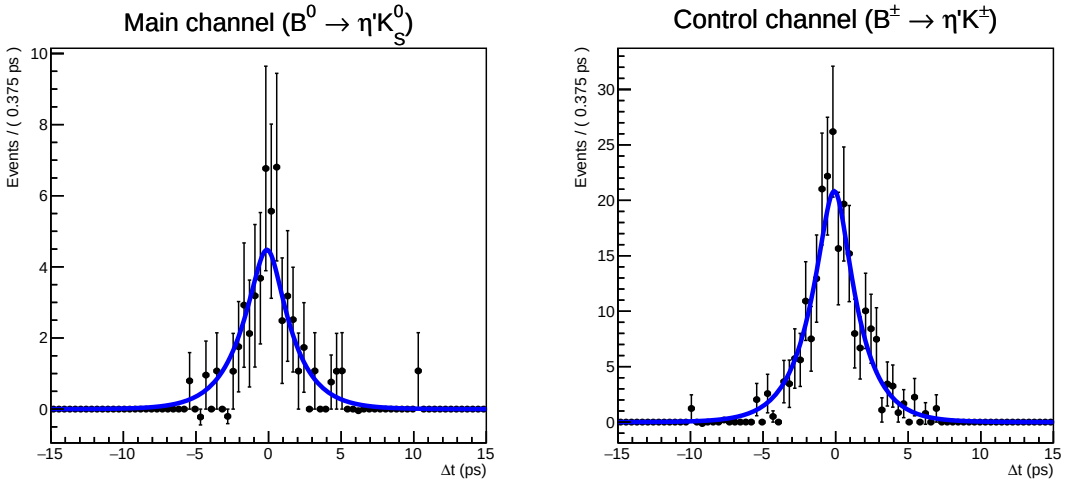


Figure 6.6: The background-subtracted Δt distribution of experimental data of full dataset using $s\mathcal{P}lot$ and a curve of signal-only PDF with the fitted parameter τ_B.

6.2 *CP* fitting

Before starting the determination of *CP* asymmetries, we should build a PDF for the *CP* fitting. Basically, we can think that the probability $P_{CP}^{true}(\Delta\tau, q_{tag}^{true})$ for our analysis is directly proportional to the decay width given in 3.8. Thus, we can write the probability with the true information $(\Delta\tau, q_{tag}^{true})$ by considering the PDF normalization as follows:

$$P_{CP}^{true}(\Delta\tau, q_{tag}^{true}|A_{CP}, S_{CP}) = \frac{1}{4\tau_B} e^{\frac{-|\Delta\tau|}{\tau_B}} \left(1 + q_{tag}^{true}\left(\mathcal{A}sym(\Delta\tau)\right)\right) \tag{6.2}$$

where $\mathcal{A}sym(\Delta\tau)$ is given in Eq. 1.28 and q_{tag}^{true} can be +1 for true-$B^0_{\rm tag}$ or -1 for true-$\overline{B}^0_{\rm tag}$.

Let us consider the wrong tag probability of the flavor tagging for $B_{\rm tag}$. Equation 6.2 can be rewritten as $P_{CP}^{true}(\Delta\tau, q_{tag}^{true}) = (1/(4\tau_B)) \cdot \exp(-|\Delta\tau|/\tau_B) \cdot p_{CP}^{true}(q_{tag}^{true})$, where $p_{CP}^{true}(q_{tag}^{true}) = (1 + q_{tag}^{true}\mathcal{A}sym)$. $\mathcal{A}sym(\Delta\tau)$ is abbreviated to $\mathcal{A}sym$ for simplicity. Then, a probability $P_{CP}^{FT}(\Delta\tau, q_{tag})$ of observing the reconstructed flavor q_{tag} from Flavor Tagger is given by $P_{CP}^{FT}(\Delta\tau, q_{tag}) = (1/(4\tau_B)) \cdot \exp(-|\Delta\tau|/\tau_B) \cdot p_{CP}^{FT}(q_{tag})$, where $p_{CP}^{FT}(q_{tag})$ is

$$\begin{aligned} p_{CP}^{FT}(q_{tag}) =& \epsilon_{q_{tag}}(1 - w_{q_{tag}}) \cdot p_{CP}^{true}(q_{tag}^{true} = q_{tag}) + \epsilon_{-q_{tag}} w_{-q_{tag}} \cdot p_{CP}^{true}(q_{tag}^{true} = -q_{tag}) \\ =& \epsilon_{q_{tag}}(1 + q_{tag}\mathcal{A}sym) - \epsilon_{q_{tag}} w_{q_{tag}}(1 + q_{tag}\mathcal{A}sym) + \epsilon_{-q_{tag}} w_{-q_{tag}}(1 - q_{tag}\mathcal{A}sym), \end{aligned} \tag{6.3}$$

and the convention of q_{tag} is exactly the same as q_{tag}^{true}, but it indicates the "reconstructed" information by Flavor Tagger.

We can convert $w_{q_{tag}} = w + q_{tag}\frac{\Delta w}{2}$ and $\epsilon_{q_{tag}} = \epsilon(1 + q_{tag}\mu)$ by using Eqs. 3.10 and 3.11, and Eq. 6.3 becomes

$$\begin{aligned} p_{CP}^{FT}(q_{tag}) =& \epsilon_{q_{tag}}(1 + q_{tag}\mathcal{A}sym) - \epsilon_{q_{tag}} w_{q_{tag}}(1 + q_{tag}\mathcal{A}sym) + \epsilon_{-q_{tag}} w_{-q_{tag}}(1 - q_{tag}\mathcal{A}sym) \\ =& \epsilon((1 + q_{tag}\mu)(1 + q_{tag}\mathcal{A}sym) - (1 + q_{tag}\mu)(w + q_{tag}\frac{\Delta w}{2})(1 + q_{tag}\mathcal{A}sym) \\ & + (1 - q_{tag}\mu)(w - q_{tag}\frac{\Delta w}{2})(1 - q_{tag}\mathcal{A}sym)) \\ =& \epsilon((1 + q_{tag}\mu)(1 + q_{tag}\mathcal{A}sym) - q_{tag}((2w\mu + \Delta w) + (2w + \mu\Delta w)\mathcal{A}sym)) \\ =& \epsilon((1 - q_{tag}\Delta w) + q_{tag}(1 - 2w)\mathcal{A}sym + \mu(\mathcal{A}sym + q_{tag}(1 - 2w - \Delta w\mathcal{A}sym)) \end{aligned} \tag{6.4}$$

where ϵ is a total flavor tagging efficiency and is assumed to be 1. In our analysis, μ is considered to be zero because it is zero-consistent at most of r-bin within its uncertainty. The effect of μ will be considered in the systematic study. Then, we can get an approximated form of $P_{CP}^{FT}(\Delta\tau, q_{tag}) = (1/(4\tau_B)) \cdot \exp(-|\Delta\tau|/\tau_B) \cdot p_{CP}^{FT}(q_{tag})$ as

$$\begin{aligned} P_{CP}^{FT}(\Delta\tau, q_{tag}|A_{CP}, S_{CP}) =& \frac{1}{4\tau_B} e^{\frac{-|\Delta\tau|}{\tau_B}} \left((1 - q_{tag}\Delta w) + q_{tag}(1 - 2w)\mathcal{A}sym(\Delta\tau)\right) \\ =& \frac{1}{4\tau_B} e^{\frac{-|\Delta\tau|}{\tau_B}} \left((1 - q_{tag}\Delta w) + q_{tag}(1 - 2w)\left(A_{CP}\cos(\Delta m\Delta\tau) + S_{CP}\sin(\Delta m\Delta\tau)\right)\right). \end{aligned} \tag{6.5}$$

Next, we calculate the probability with consideration of the incompleteness of the vertex determination as follows:

$$\begin{aligned} P_{CP}^{sig}(\Delta t, q_{tag}|A_{CP}, S_{CP}, S_{CP}) &= P_{CP}^{FT} \otimes R \\ &= \int_{-\infty}^{\infty} d(\Delta\tau) P_{CP}^{FT}(\Delta\tau, q_{tag}|A_{CP}, S_{CP}) R(\Delta t - \Delta\tau) \end{aligned} \tag{6.6}$$

where $R(\Delta t)$ is the resolution shown in Eq. 5.1. τ_B and Δm values are fixed to the PDG-averaged ones to the corresponding lifetimes (τ_{B^0} or $\tau_{B^\pm}$), and the mixing frequency of B-meson ($\Delta m_d = 0.507/\text{ps}$)[11], respectively.

Finally, we build the entire PDF for the *CP* fitting to A_{CP} and S_{CP} by including the effect of events in the Background category. The PDF is given as

$$P_{CP}(\Delta t, q_{tag}|A_{CP}, S_{CP}) = f_{sig} P_{CP}^{sig}(\Delta t, q_{tag}|A_{CP}, S_{CP}) + \frac{1 - f_{sig}}{2} P_{bckg}(\Delta t), \tag{6.7}$$

where f_{sig} is event-by-event signal probability defined in Eq. 4.5 and $P_{bckg}(\Delta t)$ is the Δt PDF for Background category given in Eq. 5.21.

6.2.1 Fit to the signal MC sample

We fitted the signal MC samples as a test of resolution function, wrong-tag probability given in the "Signal MC" column of Tab. 4.1, and the PDF given in Eq. 6.6. For the *CP* fitting, we used all reconstructed events of the (Signal + SxF) category from the entire signal MC samples. The results in Tab. 6.2 indicate good agreement with the simulation inputs of $(A_{CP}, S_{CP}) = (0, 0.703)$ for the main channel and zero asymmetries for the control channel. Figure 6.7 shows Δt distribution and fitted curve with separation of q_{tag} and asymmetry plots with error bar calculated with Clopper-Pearson method. The curve fits well with the data points from the signal MC samples.

Samples	Channels	A_{CP}	S_{CP}
Signal MC	Main	$-0.015^{+0.005}_{-0.005}$	$0.707^{+0.006}_{-0.006}$
	Control	$-0.004^{+0.005}_{-0.005}$	$0.007^{+0.006}_{-0.006}$
Generic MC 1/ab	Main	$0.085^{+0.135}_{-0.137}$	$0.712^{+0.171}_{-0.183}$
	Control	$-0.127^{+0.079}_{-0.079}$	$-0.040^{+0.105}_{-0.104}$

Table 6.2: Determined *CP* parameters from 1/ab generic or signal MC samples.

6.2.2 Fit to generic MC sample

We fitted the generic MC sample to extract the *CP* asymmetry (A_{CP}, S_{CP}) with the PDF given in Eq. 6.7. Table 6.2 shows the results, and they are consistent with input values for simulation. Figure 6.8 shows Δt distribution and fitted curves with background components, and 6.9 shows the distribution and PDFs without backgrounds using *sPlot*.

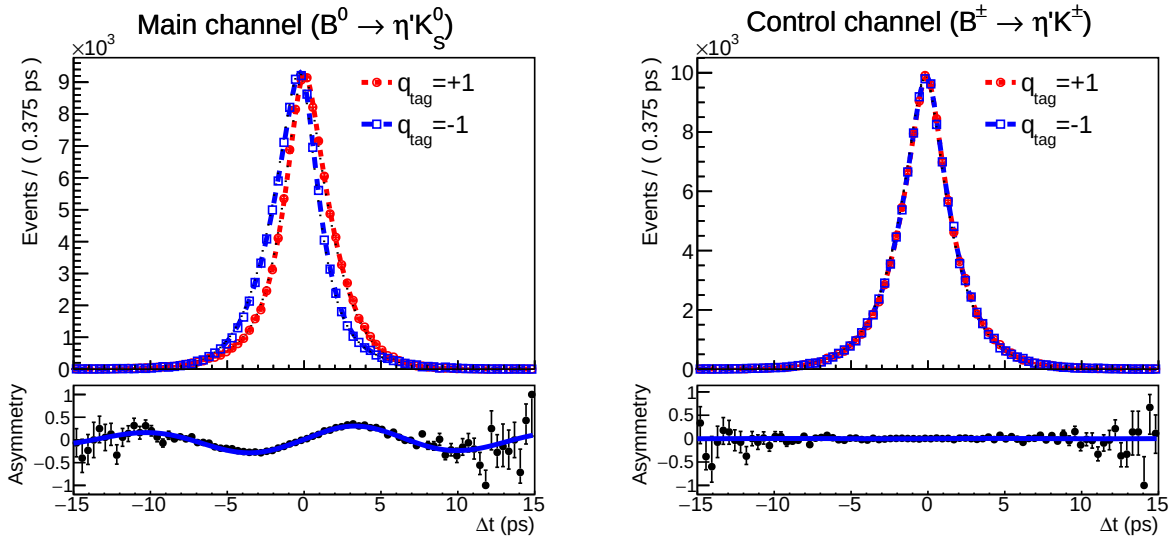


Figure 6.7: Δt distribution with q_{tag} separation from the signal MC samples and fitted curve for each q_{tag} case. The plots below show asymmetry distribution(black dot) and curves(solid blue line).

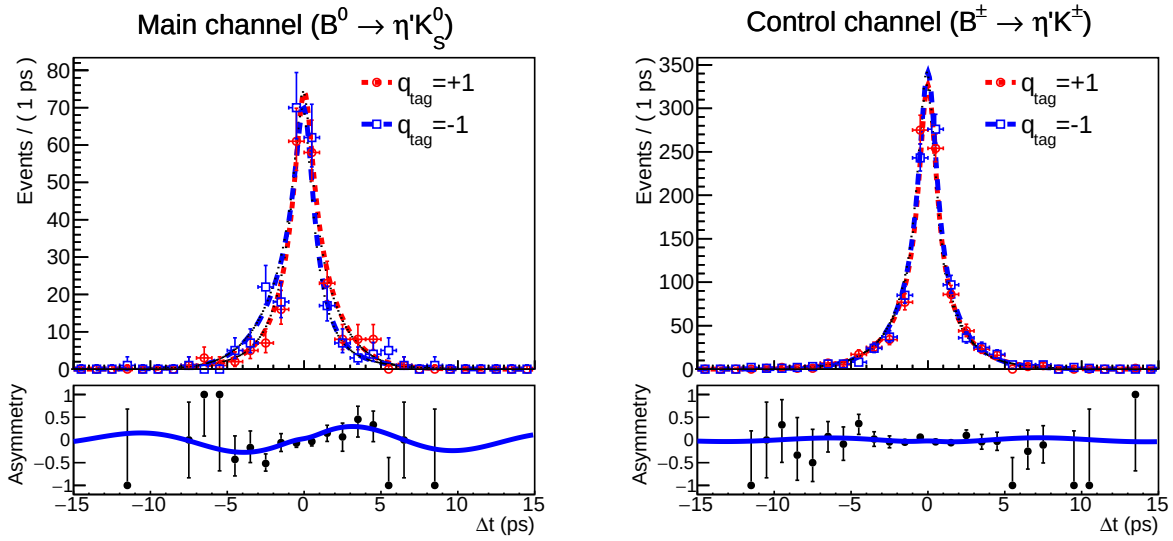


Figure 6.8: Δt distribution and fitted curve from generic 1/ab MC sample.

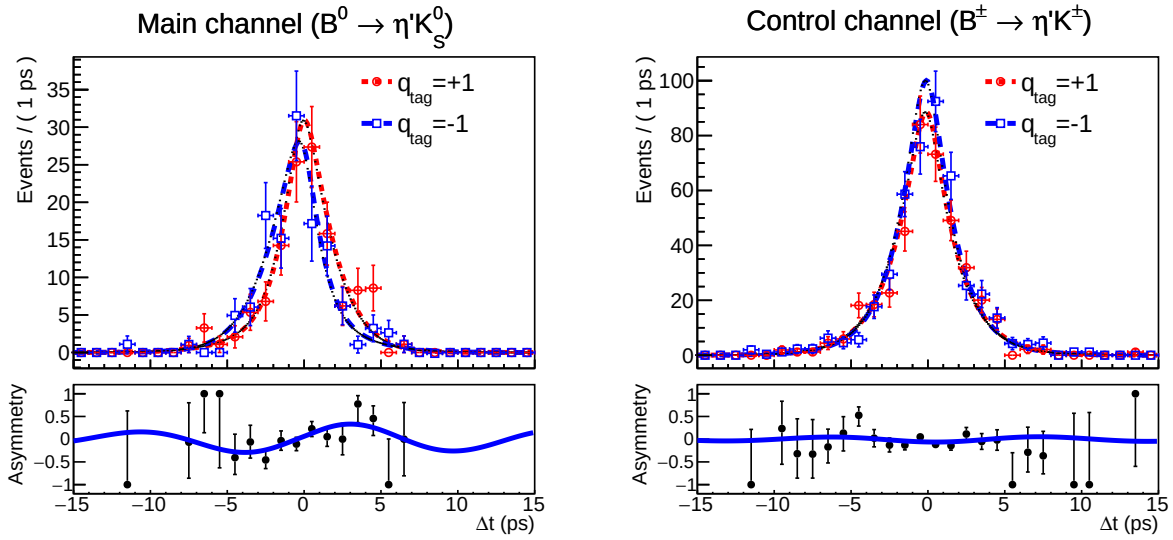


Figure 6.9: Δt distributions of the (Signal + SxF) component using *sPlot* approach and the fitted curve.

6.2.3 Linearity tests with special MC samples

We generated special MC simulation samples with various input values for *CP*-violating parameters. Input parameters of $(A_{CP}^{input}, S_{CP}^{input})$ are $(-1.0, 0), (-0.8, 0), ..., (0.8, 0), (1.0, 0)$, $(0, -1.0), (0, -0.8), ..., (0, 0.8), (0, 1.0)$, and additional points, where A_{CP} and S_{CP} are not zero, $(-0.1, -0.1), (-0.1, 0.1), (0.1, -0.1), (0.1, 0.1)$.

According to Fig. 6.10 and Tab. 6.3, we can confirm the *CP* fitting procedure reproduced the inputs well. The result indicates good linearity, but the p-value of a sub-figure of "A_{CP}^{fit} with $S_{CP}^{input} = 0$ fix" in Fig. 6.10 is lower than 0.003 (3σ), and the result from linearity tests using Toy MC dataset with non-zero μ given in App. I shows a bias in A_{CP}^{fit}. We suspect the bias is due to ignorance of an asymmetry of flavor tagging efficiencies μ. We will examine the effects of μ on extracted *CP* asymmetries in the systematic study.

$(A_{CP}^{input}, S_{CP}^{input})$	A_{CP}^{fit}	S_{CP}^{fit}
$(0.1, 0.1)$	$0.112^{+0.015}_{-0.015}$	$0.121^{0.022}_{-0.022}$
$(0.1, -0.1)$	$0.097^{+0.016}_{-0.016}$	$-0.101^{+0.023}_{-0.023}$
$(-0.1, 0.1)$	$-0.126^{0.016}_{-0.016}$	$0.081^{+0.023}_{-0.023}$
$(-0.1, -0.1)$	$-0.106^{+0.016}_{-0.016}$	$-0.123^{+0.023}_{-0.023}$

Table 6.3: Results of *CP* fitting with generated signal MC samples with both-nonzero $(A_{CP}^{input}, S_{CP}^{input})$ configurations.

6.2.4 Ensemble tests

As another test for our *CP* fitter, we performed an ensemble test only for the main channel. A toy MC dataset for the test is generated with the following procedure:

- Set inputs for pre-parameters of A_{CP}, S_{CP}, N_{sig}, and N_{bckg}.
- Set pool dataset to generate conditional variables of $\left(\frac{\chi^2}{ndf}\right)^{CP,tag}, \sigma_t^{CP,tag}$,for the (Signal + SxF) and Background categories. (Appendix A)
- For the signal events, Generate t_{CP} and t_{tag} based on exponential decay PDF with τ_B to decide Δt_{true} and T in Eq. 5.5.
- Generate true-q_{tag} with Δt_{true} from the previous step and determine r-bin and reconstructed-q_{tag} value based on $\epsilon_{B^0,\bar{B}^0}$ and $w_{B^0,\bar{B}^0}$ that can be calculated using ϵ, μ(=0 for nominal ensemble tests), w, and Δw information.
- Calculate effects of the kinematical approximation and generate Δt residual using conditional variables for vertex fitting to determine reconstructed-Δt value.

- For the background events, generate Δt value based on $P_{bckg}(\Delta t)$. and $q_{tag} = +1$ or -1 with half-half probability. $|q \cdot r|$ value is generated based on the histogram PDF.

- Generate signal extraction variable of $(M_{bc}, \Delta E, \mathcal{O}_{CS}, \cos(\theta_B^{Boost}))$ independently using PDFs for the signal extraction. This procedure is repeated N_{sig} or N_{bckg} times for each category.

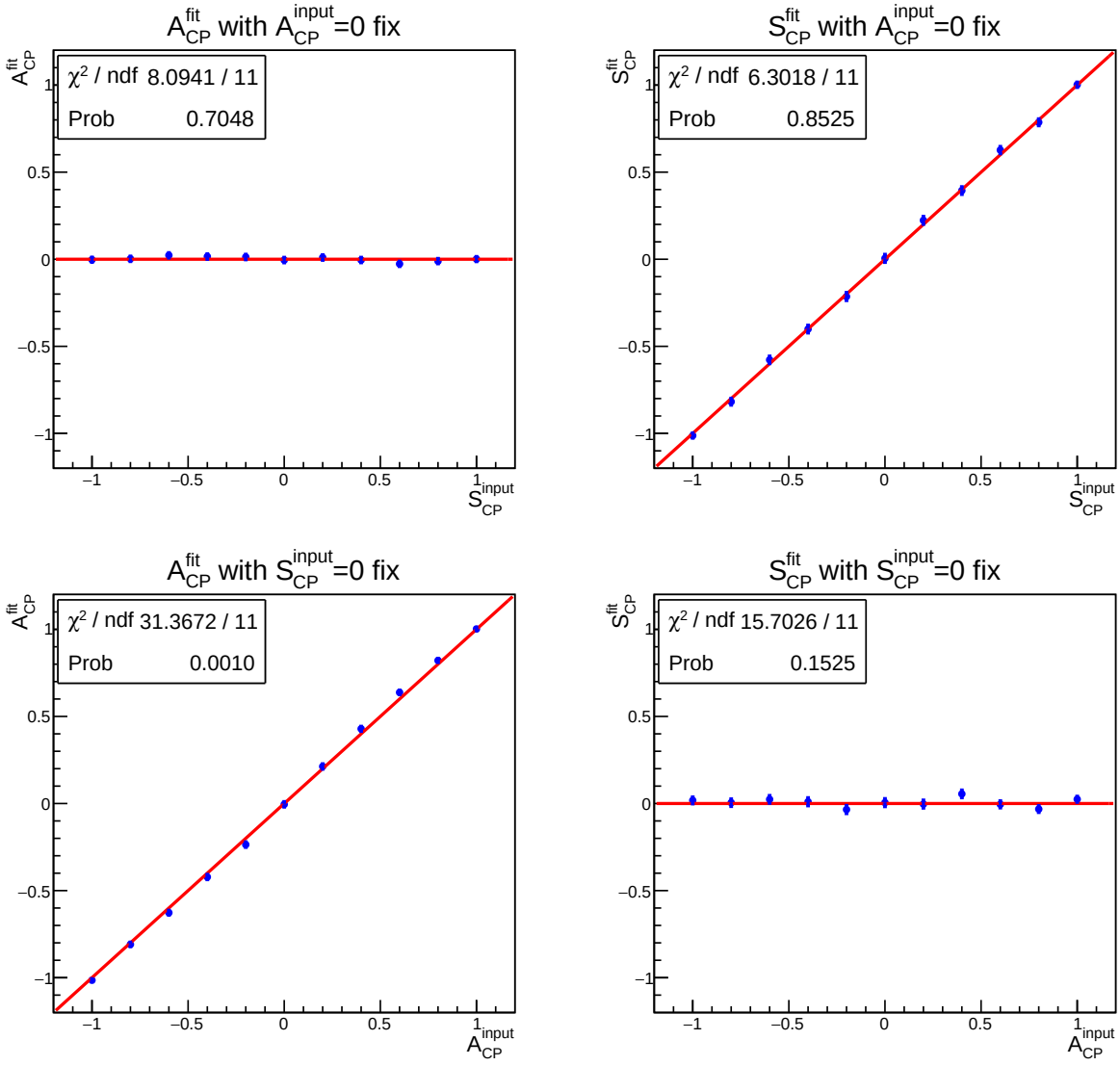
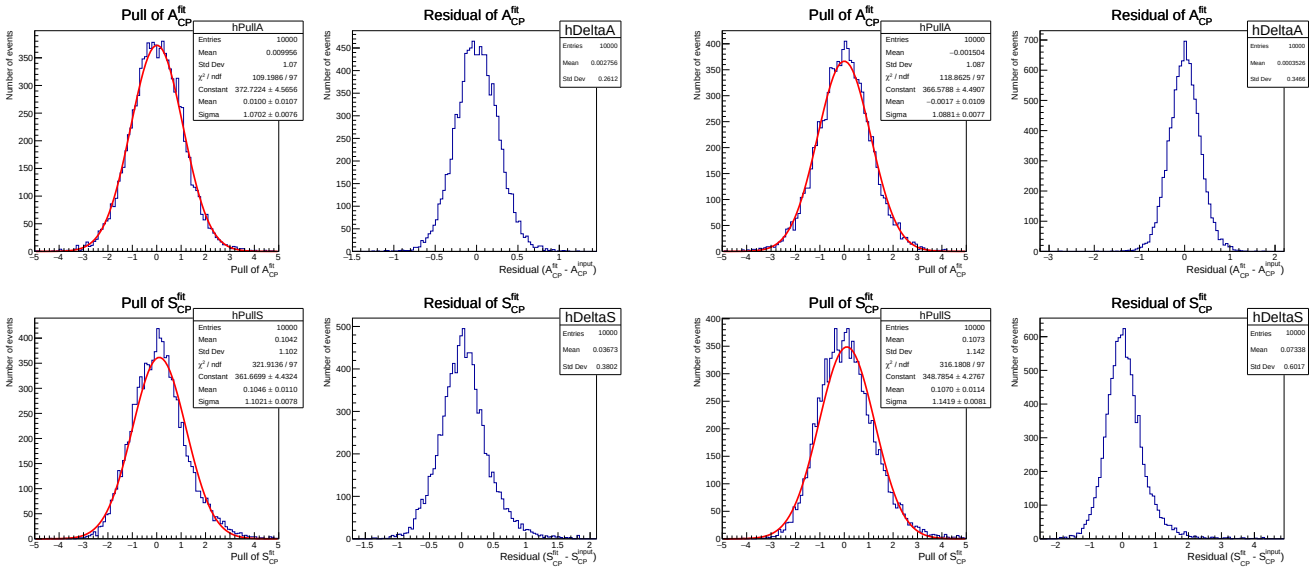


Figure 6.10: Results of the linearity test with specially generated signal MC samples with various A_{CP}^{input} and S_{CP}^{input}. Each sample in a specific input has about 26000 signal events, corresponding to about $80\,\mathrm{ab}^{-1}$, and blue bars show the CP fitting results with uncertainty. The solid red line shows the reference line of $x = y$ or $y = 0$, and the blue points with the error bar are fitting results for each input. χ^2/ndf or "Prob." value in the box indicates corresponding values with given lines and data points.

In the ensemble test, we examined two configurations for the test:

- Single test: A test with Toy MC dataset of 10000 samples with $A_{CP} = 0$, $S_{CP} = 0.703$, and N_{sig} and N_{bckg} corresponding to the amount of 400/fb and 4/ab from generic MC, and yields from signal extraction on the experimental data.
- Linearity test: A test with varying input parameters of A_{CP} and S_{CP} as $(-1, 0)$,$(-0.8, 0)$, ...,$(0.8, 0)$, $(1.0, 0)$ and $(0, -1), (0, -0.8), ..., (0, 0.8), (0, 1.0)$. Ten thousand samples are generated and fitted with configurations from the experimental data. The number of events is set by ten times N_{total} from the signal extraction for Data.

In the single test, the *CP* fitting all Toy MC samples succeeded, and Fig. 6.11 shows the results. We confirmed a bias and wrong uncertainty evaluation in the single test as well as in the linearity test using the Toy MC dataset shown in sub-figures in Fig. 6.12. We consider this due to the low statistics since other analysis shows a similar problem[45], and tests using the Toy MC dataset of large statistics show the bias and underestimation of errors are relieved.

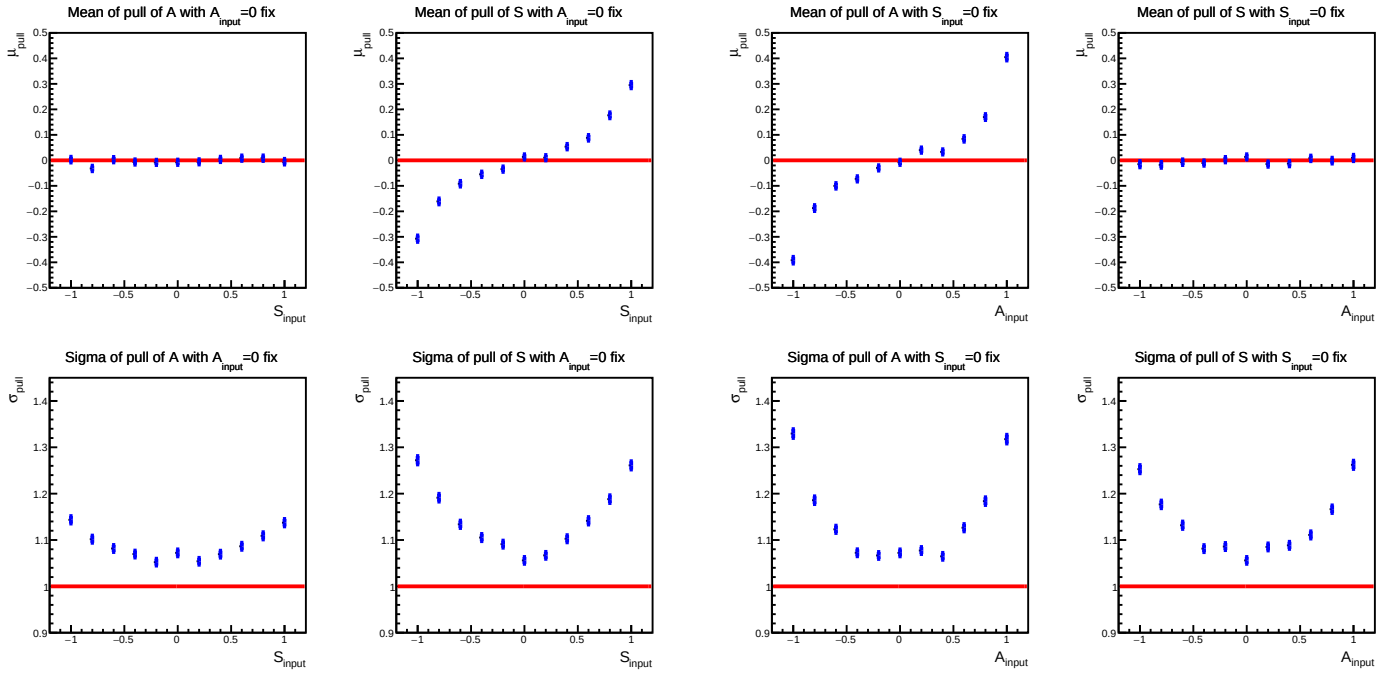


(a) 400/fb (b) Full dataset(362/fb) of experimental data

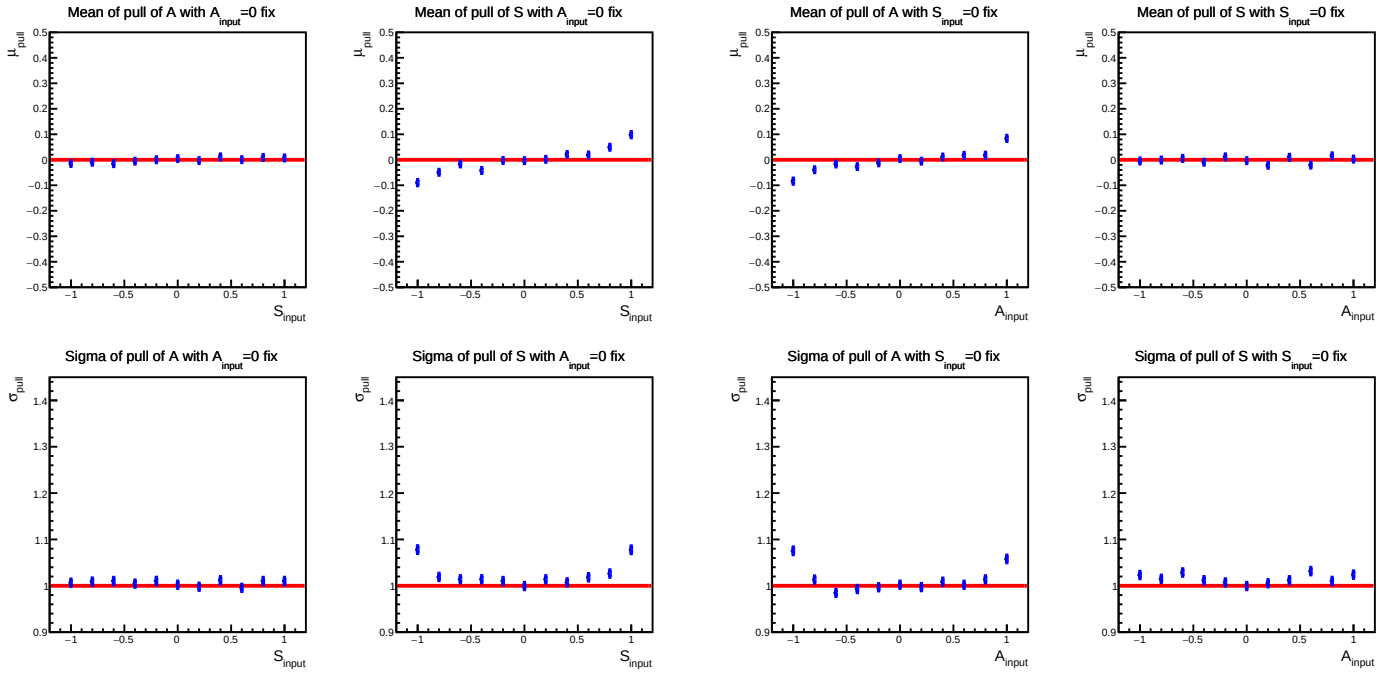
Figure 6.11: Distributions of pull and absolute bias for *CP* asymmetries from the single test with N_{sig}, N_{bckg} for various statistics equivalent. Toy MC with full dataset statistics is performed with configurations from the experimental data, and pull distributions are fitted with Gaussian.

6.2.5 Fit to the experimental data in the control channel

We extracted the *CP* asymmetries from the control channel with the experimental data as validation that our procedure also works well in the data. The extracted $A_{CP} =$



(a) Result with N_{total} from the current statistics from the experimental data ($362\,\mathrm{fb}^{-1}$). The result is biased, and its uncertainty is underestimated on both end sides.



(b) Result with $10 \cdot N_{total}$ from the current statistics from the experimental data ($3.62\,\mathrm{ab}^{-1}$). Underestimation of uncertainty and bias is largely resolved, but there still is a problem with $A_{CP}, S_{CP_{input}}$ around the physical boundary.

Figure 6.12: Linearity test using Toy MC. Blue points with error bars indicate fitting results, and solid red lines show reference values which $y = 0$ for the mean(μ) and $y = 1$ for standard deviation(σ).

$-0.091^{+0.156}_{-0.155}$ and $S_{CP} = -0.048^{+0.221}_{-0.218}$ which is consistent with the zero asymmetries as we expected. The distribution of Δt and fitted curve with separation of q_{tag} is shown in Fig. 6.13.

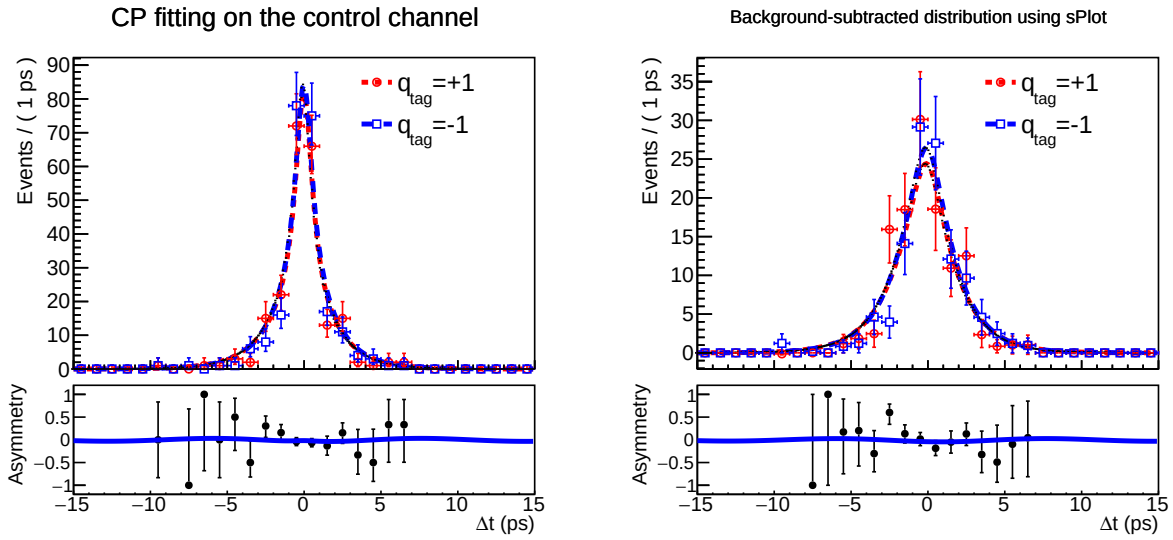


Figure 6.13: Δt distribution with q_{tag} separation of the control channel from the experimental data and fitted curve for each q_{tag} case. The plots below show asymmetry distribution(black dot) and curves(solid blue line).

6.2.6 *CP* fitting result

We performed the *CP* fitting using the full dataset of the experimental data. The estimated *CP* asymmetries from the fitting is $A_{CP} = -0.111^{+0.319}_{-0.311}$ and $S_{CP} = 0.248^{+0.470}_{-0.526}$. Figure 6.14 shows Δt distribution and the fitted curve, and Fig. 6.15 shows the same thing but separated one by the dilution factor from Flavor Tagger to see the asymmetry more clearly.

6.3 Systematic uncertainties

We examined the uncertainty due to factors inherent in our estimation procedure. We will discuss details on the factors and the effects on the *CP* fitting results from them in the following subsections.

6.3.1 Fit bias

We assume there is no correlation between M_{bc}, ΔE, and $\mathcal{O}_{CS}$ used for the f_{sig} calculation. However, we confirmed the correlation between M_{bc} and ΔE as shown in a plot on the

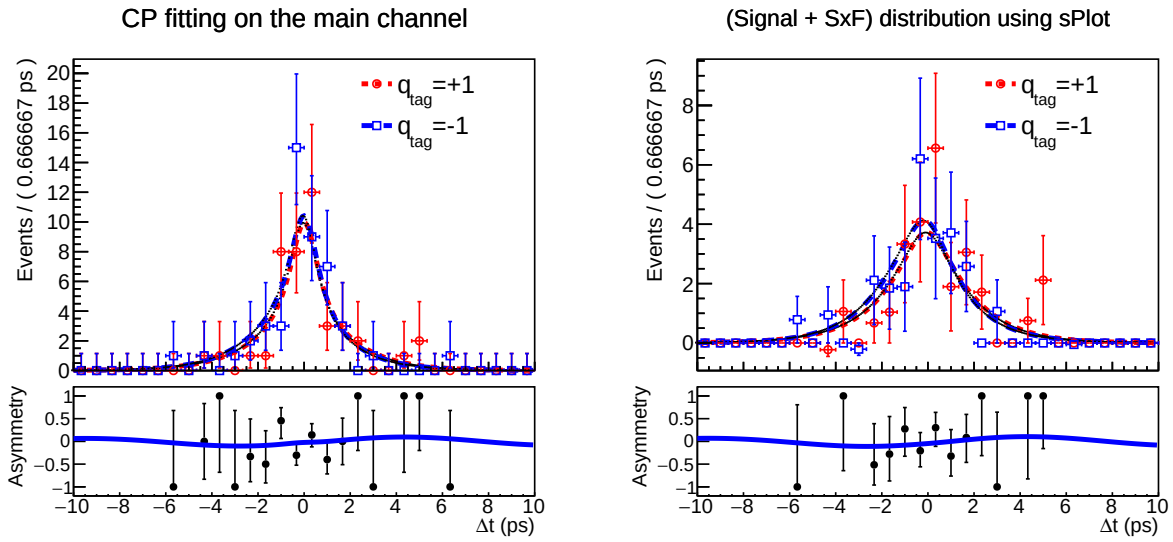


Figure 6.14: Δt distribution with q_{tag} separation of the main channel from the experimental data and fitted curve for each q_{tag} case. The plots below show asymmetry distribution(black dot) and curves(solid blue line).

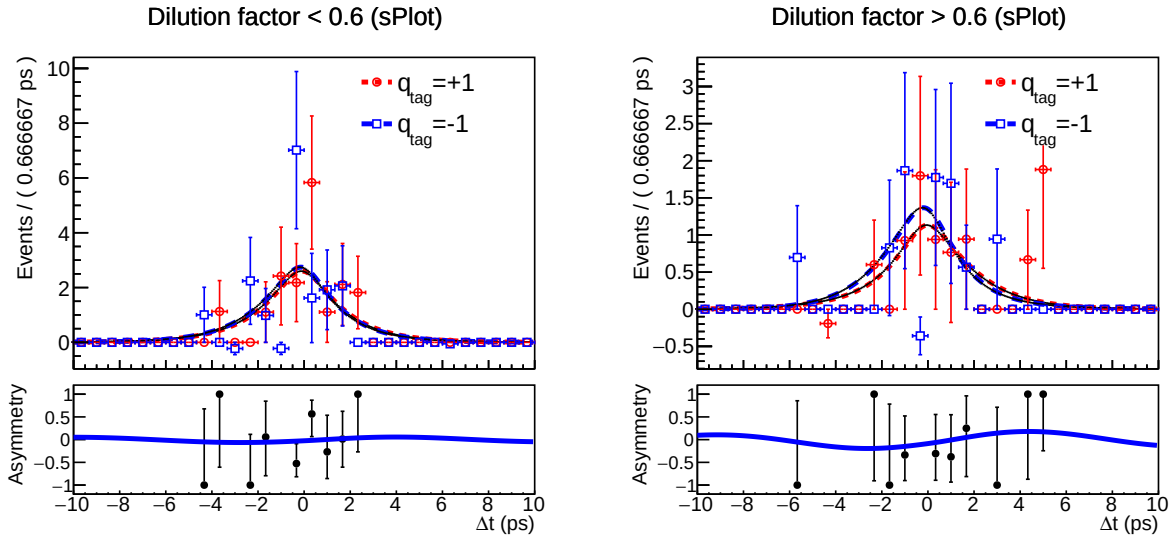


Figure 6.15: Δt distribution with q_{tag} separation of the main channel from the experimental data and fitted curve for each q_{tag} case. The plots below show asymmetry distribution(black dot) and curves(solid blue line).

right side of Fig. B.2. Also, $\mathcal{O}_{CS}$ and dilution factor($|q_{tag} \cdot r|$) from the signal events are also correlated as shown in Fig. 6.16.

We also suppose that the shape of a distribution for $\left(\frac{\chi^2}{ndf}\right)^{CP,tag}$ is the same over (Signal + SxF) and Background categories. Still, we confirmed that the shapes are different, as shown in Figs. F.14 and F.16 so Punzi effects can occur in the *CP* fitting procedure.

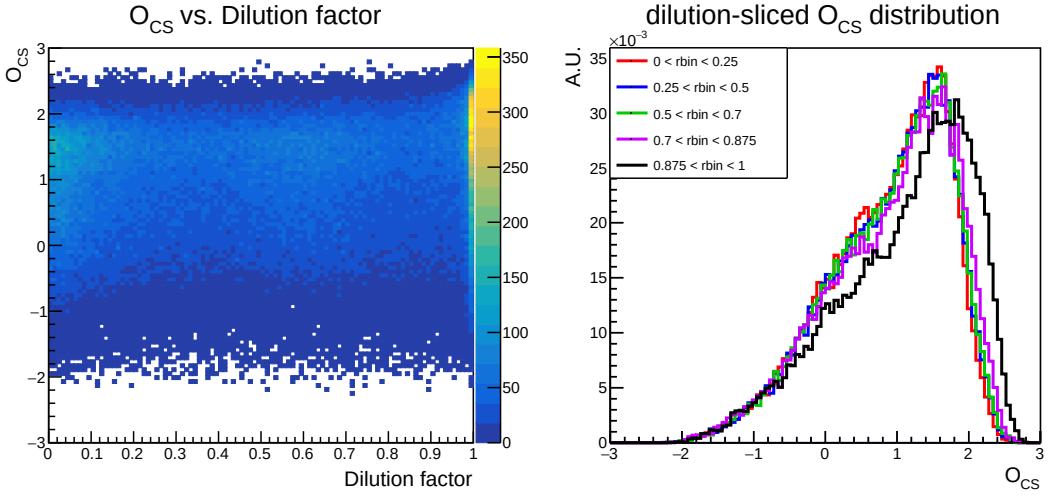


Figure 6.16: 2D scatter plot of $\mathcal{O}_{CS}$ vs. dilution factor (left) and dilution-sliced $\mathcal{O}_{CS}$ distribution (right) from signal events in the main channel

We examined the effects of these correlations on the *CP* fitting using an ensemble test with 4/ab-statistics equivalent N_{total}. A method to generate Toy MC samples that have the correlation is described in App. A. The bias due to this factor is $\delta A = +0.0076$ and $\delta S = +0.0029$.

6.3.2 Parameters for Flavor Tagger

There are two sources for this category: Asymmetries of flavor tagging efficiency and uncertainty of wrong tag probability.

(1) Asymmetry of the efficiency of Flavor Tagger (μ)

Flavor Tagger might respond differently when the true-flavor of B^0 on the tag-side q_{tag}^{true} is +1 or -1 due to the response asymmetry on charged particles of the Belle II detector or any other reason. A parameter of μ indicates an asymmetry of efficiency of the flavor tagging and is given in Eq. 3.11. The PDF for *CP* fitting given in Eq. 6.6 does not consider μ term of $\mu(\mathcal{A}sym + q_{tag}(1-2w-\Delta w\mathcal{A}sym))$ in Eq. 6.4. As both of μ and Δw are small ($\sim \mathcal{O}(1\%)$) so we can neglect $q_{tag}\mu\Delta w\mathcal{A}sym$ term, then we only have terms of $\mu\mathcal{A}sym + \mu q_{tag}(1-2w)$. If we calculate the asymmetry as given in Eq. 1.28 using Eq. 6.4, then $\mu\mathcal{A}sym$ term is canceled out in numerator, and the $\mu(1-2w)$ term has dominant effects on the A_{CP} and S_{CP} measurement. Since the term is not depending on $\mathcal{A}sym(\Delta t)$ and adds some offset on the PDF, we can expect this factor will affect A_{CP} mainly. This can be confirmed in a linearity test with a non-zero μ value stated in App. I.

We generated 10000 toy MC samples with $(A_{CP}^{input}, S_{CP}^{input}) = (0, 0.703)$ by determining reconstructed-q_{tag} with non-zero μ in the dataset generation and fitted it with the nominal

CP fitter to examine the absolute bias due to this factor. The bias due to this factor is $\delta A = -0.0052$ and $\delta S = -0.0019$.

(2) Uncertainty of wrong tag probability $(w, \Delta w)$

Wrong tag probability and its asymmetry $(w, \Delta w)$ are determined by analyzing hadronic channels with high statistics. A study for this factor considers a deviation on measured *CP* asymmetries due to the uncertainty of w and Δw. We evaluated the deviation by fluctuating w and Δw by their uncertainty given in the "Data" column of Tab. 4.1 and performed *CP* fit with the changed parameters. This procedure was repeated 1000 times with the experimental data, and a mean value of a set of $(A_{CP}^{fit}, S_{CP}^{fit})$ was calculated. We have taken the standard deviation value as an error for this factor. The value is $\delta A = \pm 0.0090$ and $\delta S = \pm 0.0088$.

6.3.3 Pre-fit parameters

CP fitting procedure requires many pre-fit parameters determined using MC samples or sideband events from the experimental data. The parameters have uncertainty, and we consider the deviation by this factor. There are two factors in this category for systematic uncertainty, which are parameters of signal PDF and background Δt PDF.

(1) Signal PDFs for signal extraction

We changed the parameters of the template PDF for the (Signal + SxF) category, which are fixed in the signal extraction. The parameters fluctuated by their uncertainty from the pre-fitting shown in Tab. 4.4 independently and performed the signal extraction. Then, the *CP* fitting on experimental data was performed with f_{sig} calculated results of the signal extraction with fluctuated parameters. This procedure was repeated 1000 times to get the deviation from the fluctuation, which is the systematic uncertainty for this factor. The deviation is $\delta A = \pm 0.0029$ and $\delta S = \pm 0.0052$.

(2) Background Δt PDF

Parameters for $P_{bckg}(\Delta t)$ were determined by fitting Δt distribution from events in the sideband region of experimental data, and they have uncertainty as shown in Tab. 5.4. The overall procedure is similar to that for the systematic study of Signal PDFs, and the deviation is $\delta A = \pm 0.0083$ and $\delta S = \pm 0.0347$.

6.3.4 Physical quantities

The PDF for *CP* fitting shown in Eq. 6.6 requires the lifetime and the mixing frequency of B-mesons. We are using the average value from PDG and fluctuating the lifetime

and frequency simultaneously by the uncertainty from PDG to examine the systematic uncertainty from this factor. A methodology for evaluating systematic uncertainty for this factor is similar to that for Pre-fit parameters. The deviation is $\delta A = \pm 0.0007$ and $\delta S = \pm 0.0050$.

6.3.5 Resolution function

We considered the effects of the resolution model on *CP* asymmetries in three aspects.

(1) Bias due to the resolution model

Although fitting results from the lifetime fit using signal samples are consistent with the input value for the Belle II full simulation, the residual fitting results shown in Figs. 5.1, 5.2, 5.3, and 5.4 suggests the residual of the fitted resolution function from MC data points. Also, we did not consider the change in vertex resolution of the Combinatorial SxF category, which is merged with the Signal category. We checked a bias due to our resolution model on (A_{CP}, S_{CP}) by examining the variance of the *CP* fitting results between `mcDeltaT` and `DeltaT` in the signal MC sample. Since the difference between the two variables is the presence or absence of smearing by resolution function, we can determine the bias by comparing *CP* fitting results using the two variables. The deviation is $\delta A = -0.0103$ and $\delta S = 0.0028$.

(2) Selection criteria for the vertex quality

We are applying some selection criteria on vertex quality variables of $(\chi^2/ndf)^{CP,tag}$ and $\sigma_z^{CP,tag}$. Threshold values for these criteria are determined based on the MC study, but as shown in Figs. 5.6, 5.8, 5.5, and 5.7, the variables from MC or experimental data indicate differences. We consider the effect by moving the threshold of the selection criteria by 10% each in the experimental data and take the quadratic average of the deviations as the systematic uncertainty. Also, we calculated this value using the control channel to minimize the statistical fluctuation, and the value is $\delta A = 0.0111$ and $\delta S = 0.0273$.

(3) Non-primary component of R_{tag}

Tag-side resolution function R_{tag} contains exponential components that express the decay of long-lived intermediate states. This parameter is determined using the MC signal sample and relates to the partial decay ratio of $b \to c$ transitions with the charged final state particles. This factor considers the possible discrepancy of the ratio between MC and Data. We scaled the τ parameter of E_n, E_p in R_{tag} by transformation of $\tau' = s_{np}^{global}\tau$ where s_{np}^{global} is a scale factor, and τ is given in Eq. 5.12. The scale factor is determined by generating a random variable using a Gaussian function with $\mu = 1$ and a specific value of σ. The σ value is determined by fitting the Δt distribution of the main channel from

the experimental data by fixing $\tau_{B^0} = 1.519$ and setting the s_{np}^{global} float in the fitting, of which the value is $0.435^{+0.793}_{-0.435}$. Estimated uncertainty for (A_{CP}, S_{CP}) is $\delta A = 0.0245$ and $\delta S = 0.0154$.

(4) Uncertainty scaling of track parameters

We apply the helix uncertainty correction given in literature[45] to consider imperfect estimation of position uncertainties of hit clusters at PXD. We reconstructed experimental data without the correction and compared the estimated *CP* asymmetries between the reconstructed result and the nominal (with the correction) one. The same parameters of $R_{CP,tag}$ are used for both samples. The deviation is taken as a systematic error for this factor, and the value is $\delta A = +0.0002$ and $\delta S = +0.0003$.

6.3.6 Calibration on resolution parameters

We consider two factors for calibrating resolution parameters: the statistics of the cosmic sample and a bias in ω of the track parameters, which is explained in App. C. We repeated each factor 100 times and collected the fitted (A_{CP}, S_{CP}) values to evaluate the systematic error for each factor.

We consider the uncertainty of parameters of the pull distribution for track parameters from the cosmic sample as a systematic error. The main input for the calibration consists of five parameters of a double Gaussian to model the pull distribution obtained from the cosmic sample. We fluctuated the parameters by their uncertainty with Gaussian and repeated the calibration and *CP* fitting. The systematic error due to this factor is $\delta A = \pm 0.0102$ and $\delta S = \pm 0.0202$.

In addition, we ignored the bias of track parameters in the calibration procedure, but the bias in track parameter ω has been confirmed. Thus, we calibrated the resolution parameters with the biased track parameter ω to examine the effects of this factor on the *CP* fitting procedure. The systematic error due to this factor is $\delta A = \pm 0.0012$ and $\delta S = \pm 0.0020$.

The quadratic sum of the standard deviation of two factors is taken as a systematic uncertainty for the calibration of resolution parameters. The total deviation is $\delta A = \pm 0.0103$ and $\delta S = \pm 0.0203$.

6.3.7 Detector misalignment

The parameters that represent the alignment of the subdetectors for vertex detection, such as SVD or VXD, have uncertainty that indicates possible deviations from reality. We generated a dedicated simulation sample of signal decay in which the detector alignment has been changed to reflect the possible deviation from reality. Then, we performed *CP* fitting with the special sample to compare the *CP* fitting results with our nominal signal

samples. The difference was taken as a systematic uncertainty for this factor, and the value is $\delta A = \pm 0.0054$ and $\delta S = \pm 0.0083$.

6.3.8 IP profile

This factor considers the uncertainty of measurements for the IP profile. Dedicated signal samples are prepared to reproduce the uncertainty. The *CP* fitting with the samples is performed to measure deviations of results between a nominal sample and specially generated ones. Systematic uncertainty for this factor is $\delta A = \pm 0.0020$ and $\delta S = \pm 0.0024$.

6.3.9 Momentum scaling

We consider the uncertainty of calibration for the momentum of charged particles in the experimental data. We adjusted the momentum of tracks by $\pm 0.1\%$ and reconstructed B_{CP} with the adjusted tracks. Then, we performed *CP* fitting to extract *CP* asymmetries. The deviation of the momentum-adjusted result from the nominal one was taken as a systematic uncertainty, and its value is $\delta A = \pm 0.0013$ and $\delta S = \pm 0.0005$.

6.3.10 Tag-Side Interference(TSI)

As discussed in the experimental strategy for *CP* asymmetry measurements in Ch. 3, we assumed that $B^0_{\rm tag}$ could only go to f_{tag} and $\overline{B}^0_{\rm tag}$ can only go to $\bar{f}_{tag}$, but this is not true. For example, a self-tagged decay of $\overline{B}^0 \to D^+\pi^-$ is used as a clue to tag the flavor of $B_{\rm tag}$ in the actual measurement. However, B^0 also can decay into the final state through $b \to \bar{u}c\bar{d}$ process, which is doubly Cabibbo suppressed. This effect is called Tag-Side-Interference (TSI) and makes some changes to the measured *CP* asymmetries[56]. We utilized the Belle II official tools for TSI to evaluate this systematic error, which numerically calculates the possible deviation of A_{CP} and S_{CP} from the true A_{CP} and S_{CP} values. The official tool infers the true *CP* asymmetries from the observed ones using a numerical method. It needs parameters that indicate effects from TSI of $r_{TSI} = 0.0118$, which means an effective amplitude ratio by TSI, and $\delta_{TSI} = 273°$, which means an effective strong phase difference by TSI[57]. The uncertainty is $\delta A = -0.0045$ and $\delta S = +0.0038$.

6.3.11 Summary

We summarized the systematic uncertainty so far, and Tab. 6.4 shows the summary and total systematic uncertainty.

Systematic Sources	δA	δS
Fit bias	$+0.0076$	$+0.0029$
μ of Flavor Tagger	-0.0052	-0.0019
Uncertainty of $w, \Delta w$	± 0.0090	± 0.0088
Parameters of Signal PDF	± 0.0029	± 0.0052
Parameters of $P_{bckg}(\Delta t)$	± 0.0083	± 0.0347
Physical quantities of τ_B and Δm_d	± 0.0007	± 0.0050
Uncertainty scaling of the track parameters	$+0.0002$	$+0.0003$
Selection criteria for the vertex quality	± 0.0111	± 0.0273
Non-primary components of R_{tag}	± 0.0215	± 0.0154
Resolution calibration	± 0.0103	± 0.0203
Bias due to the resolution function	-0.0103	$+0.0028$
Misalignment	± 0.0054	± 0.0083
IP profile	± 0.0020	± 0.0024
Momentum scaling	± 0.0013	± 0.0005
Tag-side interference	-0.0045	$+0.0038$
Total	$^{+0.0319}_{-0.0310}$	$^{+0.0531}_{-0.0530}$

Table 6.4: Summary of systematic uncertainty.

Chapter 7

Discussion and Conclusion

In this section, we discuss the significance of the measurement of *CP* asymmetries given in the previous chapter, comparison of the result with the last result at Belle, and the prospect of *CP* asymmetry measurements in the $B^0 \to \eta' K^0_S$ channel at Belle II.

7.1 Confidence interval

In the previous chapter, we extracted *CP* asymmetries in the decay of $B^0 \to \eta' K^0_S$ with a subdecay of $\eta' \to \eta(\pi^+\pi^-\pi^0)\pi^+\pi^-$ from the full dataset of the Belle II experimental data as follows:

$$A_{CP} = -0.111^{+0.319}_{-0.311}(\text{stat})^{+0.032}_{-0.031}(\text{syst}), \text{ and} \quad (7.1)$$

$$S_{CP} = +0.248^{+0.470}_{-0.526}(\text{stat})^{+0.053}_{-0.053}(\text{syst}). \quad (7.2)$$

The result is consistent with the world average given in Section 1.3 and compatible with the Standard Model prediction within 1σ uncertainty. Table 7.1 compares the result with various reference values.

We draw the 2D confidence interval with Feldman-Cousins likelihood-ratio approach[58], which is given in Fig. 7.1. (See App. G) Then, we check the constraint on NP parameters with a model-independent method from the confidence interval given in the figure. We map the coordinates of NP parameters (ϕ_{NP}, δ, r), which ϕ_{NP} is the weak phase of New Physics, δ is the difference of the strong phase between SM and NP, and r is a ratio of the magnitude of amplitude between SM and NP, to A_{CP} and S_{CP} using Eqs. 1.37 and 1.38 to get the corresponding confidence level and draw the distribution of the 2-D confidence level to reject the parameters of NP. Figures 7.2 and 7.3 show the result, which excludes the presence of NP parameters with high r and δ value and implies dependence on ϕ_{NP} with lower δ region.

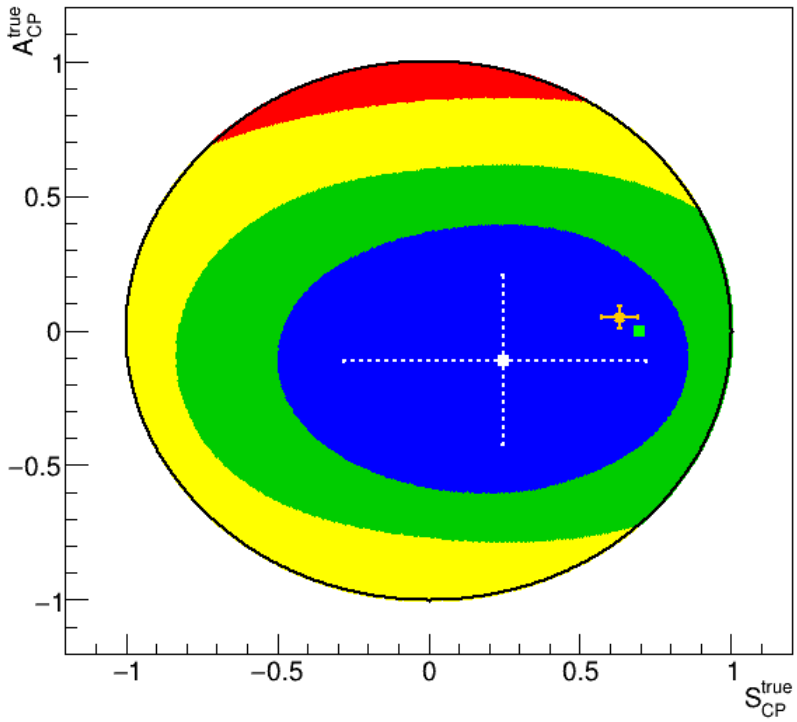


Figure 7.1: 2D confidence interval for A_{CP} and S_{CP}. The white point with the crossbar shows our measurement with statistical uncertainty, the orange cross indicates the world-average CP asymmetries in the $B^0 \to \eta' K^0$ channel, and the green rectangular point is the reference value in the charmonium decay. Legend for the color of each region is as follows: Blue is $< 1\sigma(68.27\%)$, green is $< 2\sigma(95.45\%)$, yellow is $< 3\sigma(99.73\%)$, and red is $< 4\sigma(99.99\%)$.

Experiment	Measurement	Residual
Belle[16]	$A_{CP} = 0.03 \pm 0.06$	-0.44σ
	$S_{CP} = 0.68 \pm 0.08$	-0.91σ
BaBar[17]	$A_{CP} = 0.08 \pm 0.06$	-0.60σ
	$S_{CP} = 0.57 \pm 0.08$	-0.68σ
All charmonium[19]	$A_{CP} = 0.005 \pm 0.015$	-0.36σ
(SM prediction)	$S_{CP} = 0.699 \pm 0.017$	-0.95σ

Table 7.1: Comparison of the result from this analysis from another measurement in $B^0 \to \eta' K_S^0$ channel. The Residual column indicates the residual value of our measurement to the value of each row.

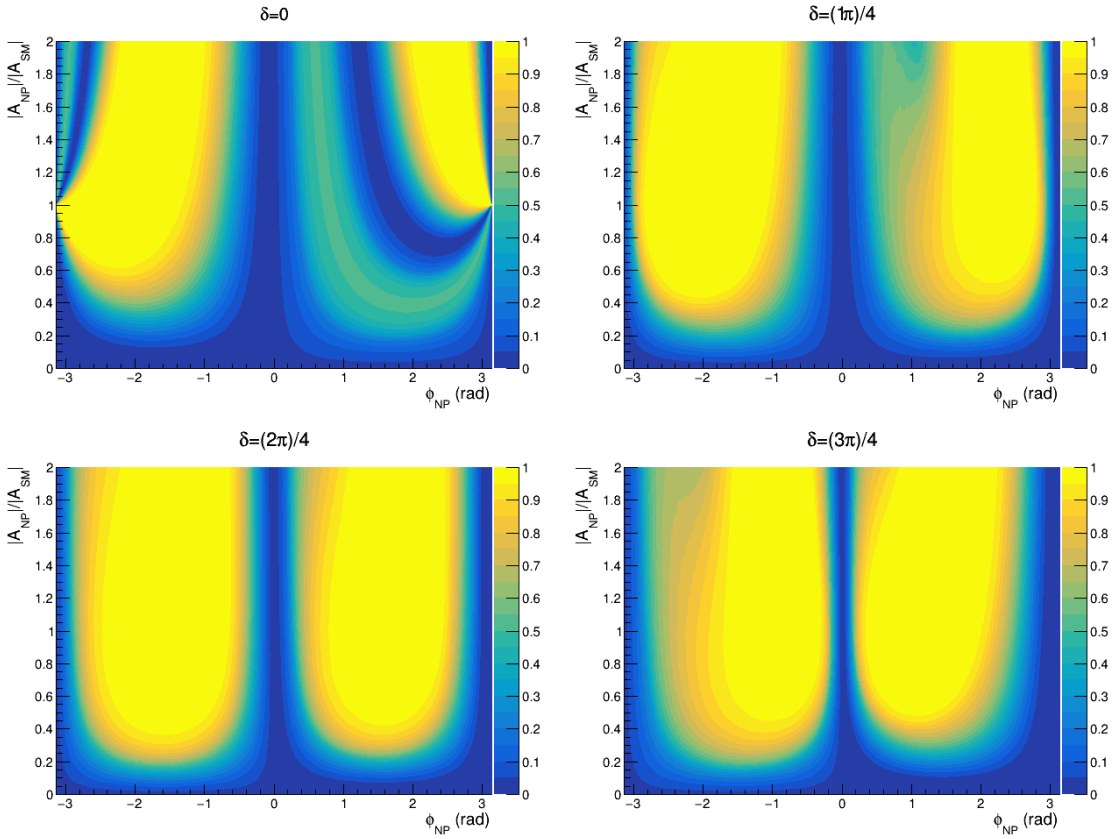


Figure 7.2: 2-D confidence level to reject the weak phase of NP (ϕ_{NP}) and the amplitude ratio of NP to SM ($r \equiv |\mathcal{A}_{NP}|/|\mathcal{A}_{SM}|$) with various values of the difference of strong phase between NP and SM (δ). The color shows the confidence level.

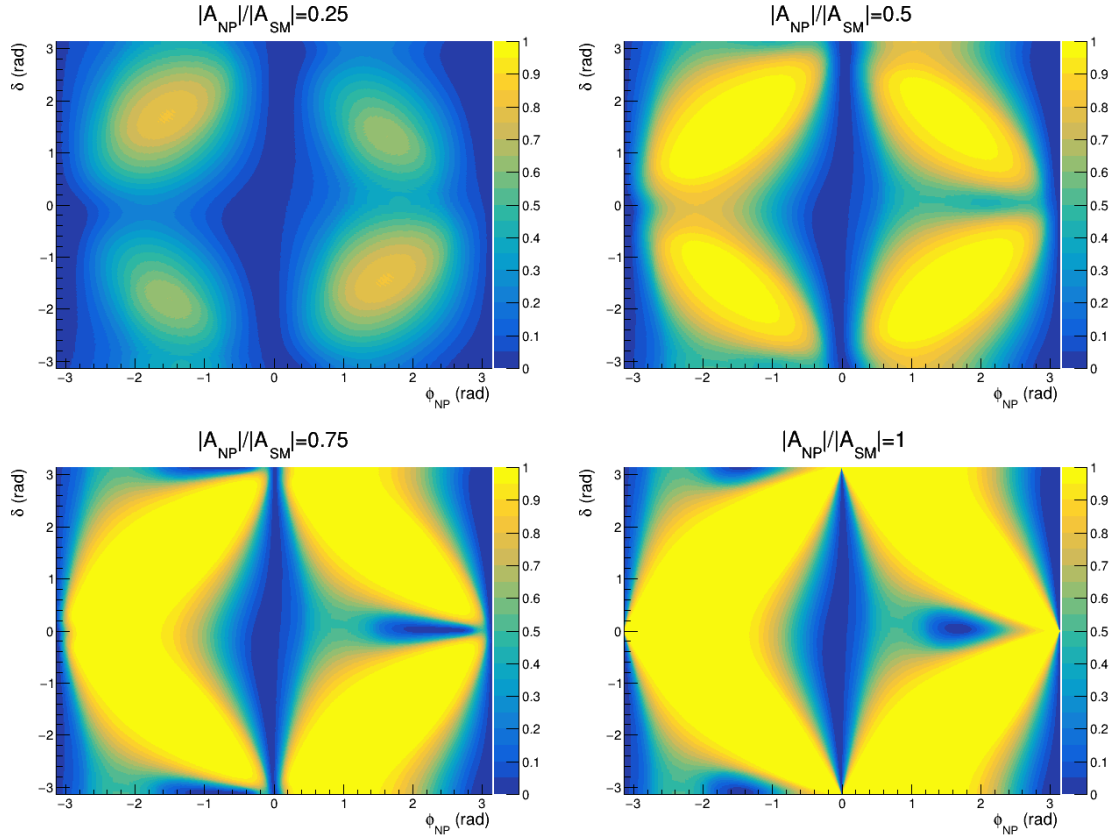


Figure 7.3: 2-D confidence level to reject the weak phase of NP (ϕ_{NP}) and the difference of strong phase between NP and SM (δ) with various values of the amplitude ratio of NP to SM ($r \equiv |\mathcal{A}_{NP}|/|\mathcal{A}_{SM}|$). The color shows the confidence level.

7.2 Impact of calibration for resolution parameters

We devised a novel method to calibrate the resolution parameters using cosmic samples, which is useful for further analysis in Belle II. This method has the following features:

- It can calibrate most of the resolution parameters,
- It can be used for any form of resolution function, and
- The systematic uncertainty due to this calibration is influenced by the statistics of cosmic samples, which can be easily increased.

However, since we focused on developing the method in an optimized shape for our target channel at the current statistics, we simplified the calibration procedure and left room for improvement. According to App. C, we confirmed there is a bias of track parameter ω that depends on the electric charge of a track. Although we suspect the problem is due to an inaccurate B-field mapping in the Belle II software framework, we did not fully understand the reason for the bias and consider the effect of it as a systematic uncertainty. Also, we simplified the modeling for the pull distribution of track parameters in the cosmic sample, such as assuming no bias in the track parameter determination and ignoring possible charge asymmetry in the reconstruction of tracks. Also, the possible improvement of Fast Simulation discussed in Ch. 5 might be needed at the higher statistics.

Nevertheless, this procedure is considered promising in terms of versatility which can be easily applied to other channels or different shapes of resolution functions and generality that can correct most of the resolution parameters. If this calibration method is improved as discussed above and validated by applying this method to other large control samples, it is expected that it will be helpful to the *CP* violation study of Belle II in other channels in the future. Also, we expect that the measurement of *CP* asymmetries in the other subchannel of $B^0 \to \eta' K^0_S$ can be corrected with little effort using the versatility of this method. According to the App. I.1.3, the sensitivity of this measurement is better than the results from Belle, and we expect such improvement in the sensitivity also in other subchannels. Thus, the impact of this method would be large and strong when it is applied to other channels for *CP* asymmetry measurement, especially to the other subchannels of $B^0 \to \eta' K^0_S$.

7.3 Comparison with Belle results

According to literature[16, 18], the latest result for *CP* asymmetry measurement of the sub-channel of $B^0 \to \eta'(3\pi)K^0_S$ with the full dataset of the Belle experiment is as follows:

$$A^{\text{Belle}}_{CP(\eta'(3\pi)K^0_S)} = -0.058 \pm 0.181, \text{ and} \tag{7.3}$$

$$S^{\text{Belle}}_{CP(\eta'(3\pi)K^0_S)} = +0.800 \pm 0.259. \tag{7.4}$$

We compare components for extracting *CP* asymmetry between this analysis and the Belle analysis given in the References and discuss the details in the following subsections.

7.3.1 The yield of (Signal + SxF) category

We compare the yield of signal events for the *CP* fitting. We calculate the signal yield per the unit number for $B\bar{B}$ events and confirm that the value in our analysis is lower than that in the Belle results. In the Belle analysis, the signal yield with 772×10^6 $B\bar{B}$ pairs is 174.3 ± 13.5[18], and the signal yield per 1×10^6 $B\overline{B}$ pairs is 0.225 ± 0.018. Meanwhile, the signal yield in the signal region from the experimental data of our analysis is 55.1 ± 8.1 with $(387 \pm 6) \times 10^6$ $B\overline{B}$ pairs and 0.142 ± 0.021 signal yields per 1×10^6 $B\bar{B}$ pairs, which are lower than that from the Belle analysis.

We investigate the reason for the low signal yield, and the main reason for such a low yield is low signal efficiency. We require a condition of the number of hits from CDC to the charged tracks, which was not required in the analysis procedure of Belle. The selection criteria can remove the low-momentum pions, and it causes decreasing in signal efficiency. To check the effect of the criterion quantitatively, we reconstructed our target decay using the signal MC samples without the CDC selection and calculated the signal efficiency. The signal efficiency of the main channel without the selection is 13.97%, and the nominal signal efficiency in this analysis is 10.88% as shown in Table 4.3. Therefore, the main reason for the low signal efficiency is the selection criteria related to CDC, and we think this can be improved in future analysis with additional studies.

Also, the overall fraction of the (Signal + SxF) category in the fitting region (F_{sig}) is 55%, which is lower than the Belle result. Such a low fraction is due to the loose selection on η and B^0. In the analysis of Belle, selection criteria for η and η' is $0.535 < M_\eta < 0.558\,\mathrm{GeV}/c^2$ and $0.945 < M'_\eta < 0.970\,\mathrm{GeV}/c^2$, respectively, and the signal region is $-0.08 < \Delta E < 0.06\,\mathrm{GeV} \ \wedge \ 5.27\,\mathrm{GeV}/c^2 < M_{bc}$. The resultant signal fraction in the signal region is 65%[18]. In our analysis, we select η and η' candidates with selections of $0.52 < M_\eta < 0.57\,\mathrm{GeV}/c^2$ and $0.40 < \Delta M_{\eta'-\eta} < 0.42\,\mathrm{GeV}/c^2$, respectively, and the signal region is $|\Delta E| < 0.1\,\mathrm{GeV} \wedge 5.27 < M_{bc} < 5.29\,\mathrm{GeV}/c^2$. Although the selection criteria for η' selection is hard to compare, criteria for η and ΔE for the signal region in this analysis are looser than the selection criteria in Belle. We applied combinations of three criteria in the Belle analysis for η, η', and ΔE to our sample to check the signal efficiency. The result is shown in Table 7.2. Appendix L explains the effects of the number of Background category events, which indicates that the distribution of uncertainty for A_{CP}^{fit} and S_{CP}^{fit} are not affected by the number of events in the Background category. Therefore, the number of events in the Background category does not matter on the sensitivity of *CP* asymmetries in our analysis.

Category	Original	Selection criteria in Belle					
		η	η'	ΔE	η & η'	η & ΔE	η & η' & ΔE
Signal	153	142	141	153	137	142	137
π-exchanged SxF	46	36	44	46	36	36	36
Combinatorial SxF	36	27	31	32	26	23	23
Background	157	110	123	116	105	81	77
Purity ($S/(S+B)$)	59.9%	65.1%	63.7%	66.6%	65.5%	71.3%	71.8%

Table 7.2: The number of each category and the purity of the (Signal + SxF) category with various combinations of selection criteria in Belle for η, η', and ΔE.

7.3.2 Flavor Tagger and Δt resolution

We compare the performance of Flavor Tagger and vertex fitting between Belle and Belle II. The effective efficiency for the flavor tagging in the experimental data of Belle II is $(31.68 \pm 0.38)\%$, as given in Table 4.1, and according to literature[59], the efficiency of Belle is $(30.1 \pm 0.4)\%$. For the Δt resolution function, we compared the root-mean-square value from our analysis and one from Belle. In this analysis, the root-mean-square(RMS) of a residual distribution for Δt from the vertex fitting for both sides is (1.074 ± 0.0019) ps, and the mean is (-0.2318 ± 0.0026) ps. In the Belle analysis, the RMS value for the residual distribution is (1.201 ± 0.022) ps, and the mean is (-0.2057 ± 0.030) ps[18]. The resolution (RMS of the residual distributions) is slightly improved from that in the Belle analysis, although the boost factor ($\beta\gamma c$) is smaller than Belle's one by a factor of 1.5. This is thanks to the PXD detectors newly installed in the Belle II experiment and requiring at least two charged pions originating from η' should have PXD hits. Therefore, the performance of Flavor Tagger from Belle II is similar to Belle, but we have better vertex resolution than Belle.

7.3.3 Summary

We compared various factors that can affect the sensitivity of *CP* asymmetry measurements of Belle II (this analysis) and Belle ([16, 18]). For the signal efficiencies, Belle II shows a lower value than Belle, and it can affect the sensitivity of *CP* asymmetries. The lower efficiency is due to some selection criteria being too strict for studies of *CP* violation and can be improved by removing or loosening such criteria. The purity of the (Signal + SxF) category of Belle II is also lower than Belle, but we confirmed that it does not affect the sensitivity of our measurement. Belle II has a better resolution for the Δt determination than Belle and a similar effective efficiency of Flavor Tagger to Belle, and those can positively affect the sensitivity of Belle II. In the next section, we discuss how to improve the sensitivity of this analysis in Belle II and suggest the prospect of *CP* asymmetry measurement in the channel of $B^0 \to \eta' K^0_S$.

7.4 Prospect

Statistical uncertainty is the primary one in this analysis, which can be reduced by accumulating more statistics. We discuss the possible improvements and prospects of *CP* violation measurement in the $B^0 \to \eta' K^0_S$ channel at the integrated luminosity of $50\,\mathrm{ab}^{-1}$, a target luminosity of the Belle II experiment.

Recovery of low-momentum pions As discussed in the previous subsection, removing the CDC hit number criteria for charged pions will improve the reconstruction efficiency for signal events 1.3 times. From this result, we expect removing the selection criteria of the number of CDC hits for charged pions will reduce the uncertainty of A_{CP} and S_{CP} by a factor of $1/\sqrt{1.3}$. However, this may deteriorate the resolution of B_{CP} vertex fitting, so we calculated the RMS value of the residual distribution for the determination of Δt without the CDC criteria. The value is $(1.076 \pm 0.0017)\,\mathrm{ps}$, which confirms no significant degradation of the resolution for Δt by removing the CDC criteria. Therefore, we expect to fully benefit from increased sensitivity on *CP* asymmetries thanks to increasing signal efficiency while minimizing the negative effects by removing the CDC criteria.

Improvements of systematic uncertainty We expect some systematic uncertainty can be improved in the future since the systematic uncertainties contain statistical fluctuation. We scaled the below uncertainties by $1/\sqrt{\mathcal{L}}$ where $\mathcal{L}$ is the integrated luminosity.

- **Uncertainty of $w, \Delta w$:** The wrong tag fraction of Flavor Tagger is calibrated by analyzing Data in the channel of $B^0 \to D^{(*)-}\pi^+$, and we can determine the parameters precisely with larger statistics, which makes this uncertainty smaller.
- **Parameters of $P_{bckg}(\Delta t)$:** As discussed in Sec. 5.3, the parameters of $P_{bckg}(\Delta t)$ is determined with sideband events from experimental data. If we have more statistics, more precise parameters are given, which reduces this uncertainty.
- **Resolution calibration:** The calibration procedure for the resolution parameters uses the cosmic Data samples. Since the size of the cosmic sample can be grown along with the data-taking period, we consider this uncertainty will be reduced.
- **Selection criteria for the vertex quality:** We evaluate this uncertainty by moving thresholds of selection criteria for vertex quality variables, and such a method may include statistical fluctuation by removing and adding some events. Currently, we expect uncertainties by the statistical fluctuation to account for most of this uncertainty, and it will be reduced by increasing the luminosity.
- **Non-primary components of R_{tag}:** According to the Sec. 6.3.5, the estimated s^{global}_{np}, a scale factor to evaluate this uncertainty, has large uncertainty and is consistent with 1, which suggests this uncertainty also contains large statistical fluctuation. Thus, we consider this uncertainty will be reduced by increasing the luminosity.

We did not scale the remaining systematic items, such as "Fit bias", by the integrated luminosity for the uncertainty extrapolation since we consider those uncertainties are not directly connected with the statistics of the Belle II experiment. However, it might be possible to reduce these uncertainties further in the future. For example, the systematic uncertainty of "Physical quantities of τ_B and Δm_d" is expected to be reduced as the precision of the values will be improved in the future. On top of that, "Tag-side interference" can be removed if we adopt a PDF that can consider the interference in the *CP* fitting, which helps reduce the systematic uncertainty further. We summarized systematic errors expected at the Belle II target luminosity in Table 7.3, and a blue line in Fig. 7.4 shows the projected systematic uncertainty.

Systematic items	Current ($362\,\mathrm{fb}^{-1}$)		Target $\mathcal{L}$ ($50\,\mathrm{ab}^{-1}$)	
	δA	δS	δA	δS
Uncertainty of $w, \Delta w$	± 0.0090	± 0.0088	± 0.0008	± 0.0007
Parameters of $P_{bckg}(\Delta t)$	± 0.0083	± 0.0347	± 0.0007	± 0.0030
Resolution calibration	± 0.0103	± 0.0203	± 0.0009	± 0.0017
Selection criteria for the vertex quality	± 0.0111	± 0.0273	± 0.0009	± 0.0023
Non-primary components of R_{tag}	± 0.0215	± 0.0154	± 0.0018	± 0.0013
Fit bias	$+0.0076$	$+0.0029$	$+0.0076$	$+0.0029$
μ of Flavor Tagger	-0.0052	-0.0019	-0.0052	-0.0019
Parameters of Signal PDF	± 0.0029	± 0.0052	± 0.0029	± 0.0052
Physical quantities of τ_B and Δm_d	± 0.0007	± 0.0050	± 0.0007	± 0.0050
Uncertainty scaling of the track parameters	$+0.0002$	$+0.0003$	$+0.0002$	$+0.0003$
Bias due to the resolution function	-0.0103	$+0.0028$	-0.0103	$+0.0028$
Misalignment	± 0.0054	± 0.0083	± 0.0054	± 0.0083
IP profile	± 0.0020	± 0.0024	± 0.0020	± 0.0024
Momentum scaling	± 0.0013	± 0.0005	± 0.0013	± 0.0005
Tag-side interference	-0.0045	$+0.0038$	-0.0045	$+0.0038$
Total	$^{+0.0319}_{-0.0310}$	$^{+0.0531}_{-0.0530}$	$^{+0.0113}_{-0.0135}$	$^{+0.0132}_{-0.0133}$

Table 7.3: Improvements of systematic uncertainties at the current statistics ($362\,\mathrm{fb}^{-1}$) and target luminosity of Belle II ($50\,\mathrm{ab}^{-1}$). The upper five items are reduced by increasing the statistics, and those in the lower part are not affected by the statistics.

Statistical uncertainty with other subchannels We estimate the statistical uncertainty with more luminosity when we include other prominent subchannels of the $B^0 \to \eta' K^0_S$ channel. The resolution model and its calibration method developed for this analysis can be used to analyze the other subchannels of the $B^0 \to \eta' K^0_S$ decay. Especially, $\eta' \to \eta(\gamma\gamma)\pi^+\pi^-$ and $\eta' \to \rho(\pi^+\pi^-)\gamma$ channels are highly prominent since the signal yield of these two subchannels is 2-3 times higher than our target channel, as shown in Table 7.4. Therefore, we can get extensive statistics and increase sensitivity

with relatively small efforts by reusing the resolution model to the two subchannels.

Channel	$\mathcal{B}\,(\times 10^{-6})$[11]	Signal yield[16, 18]	A_{CP} [16, 18]	S_{CP} [16, 18]
$\eta' \to \eta(\pi^+\pi^-\pi^0)\pi^+\pi^-$	2.146 ± 0.168	174.3 ± 13.5	-0.058 ± 0.181	0.800 ± 0.259
$\eta' \to \eta(\gamma\gamma)\pi^+\pi^-$	3.713 ± 0.289	648.3 ± 27.9	$+0.161 \pm 0.098$	0.724 ± 0.151
$\eta' \to \rho(\pi^+\pi^-)\gamma$	6.548 ± 0.510	1410.5 ± 48.5	-0.071 ± 0.069	0.718 ± 0.098

Table 7.4: Branching ratio and signal efficiency for prominent other two subchannels.

To estimate the statistical uncertainty when we include the two subchannels above in the analysis, we conducted a study using Toy MC datasets. *CP* asymmetries from the latest result by Belle of $B^0 \to \eta' K^0_S(\pi^+\pi^-)$ subchannels are given as $A_{CP} = -0.002 \pm 0.054$ and $S_{CP} = +0.728 \pm 0.079$[18], and we multiplied the uncertainty of the measurement by improvements of the sensitivity given in App. I.1.3. We calculate a projection of statistical uncertainties with the two subchannels by scaling the multiplied uncertainty with $1/\sqrt{\mathcal{L}}$, indicated as a red line in Fig. 7.4. For example, the statistical uncertainty at $50\,\mathrm{ab}^{-1}$ is $\sigma_{A_{CP}} = 0.00628$ and $\sigma_{S_{CP}} = 0.00890$.

Summary and tasks for further improvements Based on the discussion for improvements in this subsection, we extrapolated the total uncertainty of A_{CP} and S_{CP} to the integrated luminosity of $50\,\mathrm{ab}^{-1}$, which is given as a green line in Fig. 7.4. The uncertainty at $50\,\mathrm{ab}^{-1}$ integrated luminosity is expected to be $\delta A = 0.01653$ and $\delta S = 0.01600$, which is close to the current SM theoretical error given in Tab. 7.1.

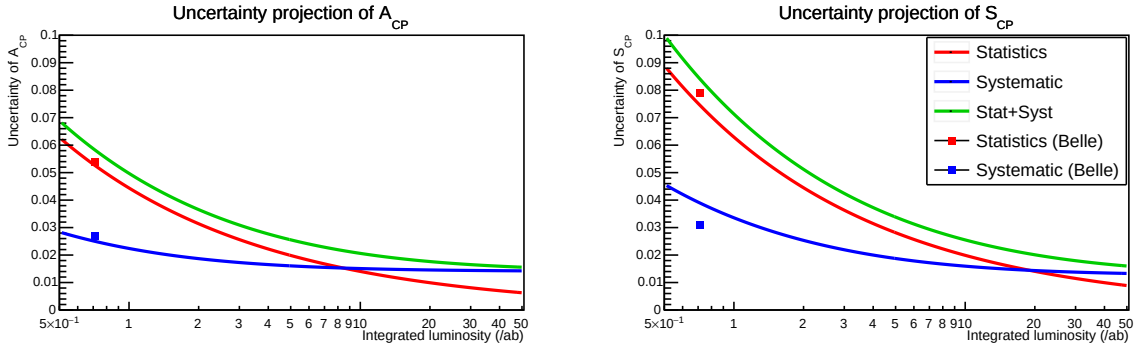


Figure 7.4: Projection of uncertainty for A_{CP} and S_{CP}. Rectangular points indicate the corresponding error of Belle[16, 18].

In the high luminosity region of $\mathcal{L} > 15\,\mathrm{ab}^{-1}$, the primary source of uncertainty becomes systematic errors, and the systematic errors not depending on the luminosity would be a main target to be reduced to achieve the higher sensitivity. For example, the "Misalignment" item in Tab. 7.3 will be the most major item in the systematic uncertainties of S_{CP}. This factor is introduced to consider the imperfection in the alignment measurement of subdetectors, and the progress of alignment methods will help to reduce the

systematic uncertainty. Also, we should consider the correlation in the evaluation for the systematic uncertainties in the high luminosity. For instance, we ignored the correlation between the parameters of the PDF when we evaluate the "Parameters of Signal PDF", which will be the second main source of the systematic uncertainty for S_{CP} in $50\,\mathrm{ab}^{-1}$, to make a conservative approach, but it should be revised to consider such correlations in the high luminosity. Another example is the correlation between items for the systematic uncertainties. Currently, we consider there is no significant correlation between each item, but "Bias due to the resolution function" is correlated with Flavor Tagger since both of them use the tag-side information. In the high luminosity, we can divide the resolution model by high and low dilution factors, and it might help to mitigate the systematic uncertainties by the resolution function bias, which is the most-contributing item to δA in $50\,\mathrm{ab}^{-1}$. Therefore, it will be important to reduce the systematic error after $\mathcal{L} \sim 15\,\mathrm{ab}^{-1}$ with advanced methods to achieve higher sensitivity.

7.4.1 Sensitivity and constraints on New Physics

We calculated the 2-D confidence interval with the sensitivity at $50\,\mathrm{ab}^{-1}$, as shown in Fig 7.5 with the measured CP asymmetries are consistent to the current world-average value from the charmonium mode ($A_{CP} = 0$ and $S_{CP} = 0.698$)[19]. Also, we recalculate the constraint on parameters of New Physics (δ, ϕ_{NP}, and $r \equiv |\mathcal{A}_{NP}|/|\mathcal{A}_{SM}|$) based on the new confidence interval, which is shown in Figs. 7.6 and 7.7. If New Physics has a sizable weak phase (ϕ_{NP}), we can exclude the ratio of amplitude r above 3-5%.

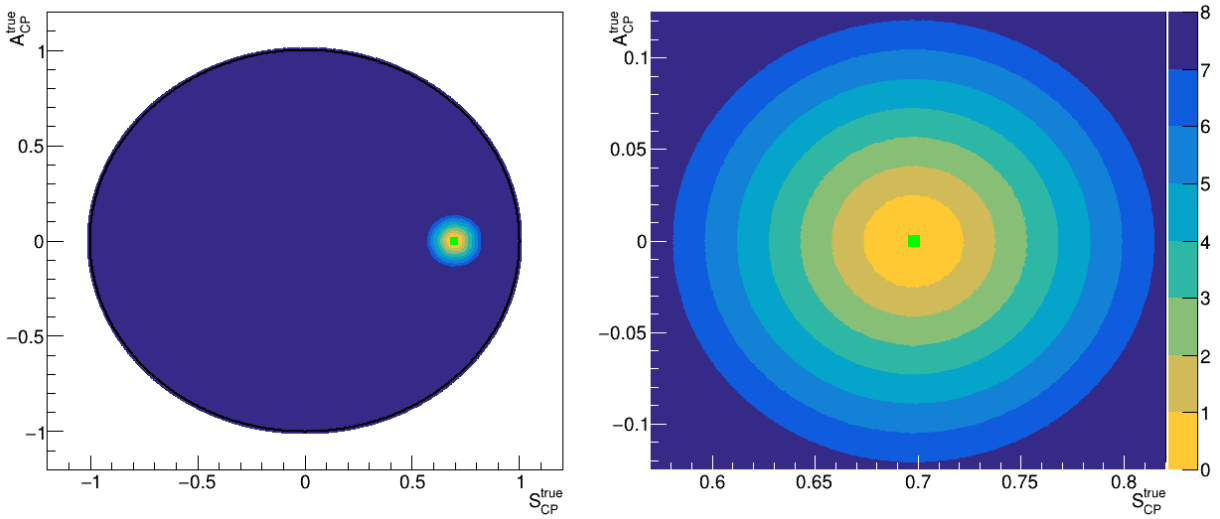


Figure 7.5: The confidence interval when the measurement results for A_{CP} and S_{CP} is the current world-average CP asymmetry in the decay of $B^0 \to c\bar{c}K^0$ (the green point) at the target luminosity of Belle II. The lightest region is $< 1\sigma$, and the darkest is $> 7\sigma$. The left plot shows the entire physical region, and the right is an enlarged version.

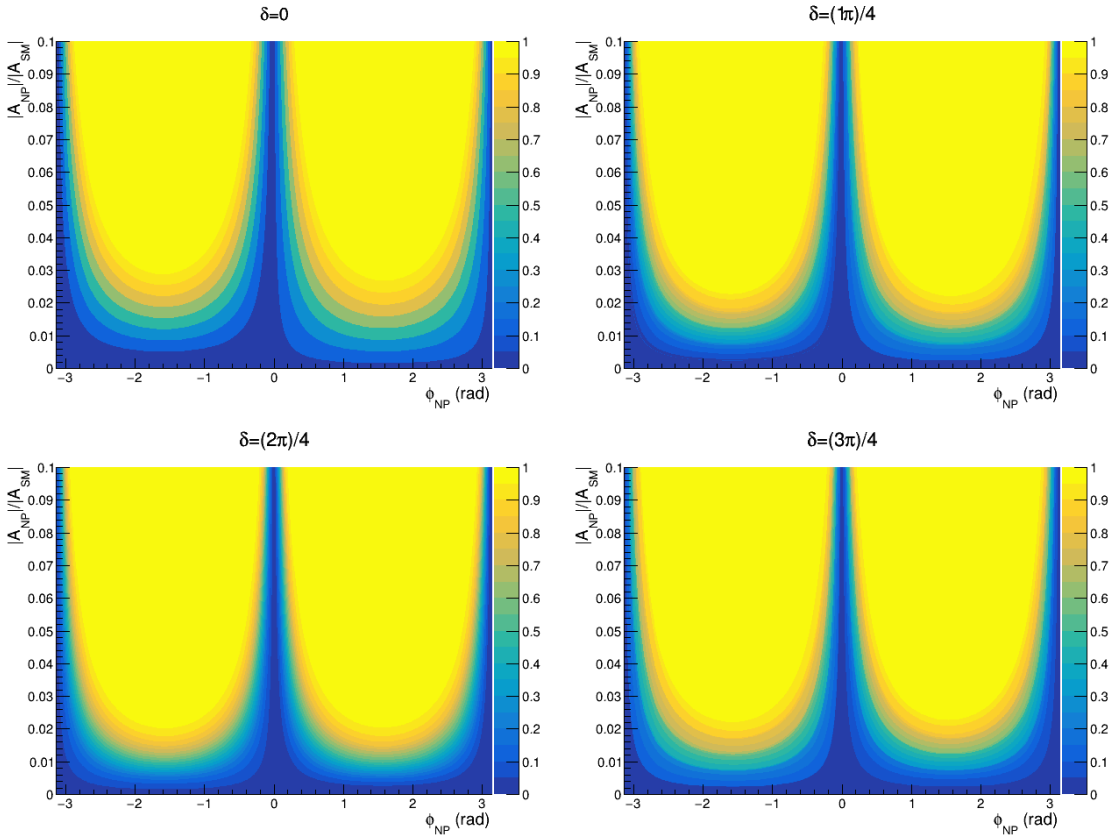


Figure 7.6: Map of confidence level to reject parameters of New Physics at the integrated luminosity of $50\,\mathrm{ab}^{-1}$. ϕ_{NP} vs. $|\mathcal{A}_{NP}/|\mathcal{A}_{SM}|$ is shown with slicing of δ. The color indicates the confidence level.

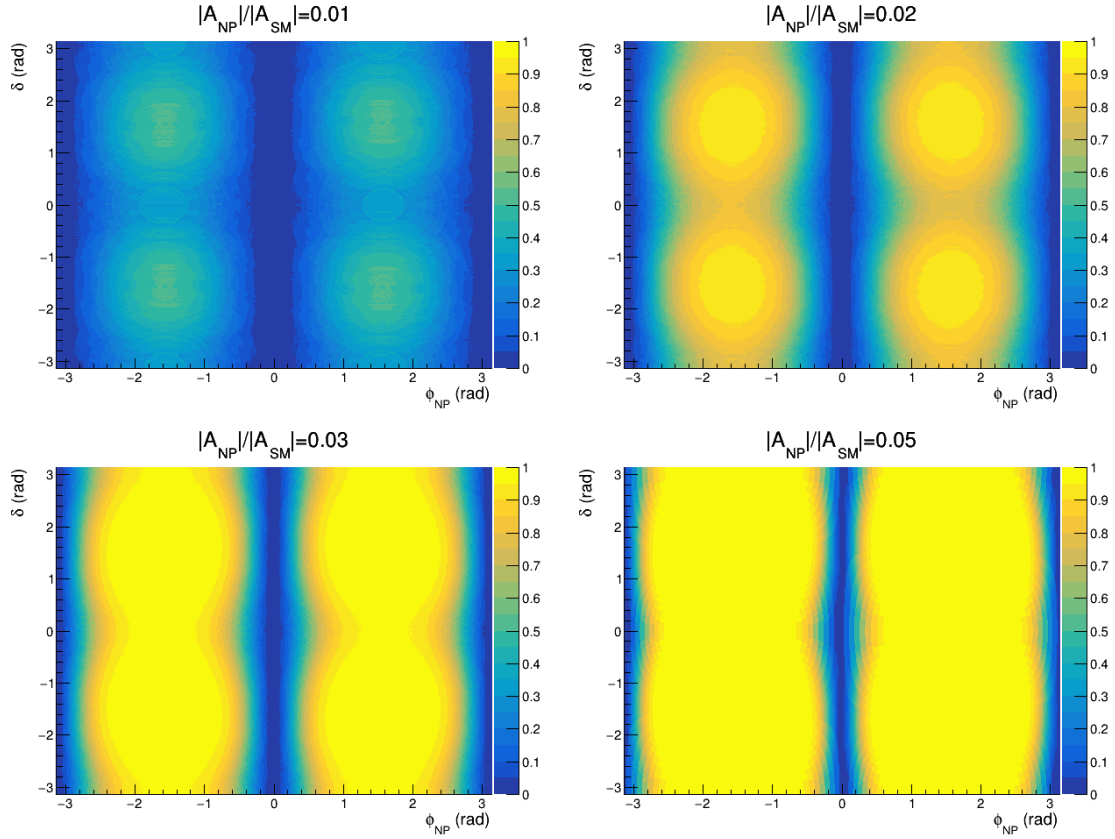


Figure 7.7: Map of confidence level to reject parameters of New Physics at the integrated luminosity of $50\,\mathrm{ab}^{-1}$. ϕ_{NP} vs. δ is shown with slicing of $r \equiv |\mathcal{A}_{NP}/|\mathcal{A}_{SM}|$. The color indicates the confidence level.

7.5 Conclusion

We developed methodologies for the whole analysis procedure to measure *CP* asymmetries in the decay of $B^0 \to \eta' K^0_S$ with the subchannel of $\eta' \to \eta(\pi^+\pi^-\pi^0)\pi^+\pi^-$, which is mediated by $b \to sq\bar{q}$ penguin diagram within the Standard Model and provides a promise mean to explore New Physics with the *CP*-violation by comparing *CP* asymmetries in the channel of $b \to c\bar{c}s$ transition, such as the $B^0 \to J/\psi K^0_S$ decay. In this analysis, we measured *CP* asymmetries in our target channel with the full-dataset of Belle II on the $\Upsilon(4S)$ resonance collected from April 2019 to June 2022, which corresponds to $N(B\bar{B}) = (387 \pm 6) \times 10^6$ and $(362 \pm 2)\,\mathrm{fb}^{-1}$.

For the measurement, we have devised a reliable methodology to determine the *CP* asymmetries with high statistics of $50\,\mathrm{ab}^{-1}$, which is the target luminosity of the Belle II project. We developed a method to obtain an optimal reconstruction result for the *CP* fitting, such as optimizing the π^0 and K^0_S selection criteria described in App. E and η' selection using the mass difference between η' and η. On top of that, we established improved techniques to consider various factors of Δt in the *CP* fitting. The calibration of the resolution parameters using cosmic samples is the first attempt to confirm the reason for the structure of the resolution function and correct the difference in the resolution between MC and data. The method has more advantages than conventional ones for calibration and can be applied to other channels or different forms of resolution functions. Then, we built a procedure for the validation of the resolution parameters using the control channel of $B^\pm \to \eta' K^\pm$ by emulating the vertex resolution of $B^0 \to \eta' K^0_S$. Also, we developed a method to extract the probability of signal events with 5-D variables, which is very sensitive for signal and background events, so that the sensitivity of our measurement is unaffected even with more background events, as discussed in App. L. Using these methodologies, we extracted *CP* asymmetries, and the results are

$$A_{CP} = -0.111^{+0.319}_{-0.311}(\mathrm{stat})^{+0.032}_{-0.031}(\mathrm{syst}), \text{ and} \tag{7.5}$$

$$S_{CP} = +0.248^{+0.470}_{-0.526}(\mathrm{stat})^{+0.053}_{-0.053}(\mathrm{syst}). \tag{7.6}$$

The result of our measurement is consistent within 1σ with the prediction from the Standard Model based on the measurements in the charmonium decay and the averaged asymmetries in $B^0 \to \eta' K^0_S$ from the previous analysis conducted by BaBar and Belle. Also, in the target luminosity of Belle II, the statistical error becomes a minor source of uncertainty, and we should focus on reducing the systematic errors mainly. At the target integrated luminosity, the total uncertainty will be 0.01653 for A_{CP} and 0.01600 for S_{CP}, similar to the current SM theoretical errors, and we can reject the amplitude ratio r above 3-5% when New Physics has a sizable weak phase.

We hope and believe all the development, improvement, and discussion from this study will be highly useful for the future measurement of *CP* asymmetries in the $B^0 \to \eta' K^0_S$ channel as well as in other channels at Belle II.

Appendix A

Algorithm for Continuous Variable Generation

In our *CP* fitting procedure, we use conditional variables such as $\chi^2_{CP,tag}$ or E^*_{beam}. To confirm the procedure, we perform an ensemble test and need to generate the conditional variables in the sample generation step. It might be good to resample the variables from the generic MC or signal samples. However, if the pool size for resampling is limited, then repeated use of one value for the conditional variables will happen. This can negatively affect the ensemble test, so we devised an algorithm to generate conditional variables without the resampling from the populations with limited statistics.

A.1 1-D case

We generated the variable with the following procedure:

1. Get the list of a variable from a given population.
2. Sort the list and pick two consecutive elements from the list (a_i, a_{i+1}).
3. Generate a uniform random number r within the range of $(0, 1)$.
4. Calculate the formula of $(1 - r) \cdot a_i + r \cdot a_{i+1}$. This value will be generated value from this procedure.

A.2 2-D case

To reproduce some correlation between two variables, a 2-D variable generator is needed. We generated a pair of variables (x, y) as following procedure:

1. Build a 1-D variable generator for variable x.

2. Make $N = 100$ uniform bins for x within the range of $(x_{\min}, x_{\max})$ and divide (x, y) pairs to the corresponding bin for x value.

3. Build 1-D variable generators for y per each bin in Step 2.

4. Generate x value from the 1-D generator built in Step 1.

5. Generate y value using the 1-D generator from the corresponding bin to the x value generated from the previous step.

We tested the 2-D generator by generating $(M_{bc}, \Delta E)$ pair from signal samples. Figure A.1 indicates the generator can reproduce the original population.

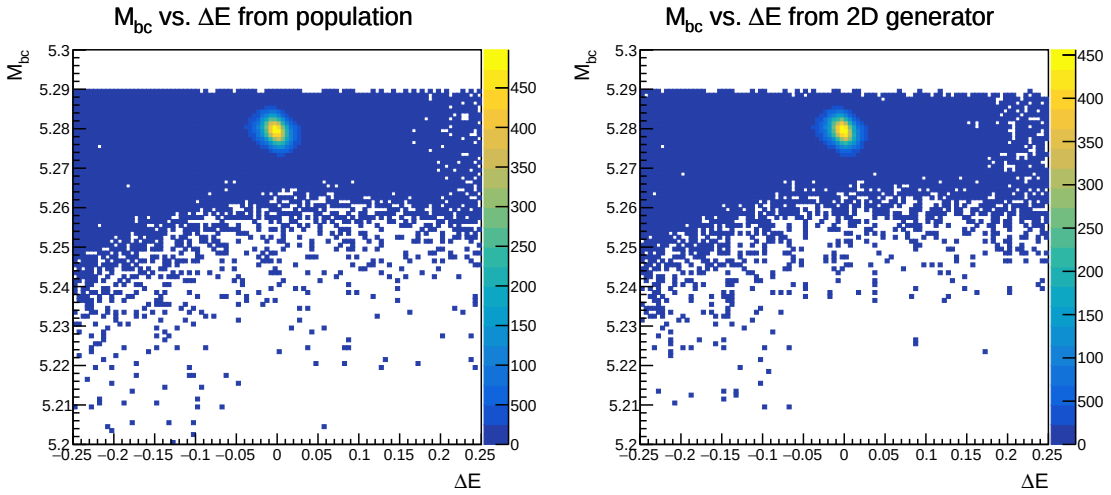


Figure A.1: Comparison of 2D scatter plots of M_{bc} vs. ΔE between signal sample(left) and 2-D case generator(right) using the signal sample.

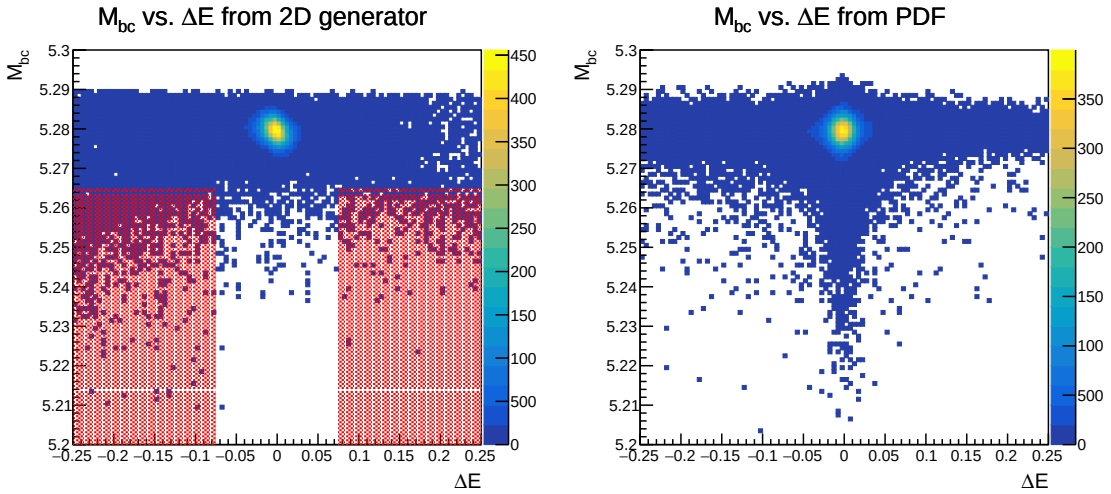


Figure A.2: Comparison of 2D scatter plots of M_{bc} vs. ΔE from Toy MC samples between using 2-D case generator (left) and signal extraction PDF (right). The probability for events in the red-hatched area in the left-side plot will be underestimated by the PDF.

Appendix B

Supplemental Studies for the Reconstruction & Signal Extraction

B.1 Reconstruction efficiency for the control channel

Selection	Signal + SxF events	Background events
η', η Selection, $M_{bc}, \Delta E$ and the best candidate selection	15.76% (15.76 %)	– (100 %)
$p_{\pi^0} > 0.204$ GeV/c	92.33% (14.55 %)	84.63% (84.63 %)
Selection for PXD Hits	99.8% (14.52 %)	99.52% (84.22%)
$\left(\frac{\chi^2}{ndf}\right)^{CP} < 10$	90.57% (13.15 %)	75% (63.17 %)
$\sigma_\ell^{CP} < 50\mu$m	94.68% (12.45 %)	88.71% (56.04 %)
Tag-side vertex fitting succeeded	98.71% (12.29 %)	98.2% (55.03 %)
$\sigma_\ell^{tag} < 100\mu$m	97.81% (12.02 %)	98% (53.93 %)
Flavor tagging succeeded	100% (12.02 %)	100% (53.93 %)

Table B.1: The efficiency for the reconstruction of $B^\pm \to \eta' K^\pm$. The percentage in parentheses is a cumulative efficiency, and the value for the Background category is the relative one for efficiency after the best candidate selection for $B^\pm_{CP}$.

B.2 Ensemble test using Toy MC dataset

We performed the ensemble test using the Toy MC generator to check the validity of the fitting procedure. We generated 1000 toy samples with parameters from signal extraction of experimental data for the main channel only. There are no failed trials, and the pull distribution for all parameters shows the standard normal distribution (Figure B.1).

B.3 Correlation between variables

We investigated the correlation between variables for the signal extraction 3D fitting (M_{bc}, ΔE, and $\mathcal{O}_{CS}$). The 2D scattering plots between M_{bc} vs. ΔE, M_{bc} vs. $\mathcal{O}_{CS}$, and ΔE vs. $\mathcal{O}_{CS}$, and sliced distribution from signal events in the main channel are shown in Figure B.4, B.5, and B.6 (Figure B.7, B.8, and B.9 for the control channel.

We can confirm that variables of M_{bc} and ΔE are correlated. Still, we are ignoring such correlation in our signal extraction, so we estimated the amount of bias by performing an ensemble test that can reproduce the correlation in the sample generation step. An algorithm for generating a pair of $(M_{bc}, \Delta E)$ variables is described in Appendix A. We generated 1000 Toy MC samples using pre-fit parameters for the main channel from the experimental data. We expect F_{sig} will be underestimated because our PDF for the signal extraction predicts a low probability for the correlated events. (The red-hatched area in Figure A.2.) Figure B.2 shows pull distributions from the ensemble test and Figure B.3 shows distributions of the absolute bias. We can confirm that the pull distribution for F_{sig} is negatively shifted and other parameters are slightly biased compared to the pull distribution of B.1. The average value of the absolute bias in F_{sig} due to the correlation is $\mathcal{O}(0.1\%)$. We will consider the effects of this correlation on the *CP* parameter estimation in systematic error studies.

Although we can see a weak correlation between $\mathcal{O}_{CS}$ and M_{bc} or $\mathcal{O}_{CS}$ and ΔE, we can neglect the correlation because the correlation is negligible. The source of this correlation is the Combinatorial SxF events. In the Background category, there is no correlation between variables as shown in Figure B.10 (main channel) and B.11 (control channel).

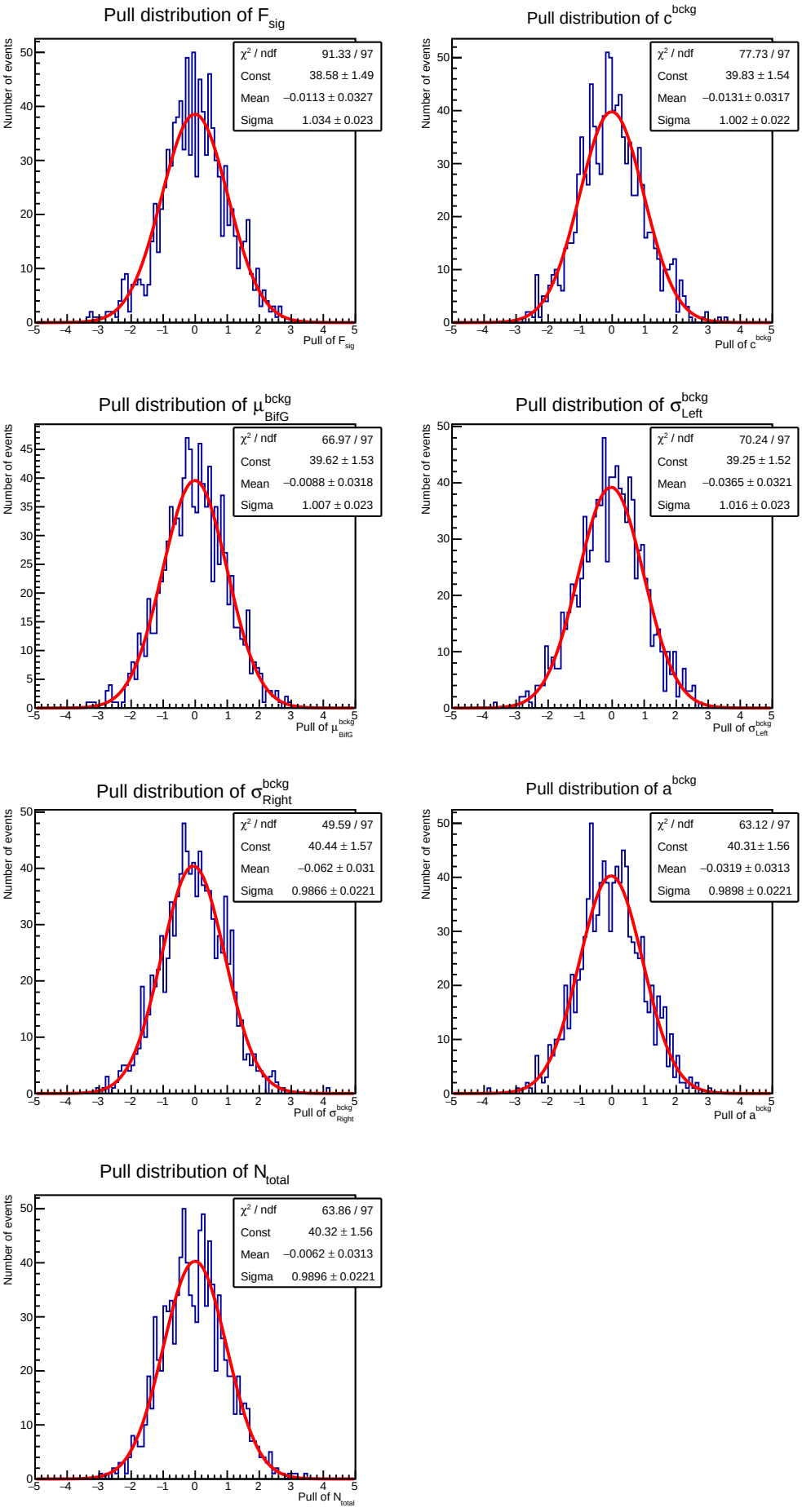


Figure B.1: Pull distribution of all parameters from the ensemble test for the signal extraction.

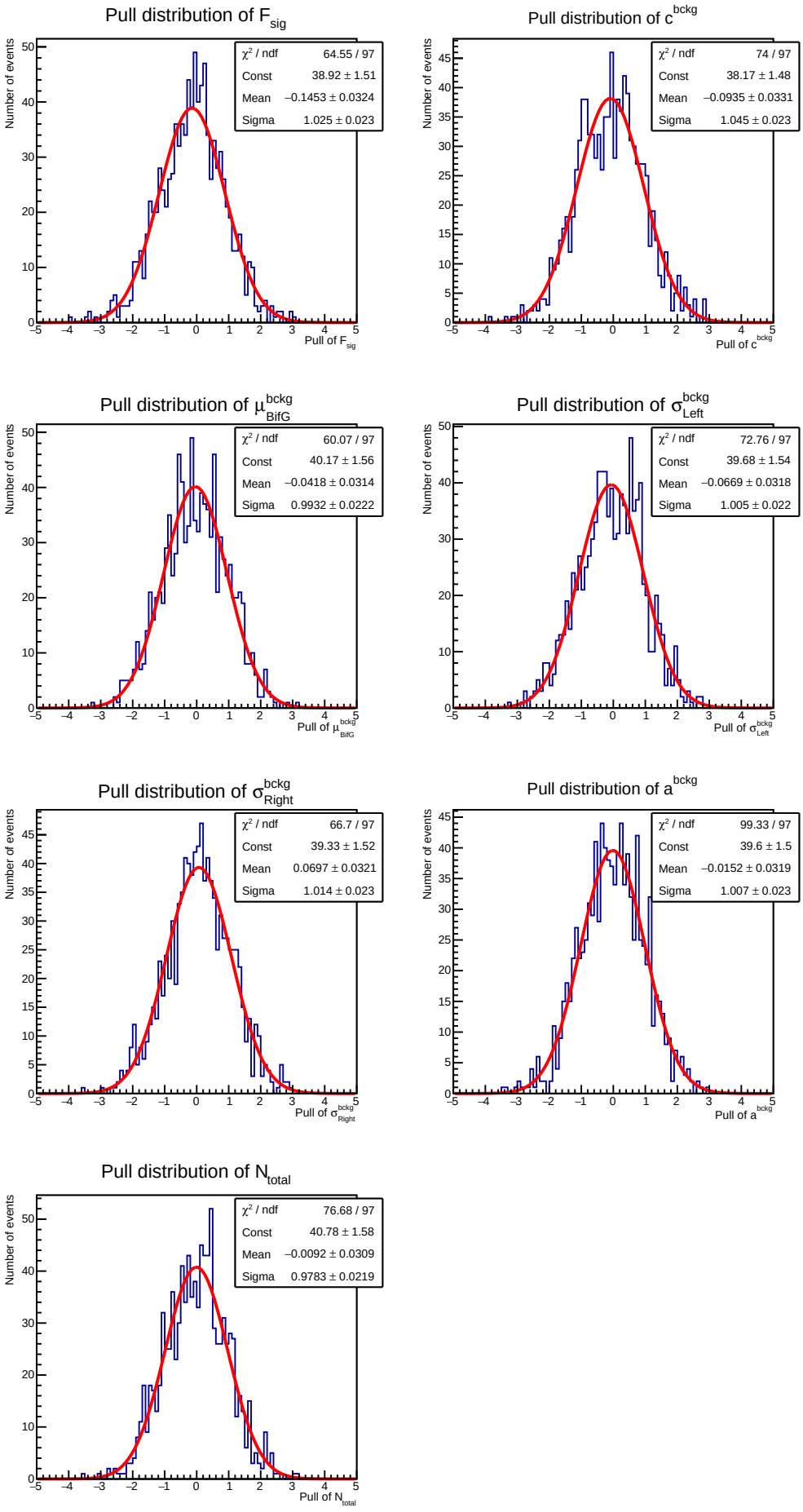


Figure B.2: Pull distributions of all parameters for the signal extraction from the ensemble test with the M_{bc}-ΔE correlation.

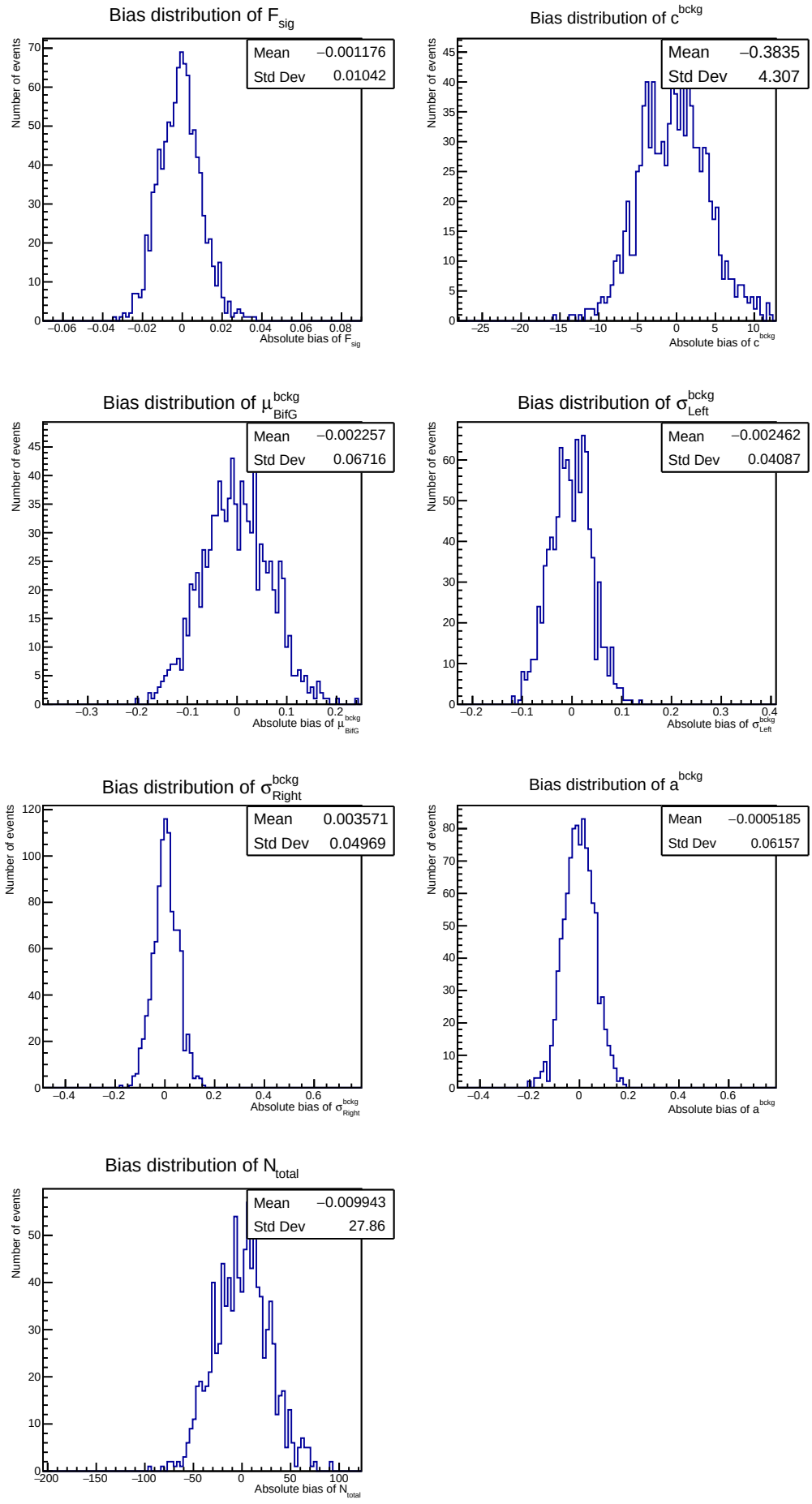


Figure B.3: Absolute bias distributions of all parameters for the signal extraction from the ensemble test with the M_{bc}-ΔE correlation.

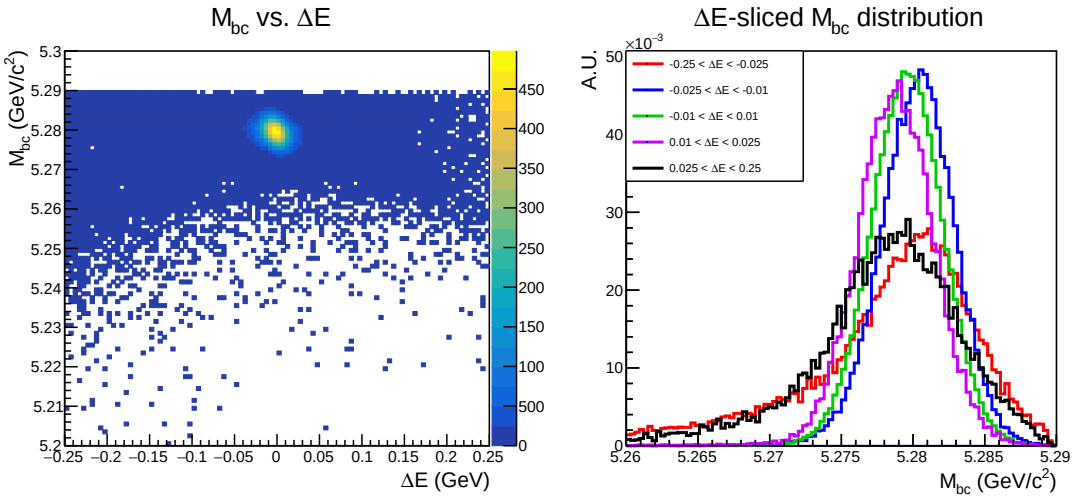


Figure B.4: 2D scatter plot of M_{bc} vs. ΔE (left) and ΔE-sliced M_{bc} distribution (right) from signal events in the main channel.

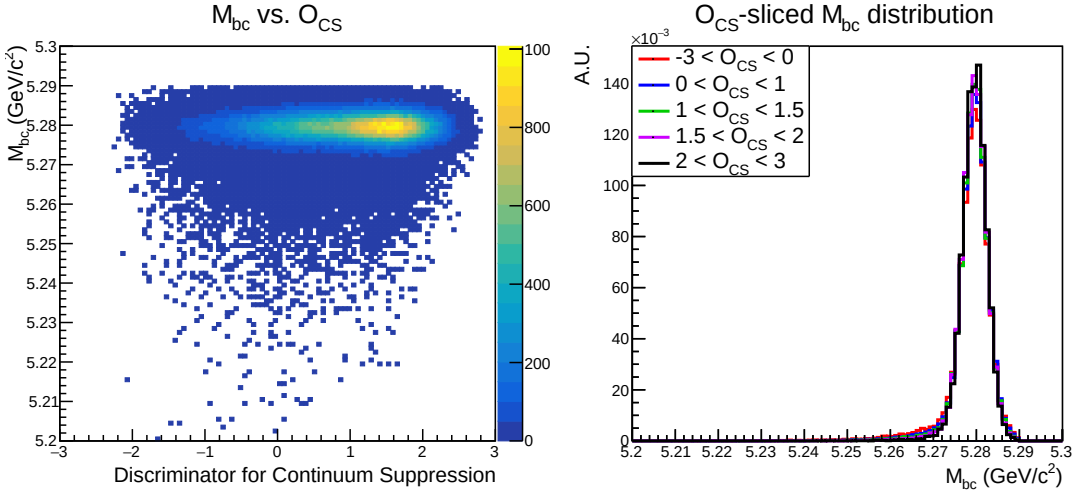


Figure B.5: 2D scatter plot of M_{bc} vs. $\mathcal{O}_{CS}$ (left) and $\mathcal{O}_{CS}$-sliced M_{bc} distribution (right) from signal events in the main channel.

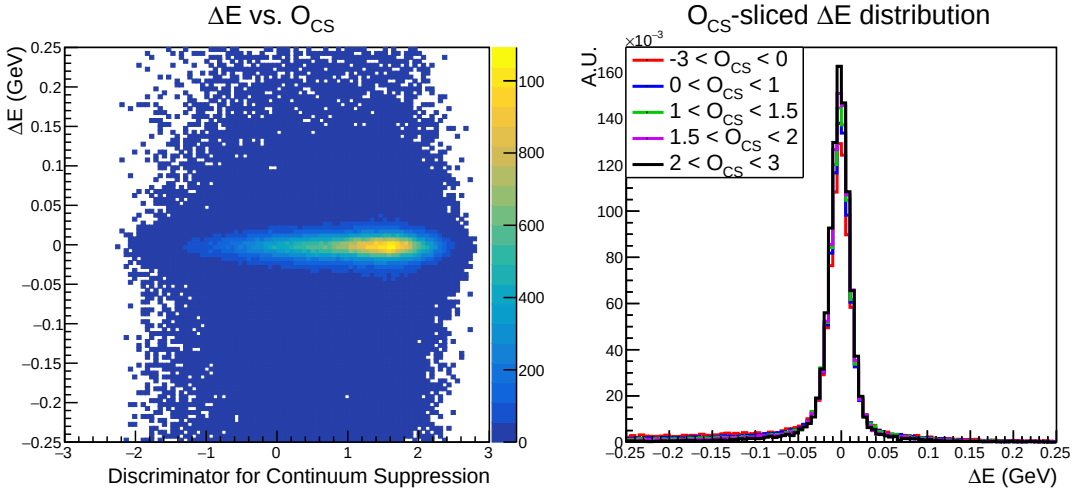


Figure B.6: 2D scatter plot of ΔE vs. $\mathcal{O}_{CS}$ (left) and $\mathcal{O}_{CS}$-sliced M_{bc} distribution (right) from signal events in the main channel.

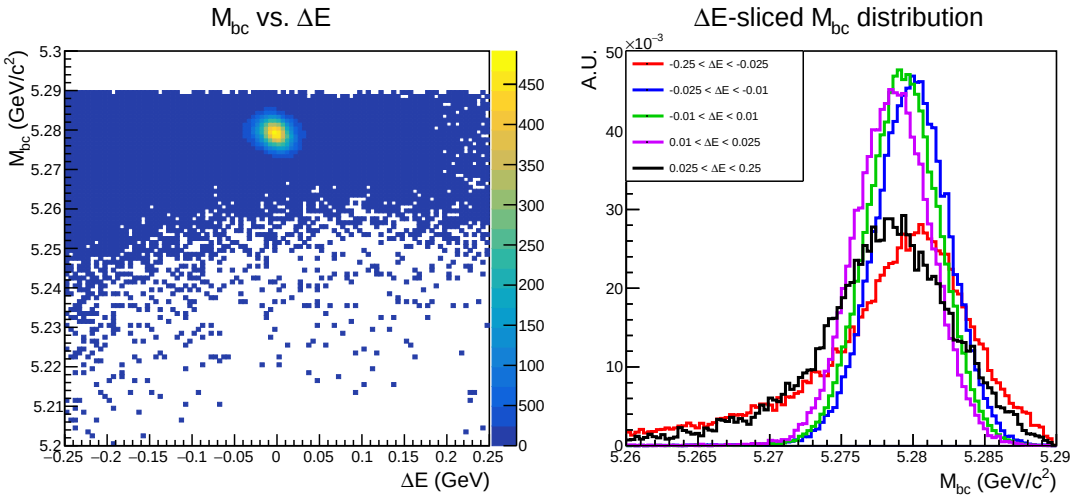


Figure B.7: 2D scatter plot of M_{bc} vs. ΔE (left) and ΔE-sliced M_{bc} distribution (right) from signal events in the control channel.

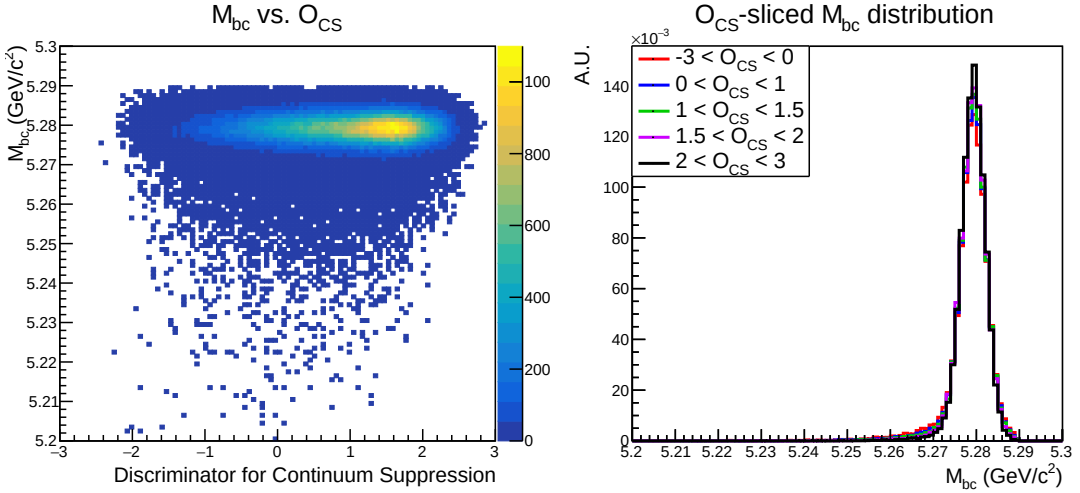


Figure B.8: 2D scatter plot of M_{bc} vs. $\mathcal{O}_{CS}$ (left) and $\mathcal{O}_{CS}$-sliced M_{bc} distribution (right) from signal events in the control channel.

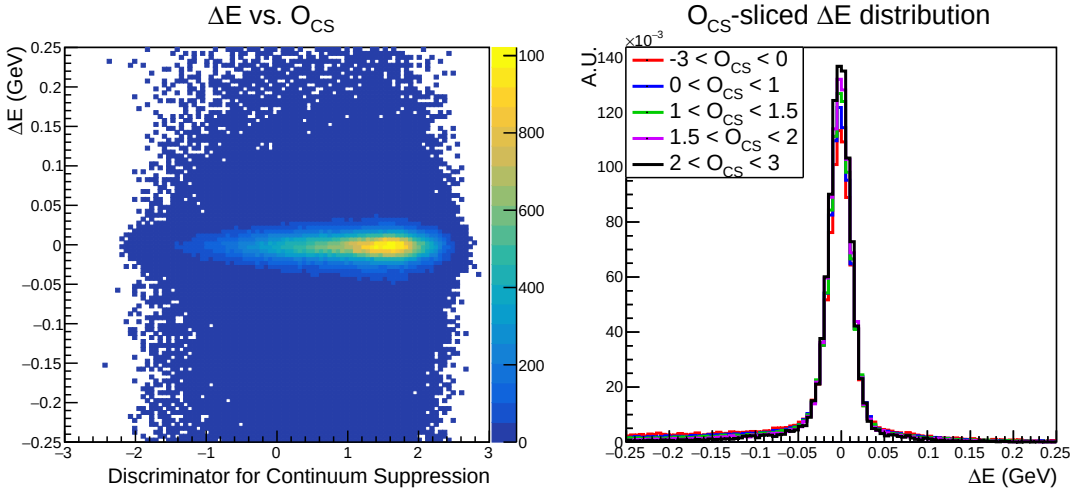


Figure B.9: 2D scatter plot of ΔE vs. $\mathcal{O}_{CS}$ (left) and $\mathcal{O}_{CS}$-sliced M_{bc} distribution (right) from signal events in the control channel.

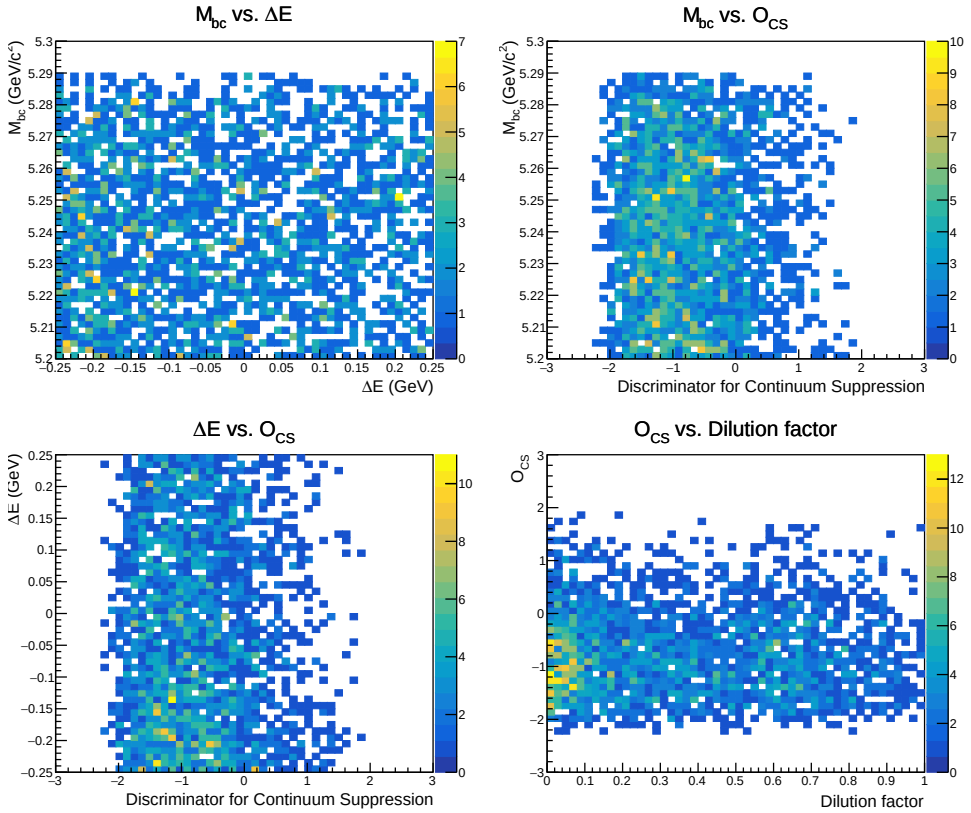


Figure B.10: 2D scatter plots between variables for the signal extraction in the main channel

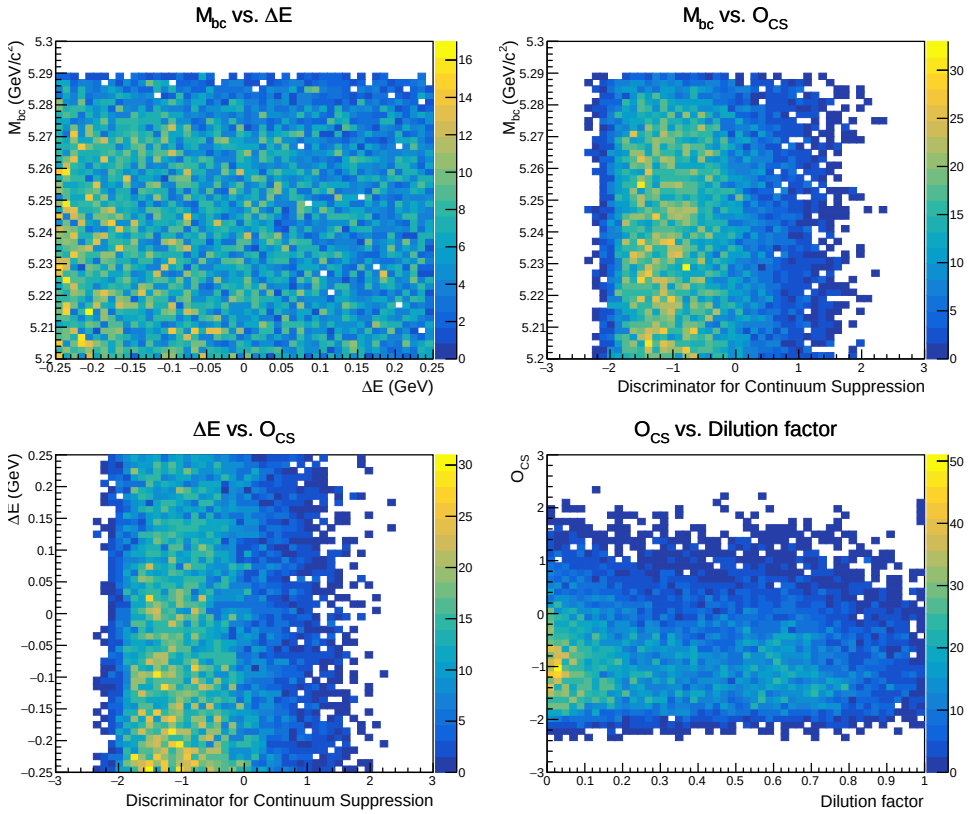


Figure B.11: 2D scatter plots between variables for the signal extraction in the control channel

Appendix C

Procedure of Calibration for Resolution Parameters

In this appendix section, we will discuss the details of the calibration procedure. The detailed procedure of calibration for the resolution parameters will be explained in the following sections.

C.1 Signal MC sample and the pull distribution for the track parameters

The main inputs for the vertex fitting are charged tracks, and each track is parameterized by five track parameters of $(d_0, z_0, \phi_0, \omega, \tan(\lambda))$. We can calculate a pull value p_x of a specific track parameter x of a track using MC-truth information as the following equation:

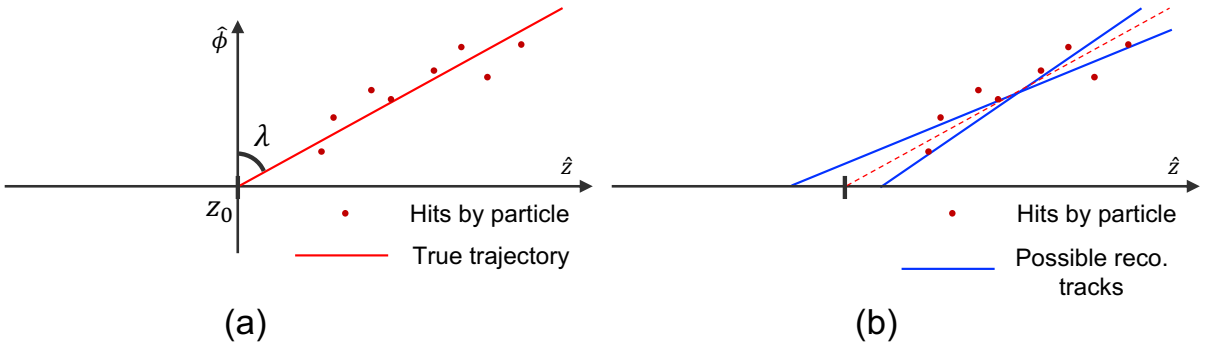


Figure C.1: Graphical interpretation for the situation of $(z_0, \tan(\lambda))$ correlation. Sub-figure (a) shows the true trajectory and hits made by a particle, and Sub-figure (b) indicates z_0 and $\tan(\lambda)$ of possible reconstructed tracks from the track fit are constrained by the hits and negatively correlated.

$$p_x = \frac{x - x_{truth}}{\sigma_x} \tag{C.1}$$

where x_{truth} is a MC-truth value from MC generator information and σ_x is uncertainty of a track parameter x. As shown in Figure C.2, we can confirm that the pulls of track parameters are significantly correlated: (d_0, ϕ_0) and $(z_0, \tan(\lambda))$. We think the correlation is related to the parameterization method of track parameters.

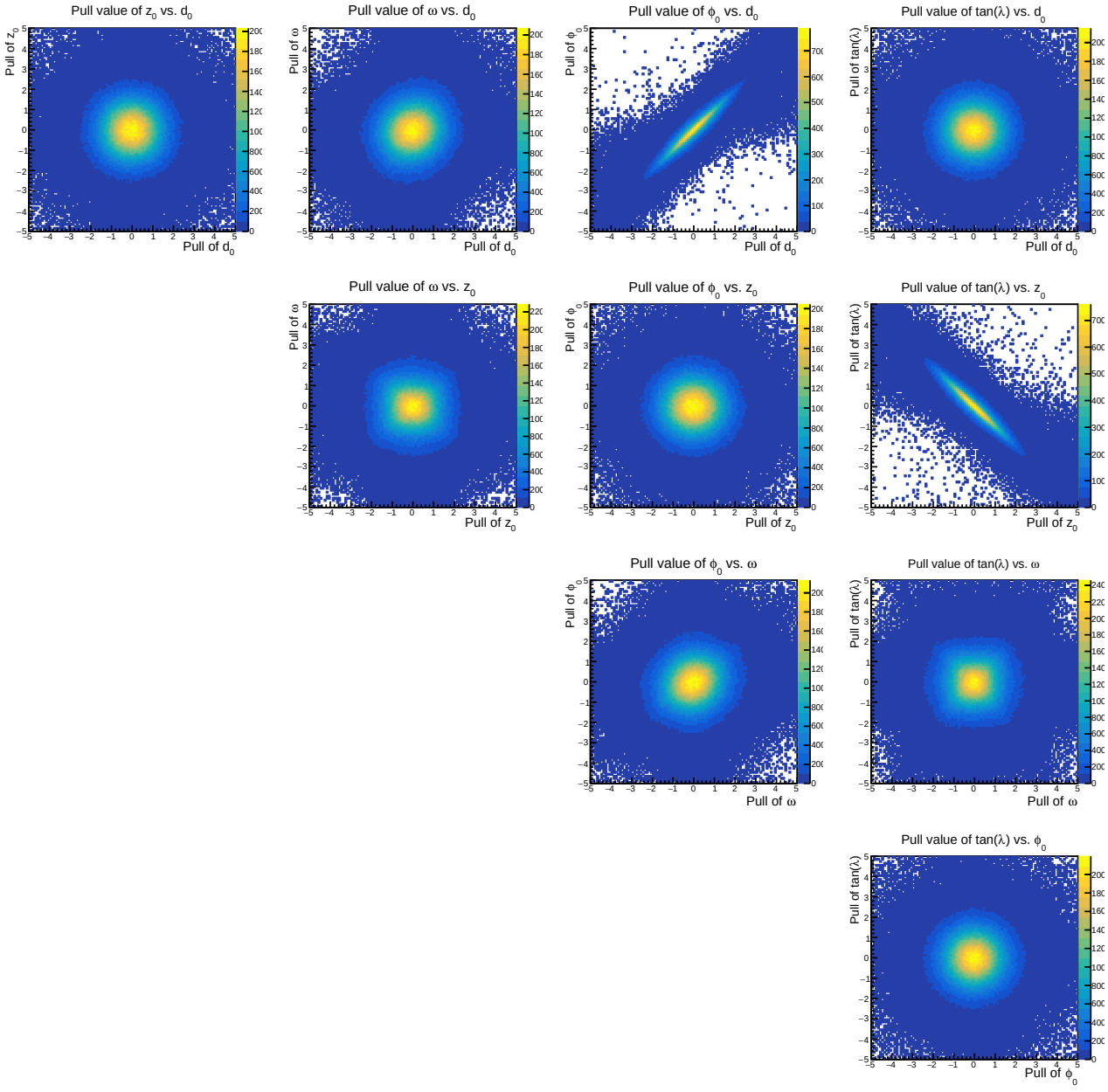


Figure C.2: 2D scatter plots between pull values of each track parameter from four $\pi^{\pm}$ in our signal MC. (d_0, ϕ_0) and $(z_0, \tan(\lambda))$ pairs are strongly correlated.

In the current parameterization, z_0 is the z-coordinate of POCA (Point Of the Closest Approach), and $\tan(\lambda)$ is a kind of the z-direction component of the momentum at POCA.

If we assume the true trajectory of a particle is the red line in Figure C.1(a), then track parameters of z_0 and $\tan(\lambda)$ are constrained by hits made by the particle shown in the figure. Because the hits are made only on one side of the figure, the residual of z_0 and $\tan(\lambda)$ from the true values are negatively correlated, as shown in Figure C.1(b). The correlation of (d_0, ϕ_0) also emerges in a similar fashion. An example plot that describes the situation of $(z_0, \tan(\lambda))$ is shown in Figure C.1.

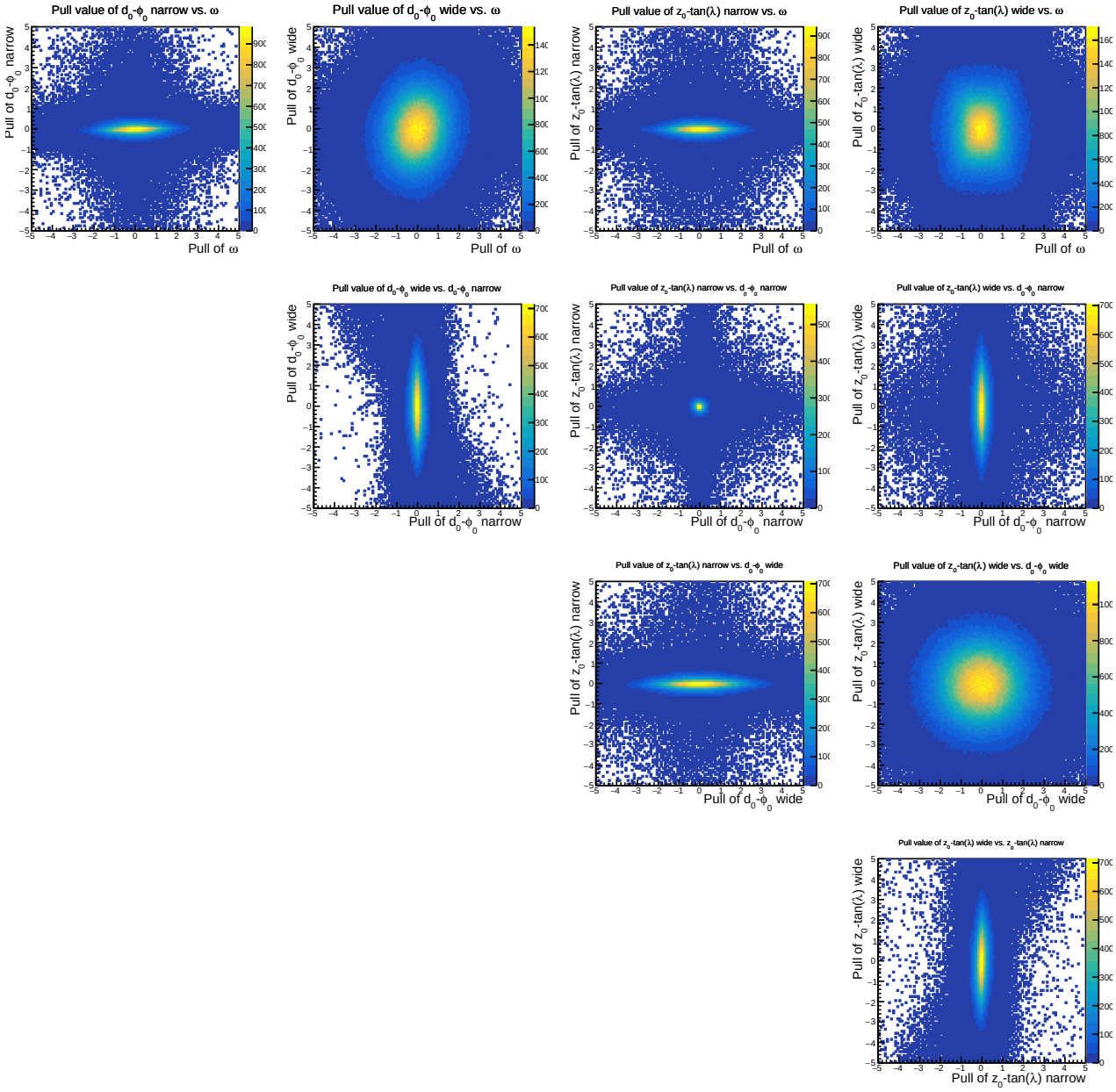


Figure C.3: Transformed distribution for pull values from the signal MC samples. There is a negligible correlation between the pull variables.

We can remove the correlation by rotating 2D distributions of correlated pairs of (d_0, ϕ_0) and $(z_0, \tan(\lambda))$ by $-45°$ and $45°$ with the following equations, respectively.

$$p_{(d_0-\phi_0)\text{narrow}} = \frac{1}{\sqrt{2}}(p_{\phi_0} - p_{d_0}) \quad , \quad p_{(d_0-\phi_0)\text{wide}} = \frac{1}{\sqrt{2}}(p_{\phi_0} + p_{d_0}) \tag{C.2}$$

$$p_{(z_0-\tan(\lambda))\text{narrow}} = \frac{1}{\sqrt{2}}(p_{z_0} + p_{\tan(\lambda)}) \quad , \quad p_{(z_0-\tan(\lambda))\text{wide}} = \frac{1}{\sqrt{2}}(p_{z_0} - p_{\tan(\lambda)}) \tag{C.3}$$

Figure C.3 shows that the correlations are successfully removed. For the further procedure of the calibration, we will only handle the rotated pull value from now.

C.2 Cosmic MC and reproducing the correlation of track parameters

For the cosmic sample, the pull value is calculated in a different way from that in the signal MC sample because there is no reliable MC truth information in the cosmic sample. In the Belle II tracking system, one cosmic particle is reconstructed by two charged tracks by the upper and lower one. Because the two tracks are independent measurements for a single particle (cosmic muon), the pull value p'_x of each track parameter x can be calculated as the following equation.

$$p'_x = \frac{x^{upper} - x^{lower}}{\sqrt{(\sigma_x^{upper})^2 + (\sigma_x^{lower})^2}} \tag{C.4}$$

where σ_x is the uncertainty of a tracking parameter x and $'$ on superscript indicates the pull value is the cosmic type.

We simulated cosmic MC events or analyzed cosmic events collected during the physics run and reconstructed the events by using a particular reconstruction module in basf2 named `add_cosmics_reconstruction` and `CDCCosmicAnalysis` that are stored in our git repository. For the reconstruction of the cosmic events, we required one PXD hit at least and $|d_0| < 1\text{cm} \wedge -2\text{cm} < z_0 < 4\text{cm}$ for all tracks. We fitted the rotated cosmic pull of $(\omega, N_{(d_0-\phi_0)}, W_{(d_0-\phi_0)}, N_{(z_0-\tan(\lambda))}, W_{(z_0-\tan(\lambda))})$ to deal with the cosmic pull values in the same footprint as in the signal MC sample. N means narrow-side variables and W means wide-side variables.

$$P(p'_x) = f^x_{major}\text{Gaus}(p'_x|\mu = \mu^x_{major}, \sigma = \sigma^x_{major}) + (1 - f^x_{major})\text{Gaus}(p'_x|\mu = \mu^x_{minor}, \sigma = \sigma^x_{minor}) \tag{C.5}$$

where the fraction of $f_{major} > 0.5$. Distributions of cosmic-type pull value and fitted curves are shown in Figures C.4 and C.5 from cosmic MC and Data, respectively.

The important feature of the cosmic pull value is that there are no correlations in the 2D distributions, as shown in Figure C.6. Thus, if we want to consider the structure of pull distribution for track parameters properly in the calibration, we should reproduce the

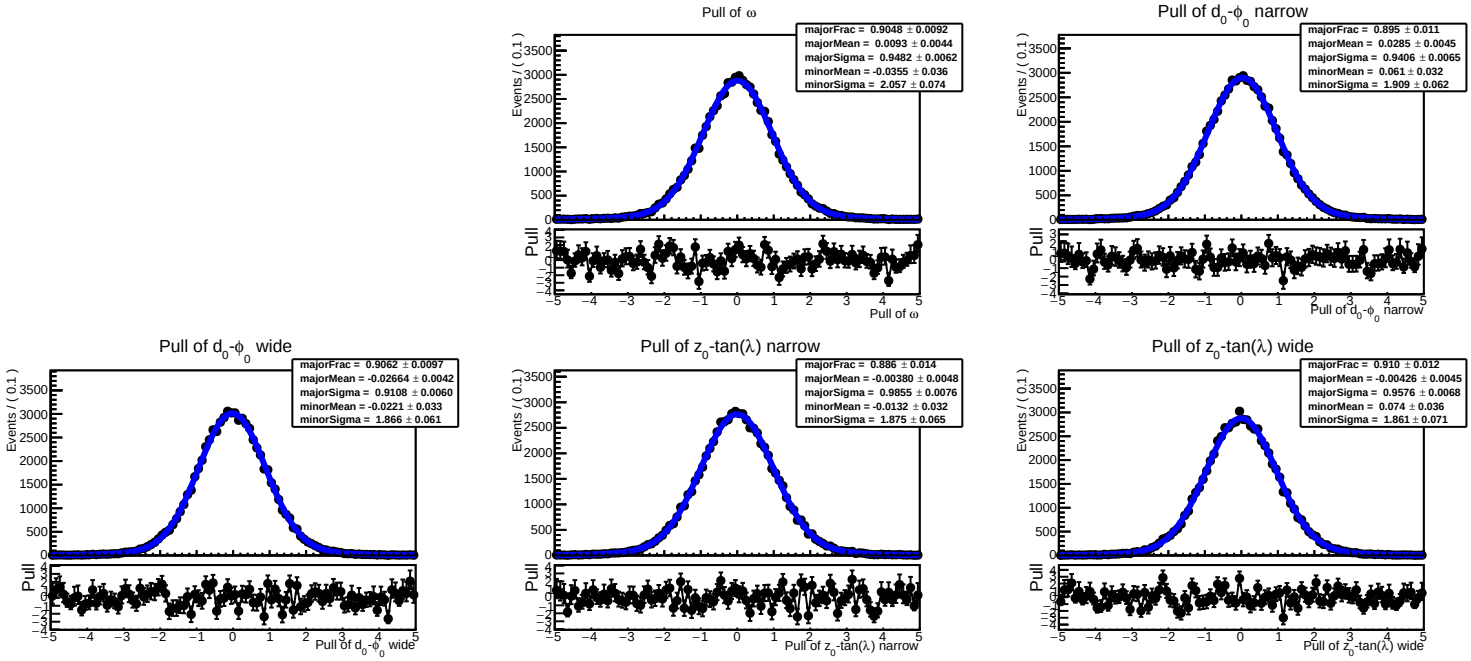


Figure C.4: Cosmic-type pull distribution of track parameters and its fitting result(curve and parameters) from cosmic MC.

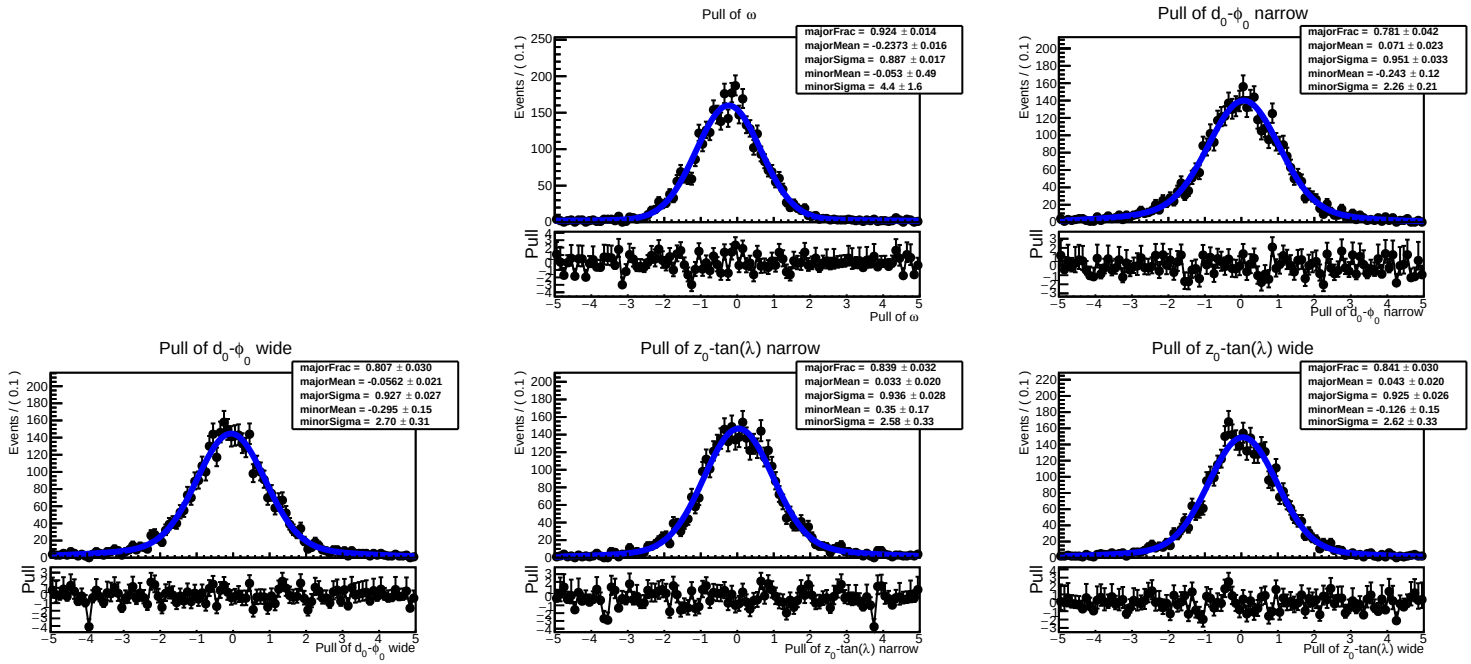


Figure C.5: Cosmic-type pull distribution of track parameters and its fitting result(curve and parameters) from cosmic data.

correlation of track parameters in the signal MC samples. We reproduced the correlation using specially simulated events in the Belle II simulation framework, named “Back-to-Back MC sample”.

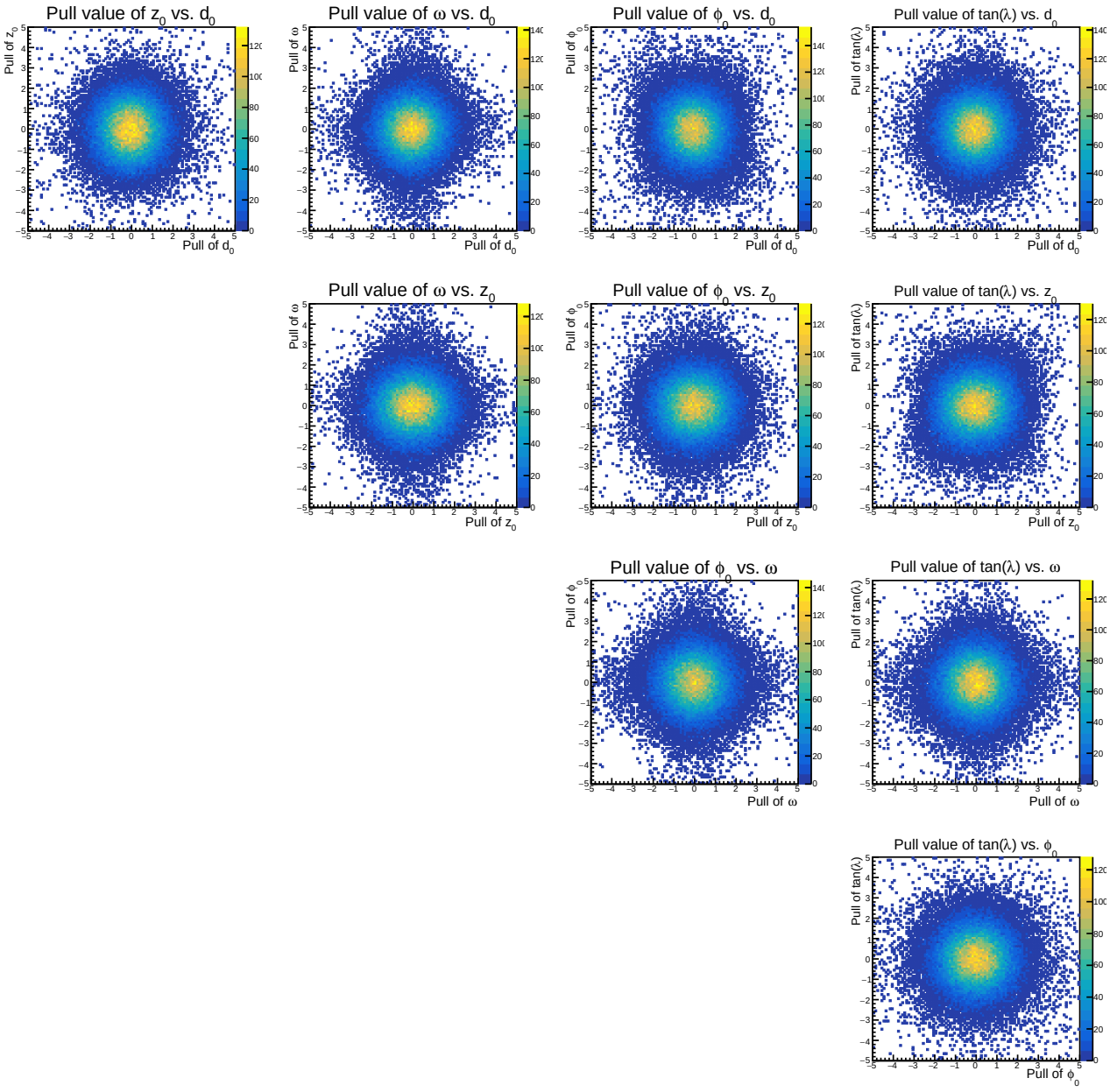


Figure C.6: 2D scatter plots between cosmic-type pull values of each track parameter from cosmic MC.

C.2.1 Back-to-Back MC sample

A Back-to-Back event is an MC event that oppositely-charged μ goes in the exact reverse direction from a point around the IP within the beam pipe. In detail, we generated a pair of $\mu^{\pm}$ in a random point within the range of $\sqrt{x^2+y^2} < 1\text{cm}$ and $-2\text{cm} < z < 4\text{cm}$. Then, the momentum of each μ is determined from the uniform distribution within a range of $0.3\,\text{GeV}/c < p_\mu < 3.0\,\text{GeV}/c$ without any boost from e^+e^-. An example event from the Back-to-Back MC generator is shown in Figure C.7.

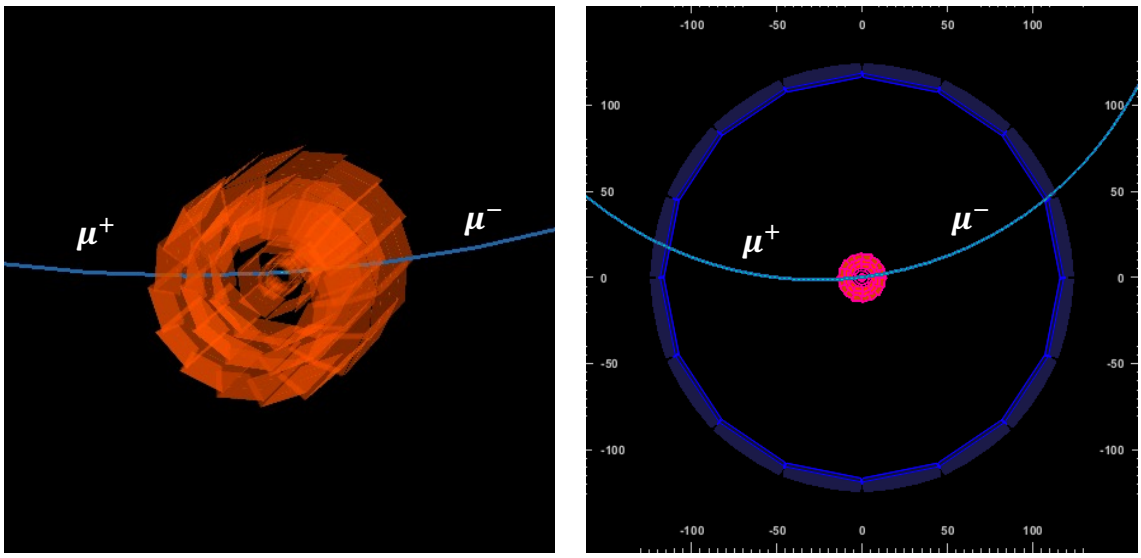


Figure C.7: An example event from Back-to-Back sample. The pair of oppositely-charged μ is generated around the IP and goes in the opposite direction.

These events are similar to cosmic events as shown in Figure 5.16, but we can calculate both pull values of signal MC (p) and that of cosmic events (p'). Using this feature, we extracted some scale factors to reproduce the correlation of the signal type pull distribution p for the cosmic MC or Data.

C.2.2 Extraction of the scale factors

We simulated and reconstructed the Back-to-Back events and got distributions of p and p' for each transformed track parameter of $(\omega, N_{(d_0-\phi_0)}, W_{(d_0-\phi_0)}, N_{(z_0-\tan(\lambda))}, W_{(z_0-\tan(\lambda))})$. In the reconstruction, we applied the reconstructed tracks should have one PXD hit at least. Then, we fitted the pull distributions with the function in Equation C.5 within $(-5, 5)$ range but $(-1, 1)$ for narrow-side variables of signal type pulls and calculated correction factors using the fitting results. The fitted distribution and parameters from Back-to-Back MC are shown in Figures C.9 and C.8 for the signal type pull p_x and cosmic type pull p'_x, respectively.

Because of that, the correlation originates from the parameterization of a track, and MC and Data use the same parameterization; we can make an assumption that the ratio

of sigma of the major Gaussian component for the wide variable to that for the narrow one will be the same between MC and Data. Also, we assumed that the overall shape of the signal type pull distribution for the narrow side variables, such as the fraction of the major Gaussian (f^N_{major}) or ratio of $\sigma^N_{minor}/\sigma^N_{major}$, is the same between the Back-to-Back and cosmic samples. Thus, we extracted a correction factor from the Back-to-Back sample as follows:

- $d^{Wide,x} = (1-f^{p_{Wide,x}}_{major})/(1-f^{p'_{Wide,x}}_{major})$: A scale factor for a fraction of major Gaussian for wide variables between the signal and cosmic type pull distribution.
- $s^{Wide,x}_{major,minor} = \sigma^{p_{Wide,x}}_{major,minor}/\sigma^{p'_{Wide,x}}_{major,minor}$: A scale factor for σ of major or minor Gaussian signal and cosmic type pull distribution.
- $f^{Narrow,x}_{major} = f^{p_{Narrow,x}}_{major}$: Fraction of major Gaussian of narrow-side signal type pull distribution.
- $s^{Wide\to Narrow}_{major,x} = \sigma^{p_{Narrow,x}}_{major}/\sigma^{p_{Wide,x}}_{major}$: A scale factor for σ of major Gaussian between wide and narrow-side signal type pull distribution.
- $s^{Narrow,x}_{major\to minor} = \sigma^{p_{Narrow,x}}_{minor}/\sigma^{p_{Narrow,x}}_{major}$: A scale factor for σ between major and minor Gaussian for narrow-side signal type pull distribution.

where x is a pair of correlated pull variables.

Table C.1 shows each correction factor determined by Back-to-Back MC samples.

Parameter	Value	Parameter	Value
$d^{Wide,(d_0-\phi_0)}$	0.461	$f^{Narrow,(d_0-\phi_0)}_{major}$	0.824
$s^{Wide,(d_0-\phi_0)}_{major}$	1.443	$s^{Wide\to Narrow}_{major,(d_0-\phi_0)}$	0.314
$s^{Wide,(d_0-\phi_0)}_{minor}$	1.631	$s^{Narrow,(d_0-\phi_0)}_{major\to minor}$	0.409
$d^{Wide,(z_0-\tan(\lambda))}$	0.807	$f^{Narrow,(z_0-\tan(\lambda))}_{major}$	0.723
$s^{Wide,(z_0-\tan(\lambda))}_{major}$	1.397	$s^{Wide\to Narrow}_{major,(z_0-\tan(\lambda))}$	0.352
$s^{Wide,(z_0-\tan(\lambda))}_{minor}$	1.371	$s^{Narrow,(z_0-\tan(\lambda))}_{major\to minor}$	0.421

Table C.1: Determined correction factor from Back-to-Back MC sample to reproduce the correlation of signal type pull.

C.2.3 Making the correlation

We reproduced the correlation by applying the scale factor from the previous paragraph. [1] Firstly, we scaled parameters of double Gaussian for wide-side pull distributions from cosmic samples. The parameters are obtained from a fitting with a double Gaussian function to cosmic MC and Data samples, and the fitting results are shown in Figures C.4 and C.5. Then, the factors are applied to the parameters as follows:

[1] In this procedure, we did not correct parameters for ω because there is a negligible difference in the shape between residual and cosmic type pull distribution and no correlation with other track parameters.

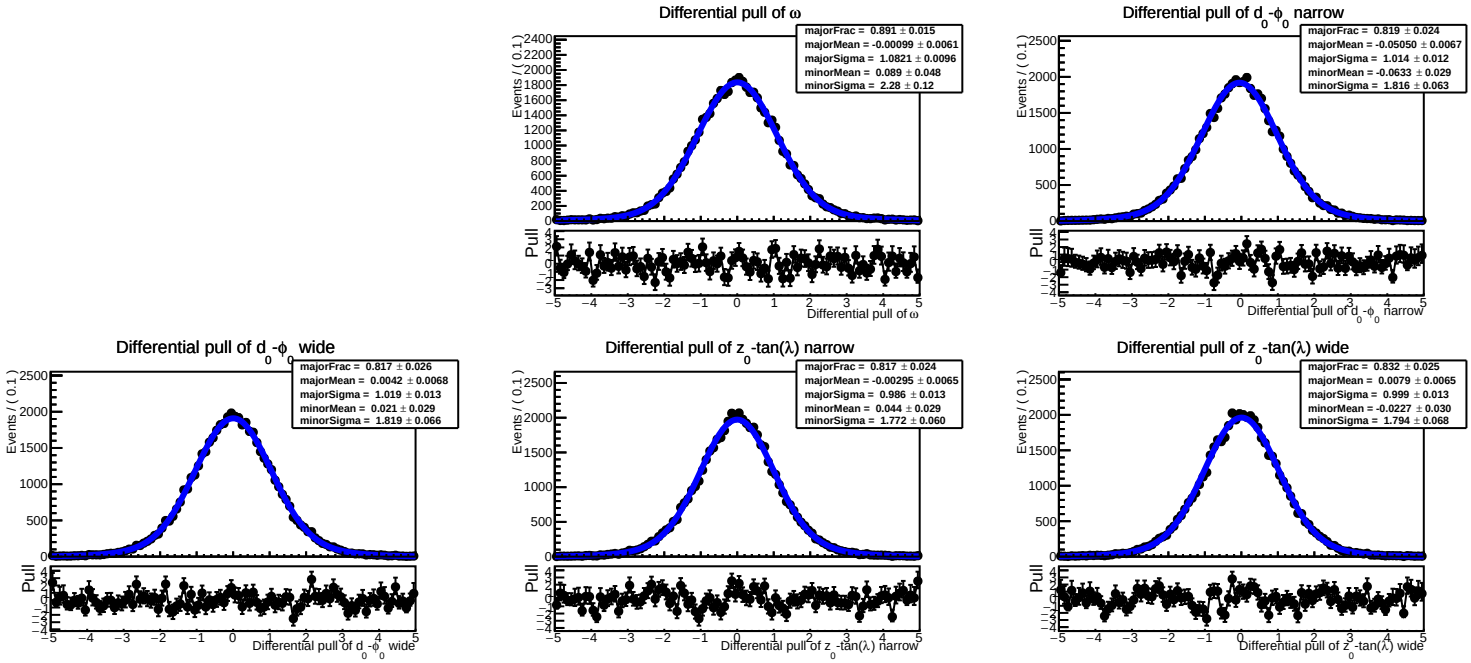


Figure C.8: Cosmic type pull distribution for transformed track parameters from Back-to-Back MC samples. Plotting and fitting range are the same as $(-5, 5)$.

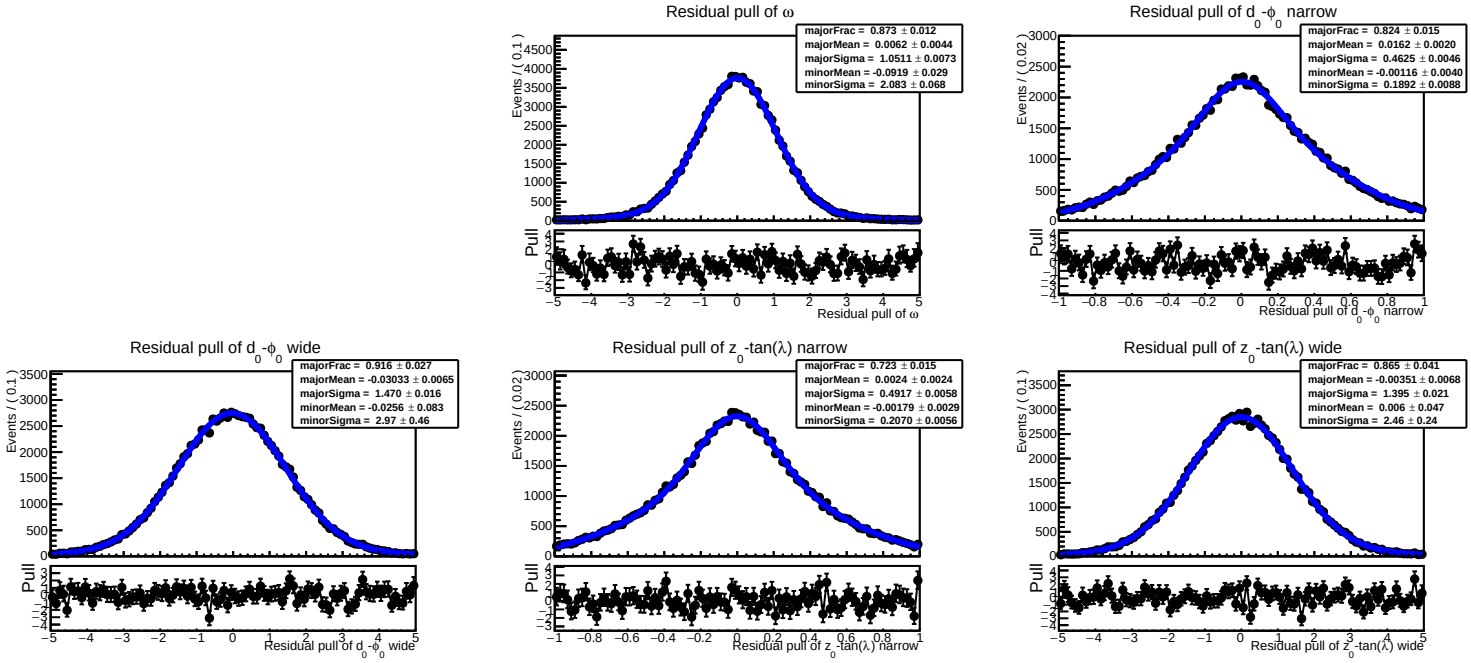


Figure C.9: Signal type pull distribution for transformed track parameters from Back-to-Back MC samples. The range for fitting narrow-side variables is $(-1, 1)$, and that for others is $(-5, 5)$. The plotting and fitting range are the same.

$$f_{major}^{Wide,x} = d^{Wide,x} \cdot (1 - (f_{major}^{Wide,x})') \quad \text{(C.6)}$$
$$\sigma_{major}^{Wide,x} = s_{major}^{Wide,x} \cdot (\sigma_{major}^{Wide,x})' \quad \text{(C.7)}$$
$$\sigma_{minor}^{Wide,x} = s_{minor}^{Wide,x} \cdot (\sigma_{minor}^{Wide,x})' \quad \text{(C.8)}$$

where x is a pair of correlated pull variables and f', σ' are parameters of the Gaussian function for the cosmic type pull distribution from the cosmic samples.

After calculating the parameters of Gaussian functions for wide-side pulls, we determined that for narrow-side pull values as following equations.

$$f_{major}^{Narrow,x} = f_{major}^{p_{Narrow,x}} \quad \text{(C.9)}$$
$$\sigma_{major,x}^{Narrow} = s_{major,x}^{Wide \to Narrow} \cdot \sigma_{major,x}^{Wide} \quad \text{(C.10)}$$
$$\sigma_{minor,x}^{Narrow} = s_{major \to minor}^{Narrow,x} \cdot \sigma_{major,x}^{Narrow} \quad \text{(C.11)}$$

Using this procedure, we got full parameters to reproduce the correlation of signal type pull distributions with parameter inputs from the pull distributions of the cosmic MC sample. We generated the wide and narrow-side pull variables independently and rotated back the generated $(N_{d_0-\phi_d}, W_{d_0-\phi_d})$ and $(N_{z_0-\tan(\lambda)}, W_{z_0-\tan(\lambda)})$ variables using the inverse transformation of Equations C.2 and C.3, respectively. Figure C.10 shows the generated and rotated back pull distribution from the procedure stated above, and we can see the correlation.

C.3 Fast Simulation using Signal MC

We performed "Fast Simulation" using distributions from cosmic MC or Data samples, to generate track parameters. Fast Simulation is a procedure to get resolution parameters that reflect the tracking resolution of the detector from MC or Data, which consists of two parts: a special reconstruction of signal events in the signal MC samples and fitting distribution of the residual variable δt to get the resolution parameters. In the first part, we replaced five track parameters of all charged tracks in an event with the randomly generated values from pull distributions from cosmic MC or Data samples that the scale factors applied to reproduce the correlation between some pull values of track parameters, as discussed in the previous subsection. For example, we generated the track parameters from distributions shown in Figure C.10 to reflect the tracking resolution of the cosmic MC sample in the signal reconstruction procedure. With the randomized track parameters, we continued the nominal processes for reconstruction described in Section 4, such as the build and selection of B^0 candidates, the vertex fitting, and the best candidate selection.

In the next part after the special reconstruction, we fitted distributions of the residual δt variable to get parameters of the resolution functions as discussed in Section 5. Those

parameters from reconstruction results of which the track parameters are randomized using cosmic MC or Data samples reflect the detector resolution in MC or Data. Table C.2 shows the resolution parameters from cosmic MC and Data. [2]

[2]Parameters for a fraction of δ and left-right exponential functions in the tag side were omitted because we will not correct those parameters

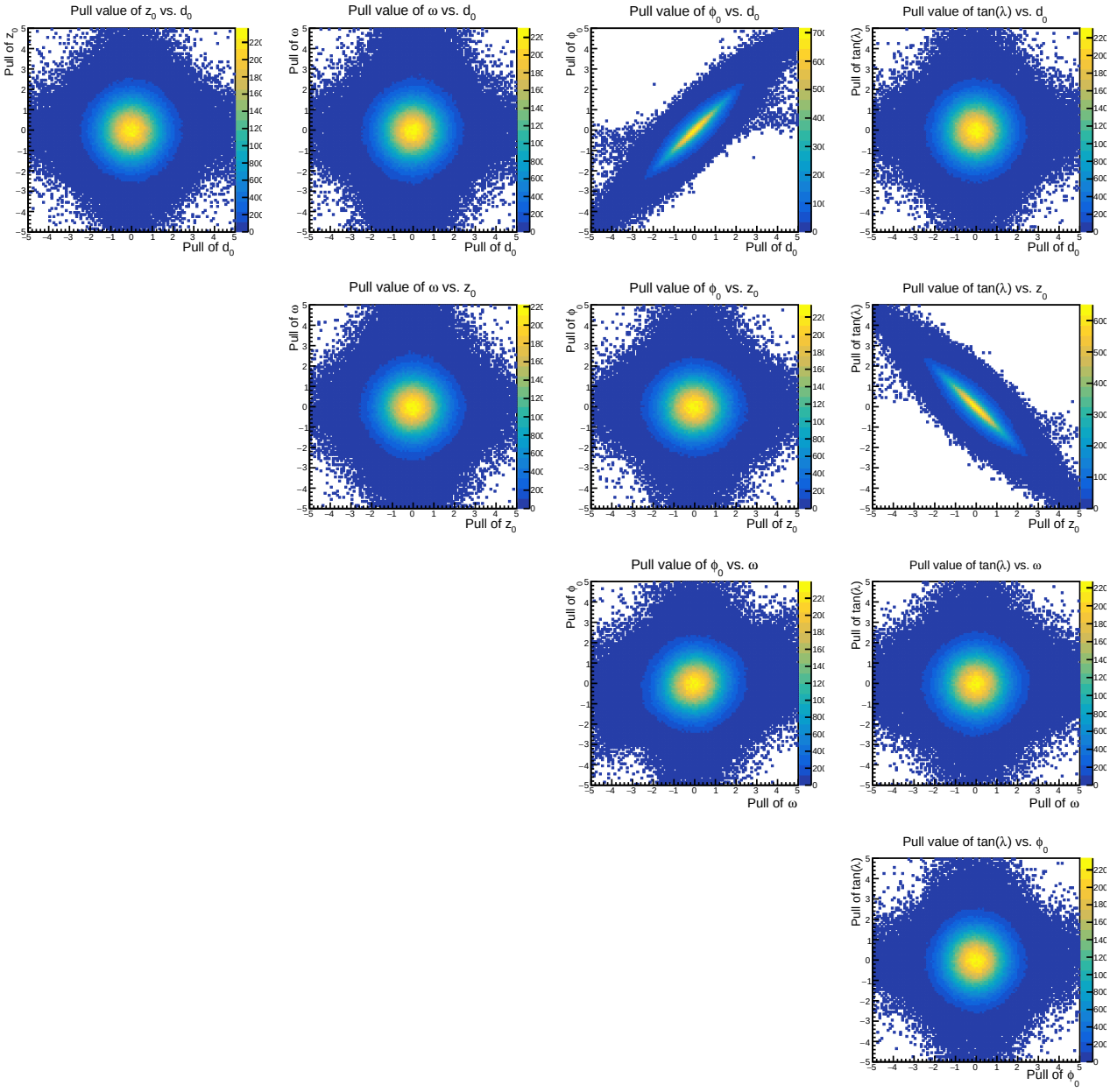


Figure C.10: 2D scatter plots between cosmic type pull values of each track parameter from cosmic MC.

Parameters	Resolution parameters from Fast Simulation using		Correction factors	
	Cosmic MC	Cosmic Data	Method	Value
f_{core}^{CP}	0.88667	0.70388	Scale factor	0.793845
c_{core}^{CP}	0.98315	0.87500	Bias factor	-0.108157
s_{core}^{CP}	0.05975	0.11199	Scale factor	1.87433
c_{tail}^{CP}	1.12654	1.06470	Bias factor	-0.0618429
s_{tail}^{CP}	0.28572	0.36984	Scale factor	1.29441
$\tau_{\mathrm{exp},const}^{tag}$	1.48926	1.59702	Bias factor	0.107755
$\tau_{\mathrm{exp},slope}^{tag}$	0.23156	0.20564	Scale factor	0.888067
$\tau_{\mathrm{exp},ceil}^{tag}$	5.40368	9.53307	Bias factor	4.1294
c_{δ}^{tag}	1.02380	0.98907	Bias factor	-0.0347335
s_{δ}^{tag}	0.12860	0.17796	Scale factor	1.38386
c_{np}^{tag}	0.42113	0.41865	Bias factor	-0.00248241
s_{np}^{tag}	0.18590	0.16968	Scale factor	0.912705

Table C.2: Resolution parameters from Fast Simulation using both cosmic samples as inputs for the track parameters and correction factors for resolution parameters.

C.4 Limitation of Fast Simulation

In this section, we explain the limitation of Fast Simulation and discuss room for improvement of the calibration.

C.4.1 Resolution parameter after Fast Simulation

The current Fast Simulation with inputs from cosmic MC cannot reproduce the resolution parameters from the signal MC samples. Currently, the resolution parameters from Fast Simulation with inputs of cosmic MC samples are inconsistent with the resolution parameters directly determined from the signal MC samples. We consider one of the reasons is inaccurate modeling of pull distributions of tracking parameters. Especially, modeling the pull distribution of narrow side parameters requires more sophisticated functions. We use the double Gaussian to model distribution for simplicity, and it may occur the loss of tail information for the pull distribution of track parameters, which leads to inaccurate reproducing of resolution parameters. Thus, improving the modeling can be helpful in implementing accurate Fast Simulation. Ultimately, perfect Fast Simulation with input from cosmic MC samples can reproduce the resolution parameters that are consistent with those from the signal MC samples, and it will simplify the entire calibration procedure by removing the use of signal MC samples in the calibration.

C.4.2 Bias of track parameter ω

In Fast Simulation using the track parameters from cosmic samples, we assume no bias in determined track parameters and set $\mu = 0$ of the Gaussian functions for Fast Simulation for the calibration since the main purpose of the calibration is to examine the difference in resolution parameters using the resolution of track parameters from cosmic MC/Data sample. However, we confirmed a significant bias in the pull value of ω for track parameters from signal MC samples, as shown in Figure C.11. We suspect the reason for this problem is an inexact B-field mapping, but we did not fully understand. Solving this problem requires a sophisticated understanding of the reconstruction process of tracks, but at the current statistics, we consider the effect of this factor negligible. Thus, the effect of this factor on the *CP* fitting is currently dealt with a systematic uncertainty. However, this limitation can be improved with a deeper understanding of the track reconstruction system of Belle II, and we expect this will reduce the systematic uncertainty regarding the resolution calibration.

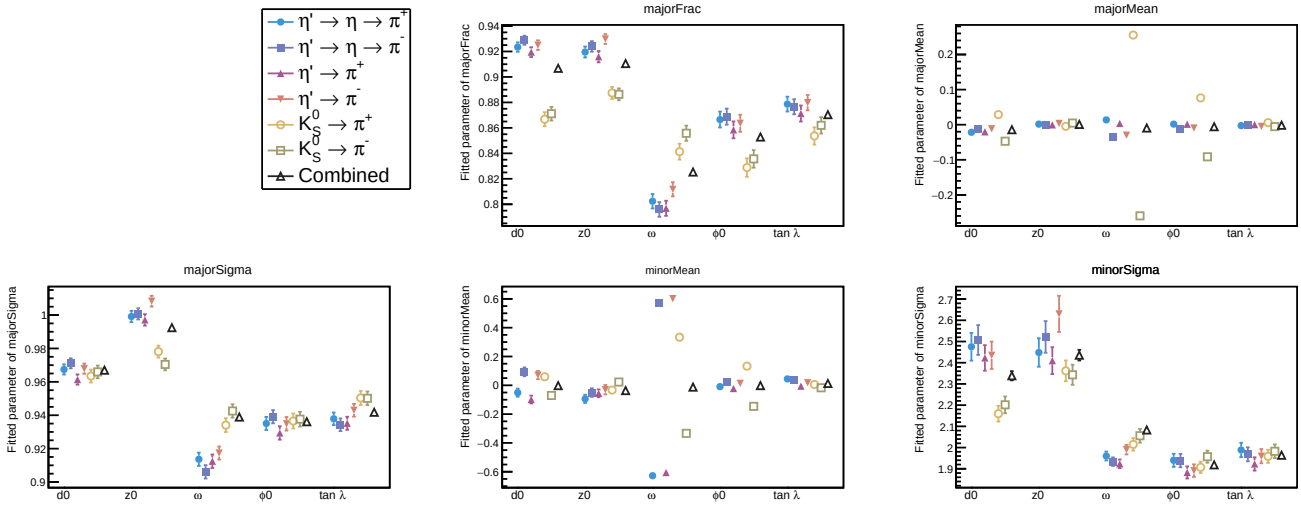


Figure C.11: Fitted track parameters for charged tracks from signal MC samples our sub-channel of $B^0 \to \eta' K^0_S$ decay. ω shows large μ values.

Appendix D

Efficiency Study for the Experimental Data

We computed Data/MC ratio for K^0_S, neutral pions, charged pions, and charged kaons to calculate proper signal yield in the experimental data. We considered the dependence of the ratio to the momentum, angle, and flight distance (only for K^0_S) in the calculation for the total Data/MC efficiency ratio. The Data/MC efficiency ratio for B^0 and $B^\pm$ is calculated with the following equations:

$$\epsilon^{B^0}_{Data/MC} = \epsilon^{\pi^\pm}_{Data/MC} \cdot \epsilon^{\pi^0}_{Data/MC} \cdot \epsilon^{K^0_S}_{Data/MC} \quad \text{(D.1)}$$

$$\epsilon^{B^\pm}_{Data/MC} = \epsilon^{\pi^\pm}_{Data/MC} \cdot \epsilon^{\pi^0}_{Data/MC} \cdot \epsilon^{K^\pm}_{Data/MC} \quad \text{(D.2)}$$

where $\epsilon^{X}_{Data/MC}$ is the Data/MC ratio of efficiency for the particle X and the value of $\epsilon^{B^0}_{Data/MC} = 0.8436 \pm 0.0488$ and $\epsilon^{B^\pm}_{Data/MC} = 0.8570 \pm 0.0691$. We will discuss the details of how to calculate each component of the equation above in the following sections.

D.1 Charged pions and kaons

We examined the Data/MC efficiency ratio for charged pions and kaons with a selection of `pionID > 0.1` for $\pi^\pm$ and `kaonID > 0.1` for $K^\pm$. The ratio was calculated using Systematic Framework provided by the Belle II PID group with binning of the following boundary for the momentum and its polar angle:

$$p_\pi : (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 1.0, 1.2, 1.8) \quad \text{(D.3)}$$

$$p_K : (1.8, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.4, 3.8) \quad \text{(D.4)}$$

$$\theta_{\pi,K} : (0, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.7, 1.9, 2.1, 2.6) \quad \text{(D.5)}$$

where $p_{\pi,K}$ is the momentum in the lab frame of particles and $\theta_{\pi,K}$ is the polar angle of the momentum.

Figures D.2 and D.3 show the Data/MC ratio for charged pions and kaons, respectively, with the binning above. We calculated the total averaged ratio for charged pions originating from η' and η, and kaons from $B^{\pm}$ by averaging the ratios of all individual particles calculated with the momentum and the polar angle. The 2D distribution of ($p_{\pi,K}$ vs. $\theta_{\pi,K}$) of the particles from the signal MC samples is shown in Figure D.4. Then, the total averaged Data/MC ratio is calculated by multiplying all the ratios of particles for the reconstruction of B^0 or $B^{\pm}$. The result is shown in Table D.2.

π^0 momentum	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1.0
Efficiency	0.903 ± 0.039	0.963 ± 0.036	0.987 ± 0.082	1.013 ± 0.058
π^0 momentum	1.0-1.5	1.5-2.0	2.0-3.0	
Efficiency	1.042 ± 0.060	1.045 ± 0.065	1.011 ± 0.050	

Table D.1: Data/MC ratio for π^0 per bins of the momentum with 40% π^0 efficiency list[60].

Main channel		Control channel	
Particle and its parents	Data/MC ratio	Particle and its parents	Data/MC ratio
$\eta' \to \pi^+$	0.9816 ± 0.0095	$\eta' \to \pi^+$	0.9816 ± 0.0095
$\eta' \to \pi^-$	0.9819 ± 0.0096	$\eta' \to \pi^-$	0.9819 ± 0.0096
$\eta' \to \eta \to \pi^+$	0.9817 ± 0.0096	$\eta' \to \eta \to \pi^+$	0.9822 ± 0.0097
$\eta' \to \eta \to \pi^-$	0.9819 ± 0.0097	$\eta' \to \eta \to \pi^-$	0.9823 ± 0.0097
Total Data/MC ratio for $\pi^{\pm}$	0.9292 ± 0.0182	Total Data/MC ratio for $\pi^{\pm}$	0.9300 ± 0.0182
		$B^{\pm} \to K^{\pm}$	0.9676 ± 0.0547

Table D.2: Data/MC ratio for each charged particle and total averaged Data/MC ratio for the main and control channel.

D.2 Neutral pions

Using the momentum of π^0, we checked the Data/MC ratio for the particle. The binning and ratio for π^0 with 40% efficiency list are shown in Table D.1. We calculated the ratio with π^0 momentum using a method similar to that in charged particles. The distribution of π^0 momentum is given in the left plot of Figure D.1, and the total averaged Data/MC efficiency is 0.9533 ± 0.0519 for the main channel and 0.9523 ± 0.0516 for the control channel.

D.3 K_S^0 (only for the main channel)

The Data/MC ratio for K_S^0 is evaluated with the flight length of the particle. A dependency of the ratio on the flight length is given as a formula of $\epsilon_{Data/MC} = c_0 + s_0 \cdot \ell_{K_S^0}$ where

$\ell_{K_S^0}$ is the flight length of K_S^0 and $c_0 = 1.008, s_0 = -0.00546/\text{cm}$ which are determined in K_S^0 efficiency study in the channel of $D^{*+} \to [D^0 \to K_S^0\pi^+\pi^-]\pi^+$. We calculated the ratio with the distribution of the flight distance of K_S^0 from the signal MC sample shown in the right plot of Figure D.1, and the result is 0.9523. The value is also consistent with the equivalent result estimated in the K_S^0 reconstruction study in the channel of $B^0 \to K_S^0\pi^0$, $0.9524 \pm 0.0034(\text{stat.}) \pm 0.0118(\text{norm.}) \pm 0.0151(\text{fit model})$.

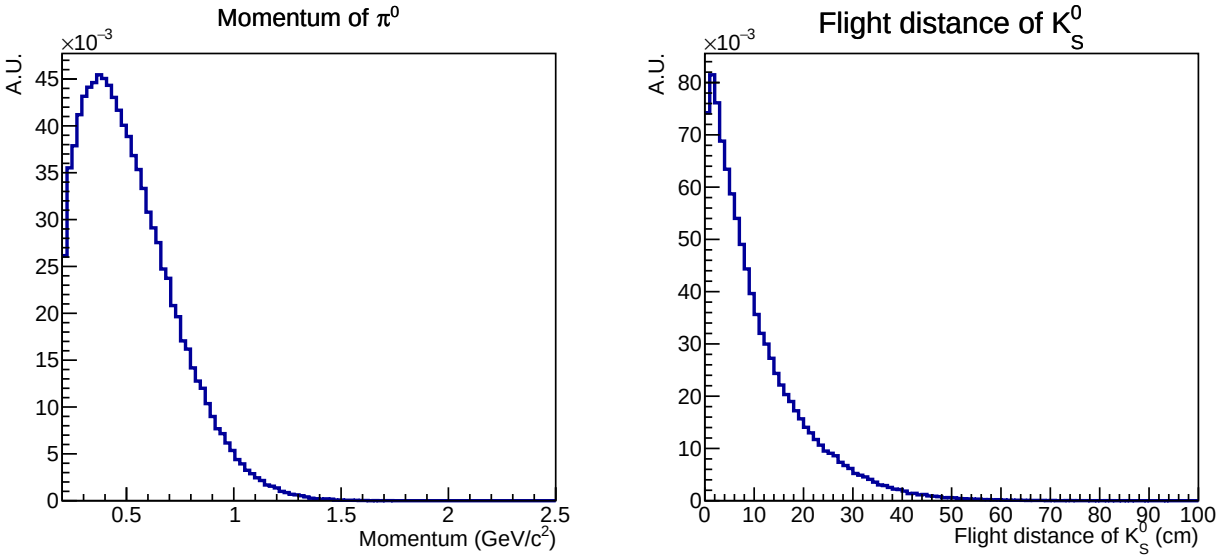


Figure D.1: Distribution for π^0 momentum (left) and flight distance of K_S^0 (right) in the main channel from the signal MC sample.

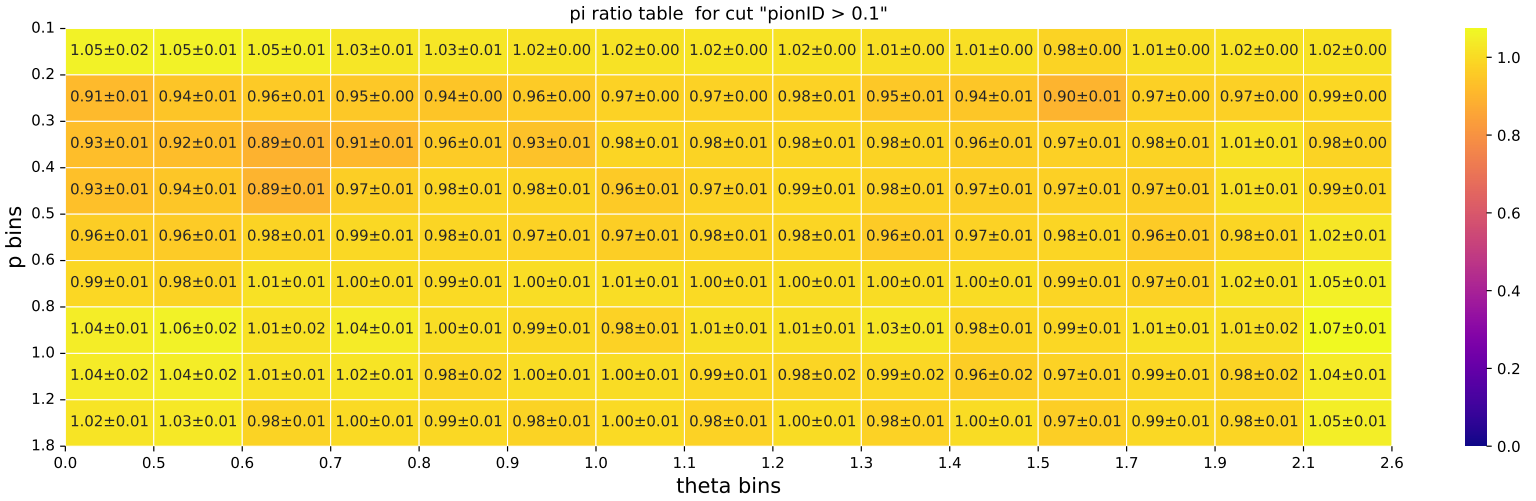


Figure D.2: The Data/MC ratio for charged pions with bin boundaries given in D.5 for the polar angle (X-axis) and D.3 for the momentum (Y-axis).

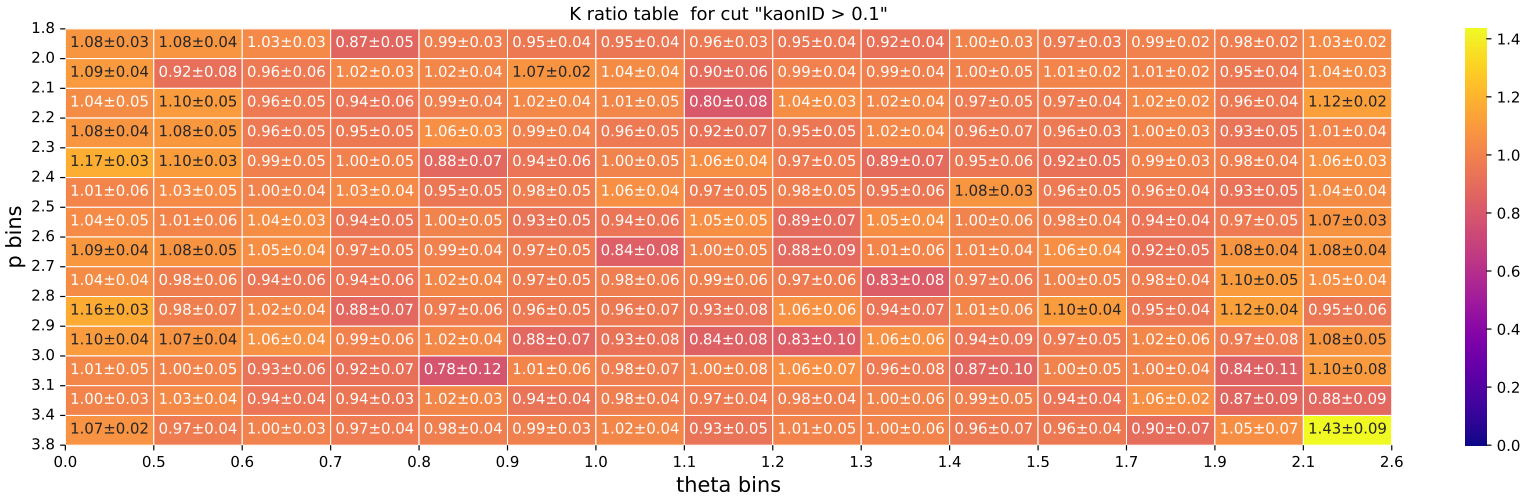


Figure D.3: The Data/MC ratio for charged kaons with bin boundaries given in D.5 for the polar angle (X-axis) and D.4 for the momentum (Y-axis).

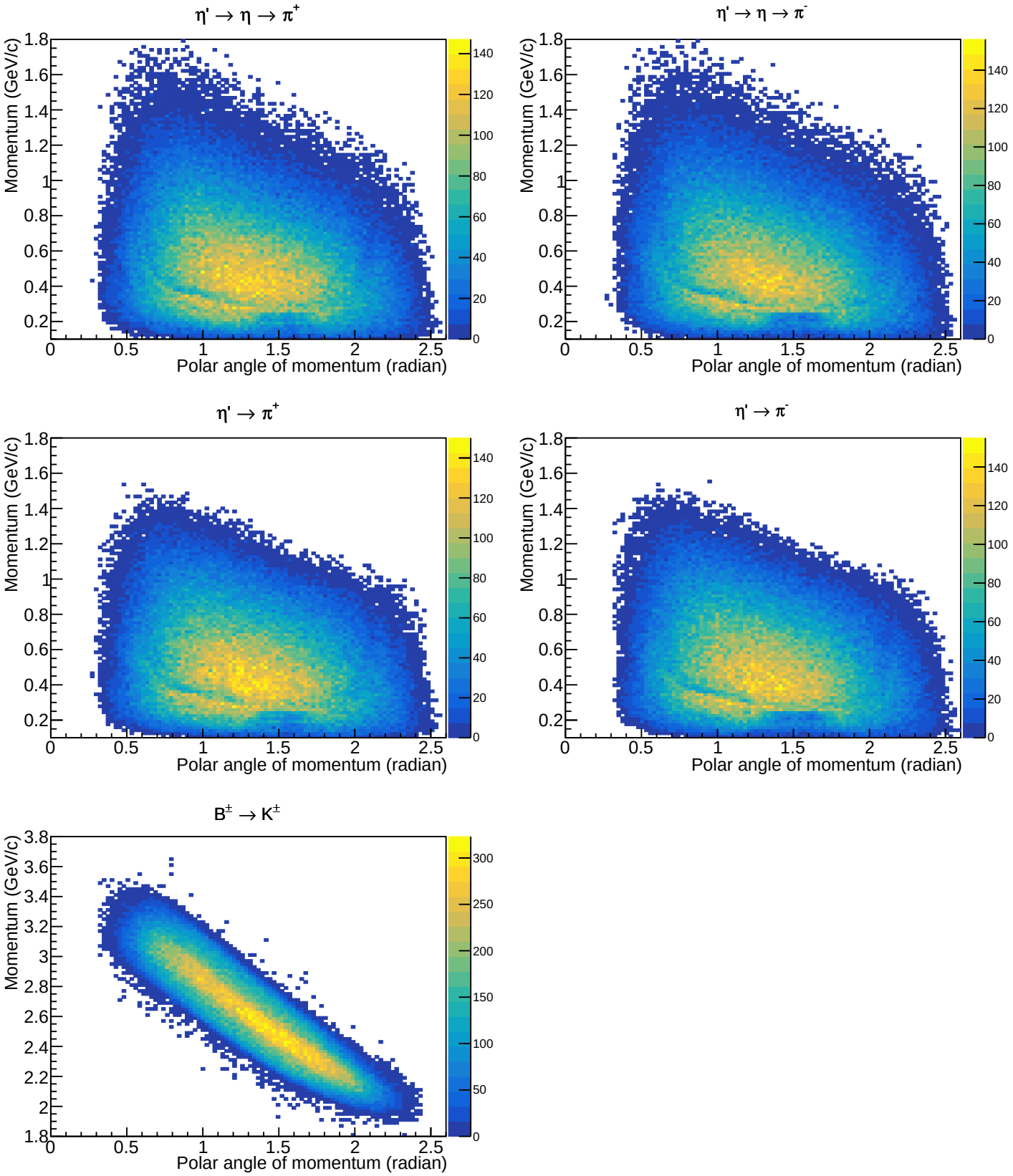


Figure D.4: 2D distribution of $p_{\pi,K}$ vs. $\theta_{\pi,K}$ for each charged particle. The decay chain in the plot's title indicates which particle the plot is for and the parents of the particle. Note that the direction of the Y-axis of this plot is inverse of that in Figures D.2 and D.3.

Appendix E

Optimization for Selection Criteria of π^0 and K_S^0

The momentum distribution of π^0 and invariant mass distribution of K_S^0 differs from signal and background π^0 candidates. (Figures E.1 and E.2). `KsSelector`[47], [45] is a FastBDT discriminator to select true K_S^0 and to reject neutral particles such as Λ from IP or around IP. We devised a method to get optimized selection criteria for these physical quantities and will elaborate on it in this subsection.

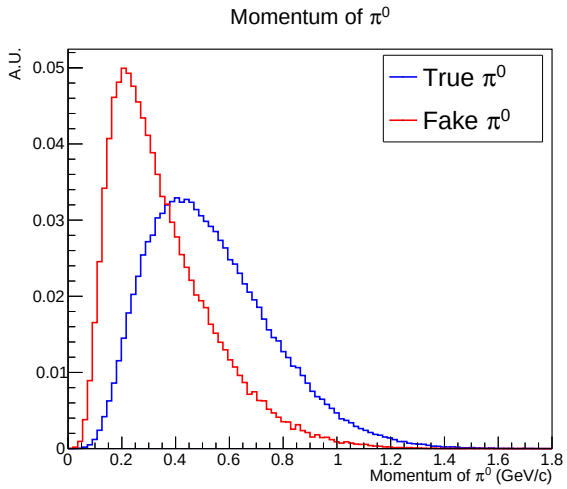


Figure E.1: Distribution of momentum from true and fake π^0

We calculated the Figure of Merit (FoM; $S/\sqrt{(S+B)}$) in the signal region given in Eq. 4.12. A special treatment was used for the optimization to deal with limited statistics of the number of events in the area. We assumed that the ratio $\alpha^{\mathrm{Sig}} = N_{FR}^{\mathrm{Sig}}/N_{SR}^{\mathrm{Sig}}$ is consistent over signal and generic MC samples. N_{FR}^{Sig} is the number of signal B^0 events that can pass the additional selection for π^0 and K_S^0 in the entire fitting region for signal extraction and N_{SR}^{Sig} is that in the signal region. We also assumed that the efficiency

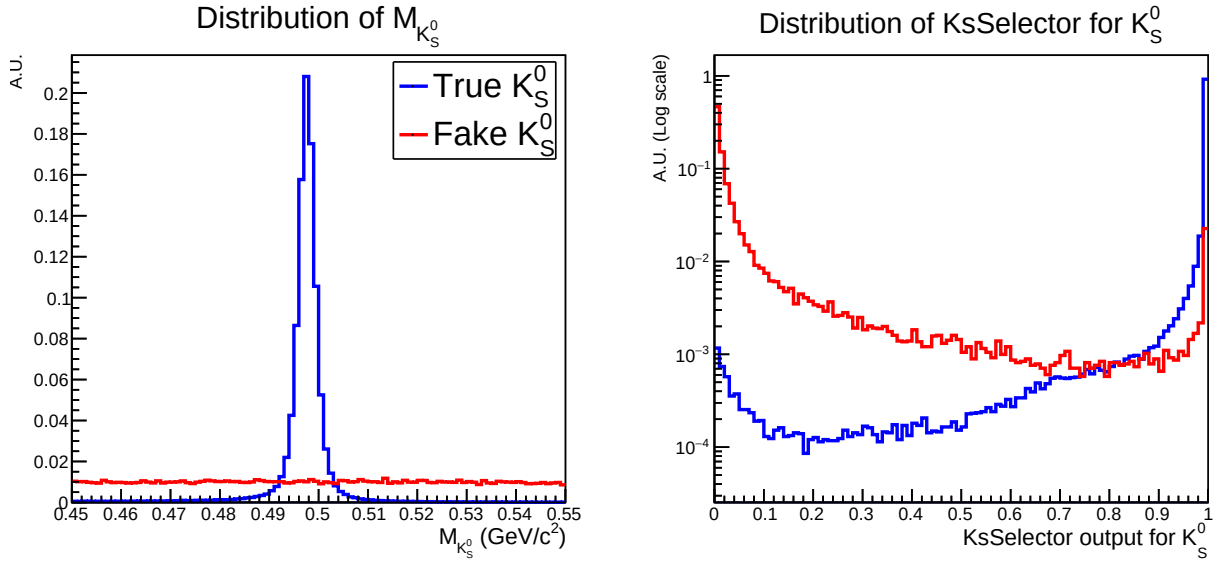


Figure E.2: Distribution of invariant mass and output of KsSelector from true and fake K^0_S candidates.

$\epsilon^{\text{Bckg}} = N^{\text{PassedBckg}}/N^{\text{AllBckg}}$ in the fitting region and the signal region is consistent, where $N^{\text{PassedBckg}}$ is the number of background B^0 events that can pass the additional criteria and N^{AllBckg} is all reconstructed background events. With these treatments, we can calculate the expected number of signal events (N_{Sig}) and background events (N_{Bckg}) in the signal regions with the selections by Eqs. E.1 and E.2 and minimize the statistical fluctuation so that we can get a reliable FoM curve from the discrete variable with limited statistics.

$$N_{\text{Sig}} = \alpha^{\text{Sig}} \cdot N'^{Sig}_{FR}, \tag{E.1}$$

$$N_{\text{Bckg}} = \epsilon^{\text{Bckg}} \cdot N'^{Bckg}_{SR}, \tag{E.2}$$

FoM maximizes by minimizing reciprocal FoM (1/FoM) with the additional criteria by the Genetic minimizer from TMVA[61]. The additional selection criteria consist of the lower limit of π^0 momentum (p_{π^0}) and `KsSelector`, and invariant mass window for K^0_S (lower limit and the width of the window for $M_{K^0_S}$) Floating parameters of the minimization are these four parameters, and optimization results are shown in Tab. E.1 and scan result is shown in Fig. E.3. We adjusted the lower limit of `KsSelector` from the optimized value to harmonize the selection criteria with analysis on the other subchannels for $B^0 \to \eta' K^0_S$. Also, we confirmed that the optimized thresholds for the selection with the initial version of our analysis differ from the optimization result with the latest dataset. However, it is okay to use selection criteria from the initial version because the difference in FoM is negligible. The value with old thresholds is 11.09, and that with optimized ones is 11.13.

Parameters for the selection	Optimized value	Used value	Note
Lower limit of p_{π^0}	0.144 GeV	0.204 GeV	$p_{\pi^0} > 0.204$ GeV
Lower limit of `KsSelector`	0.712	0.8	Adjusted to `KsSelector > 0.8`
Lower limit of $M_{K^0_S}$ window	0.482 GeV/c^2	0.482 GeV/c^2	
Width of $M_{K^0_S}$ window	0.028 GeV/c^2	0.029 GeV/c^2	$0.482 < M_{K^0_S} < 0.511$ GeV/c^2

Table E.1: List of optimized selection criteria.

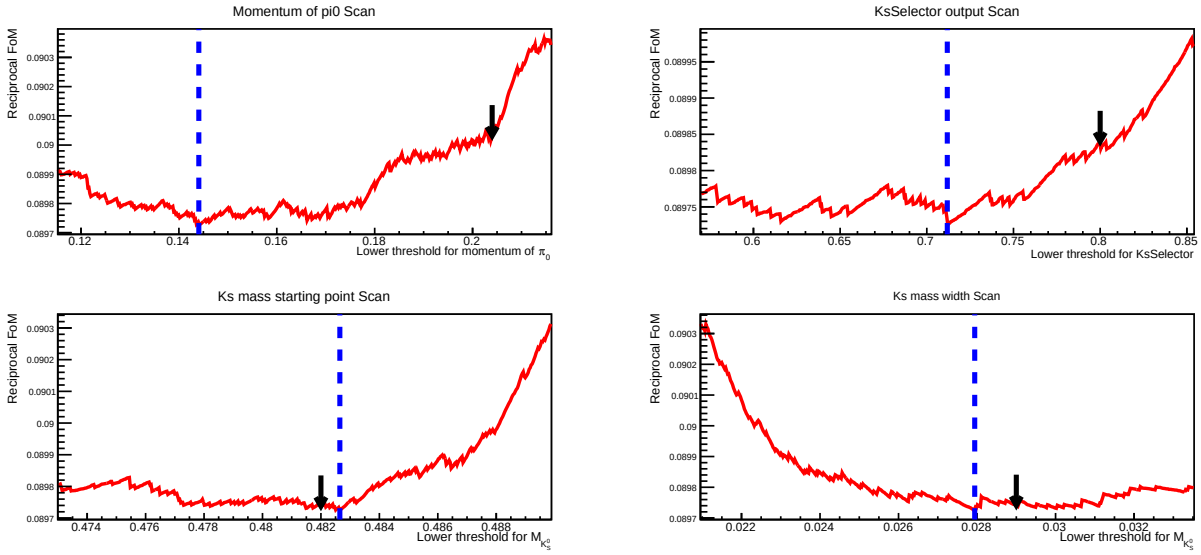


Figure E.3: Scan result for optimization of π^0 and K^0_S criteria. Blue dashed lines indicate optimized values, and black arrows show the current threshold of selection criteria.

After this optimization, we improved FoM from 6.63 to 11.09, and signal efficiency is 7.66% within the signal region. Only signal events were used for this calculation.

Appendix F

Comparison for Variables of the Vertex Quality

We compared vertex quality variables by two aspects: generic MC vs. experimental data and (Signal + SxF) vs. Background category. Detailed descriptions with plots will be provided in the following subsections.

F.1 Comparison between generic MC and experimental data using *sPlot*

We compared the distribution of variables between MC and Data that indicates the quality of the vertex fitting on that *CP* and tag side. Each component of (Signal + SxF) and Background category are separated using *sPlot* technique from results from the signal extraction for MC and Data.

Comparison of a component of the Background category shows differences, and we consider it is due to the imperfect simulation. Since we use parameters for background Δt PDF using sideband events from the experimental data, such differences are properly considered in the *CP* fitting procedure.

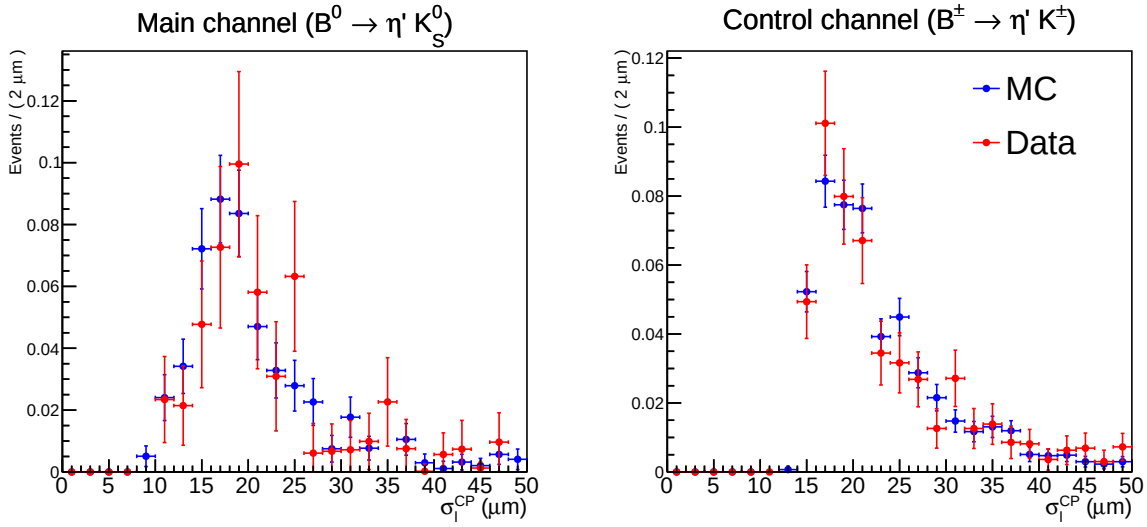


Figure F.1: Comparison of σ_ℓ^{CP} between MC and Data. (Signal + SxF) components are drawn.

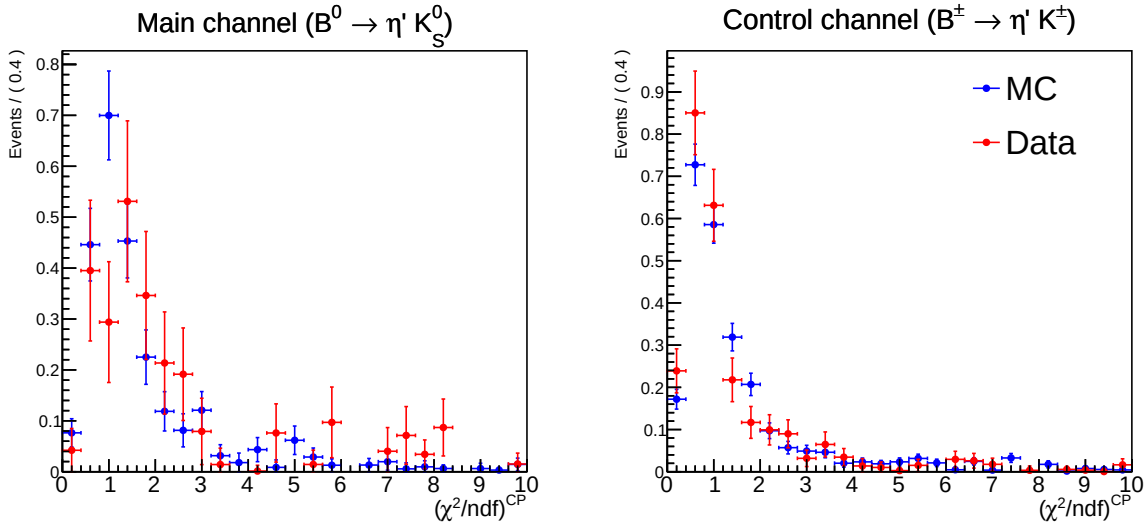


Figure F.2: Comparison of $(\chi^2/ndf)^{CP}$ between MC and Data. (Signal + SxF) components are drawn.

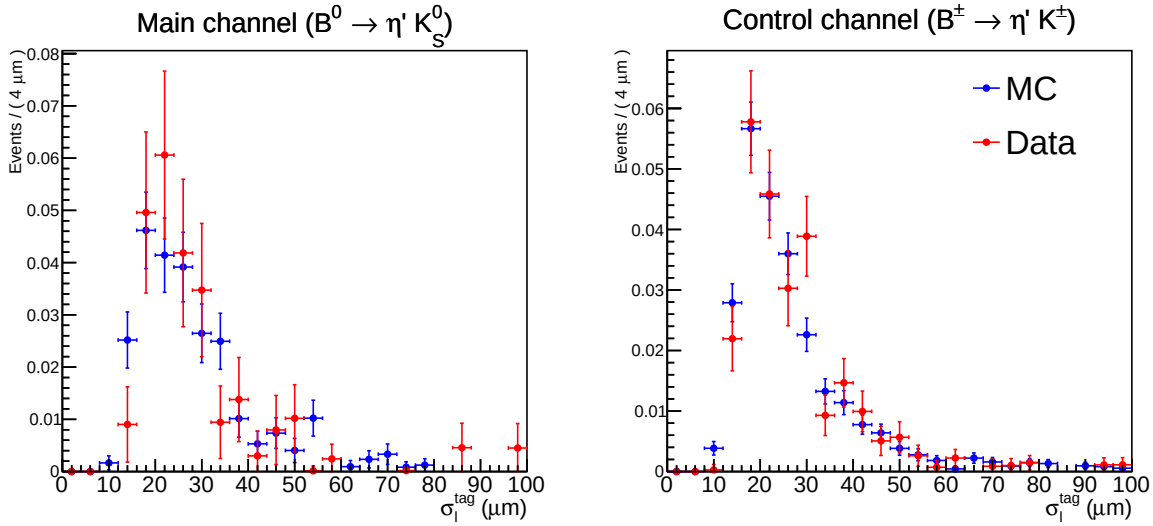


Figure F.3: Comparison of σ_ℓ^{tag} between MC and Data. (Signal + SxF) components are drawn.

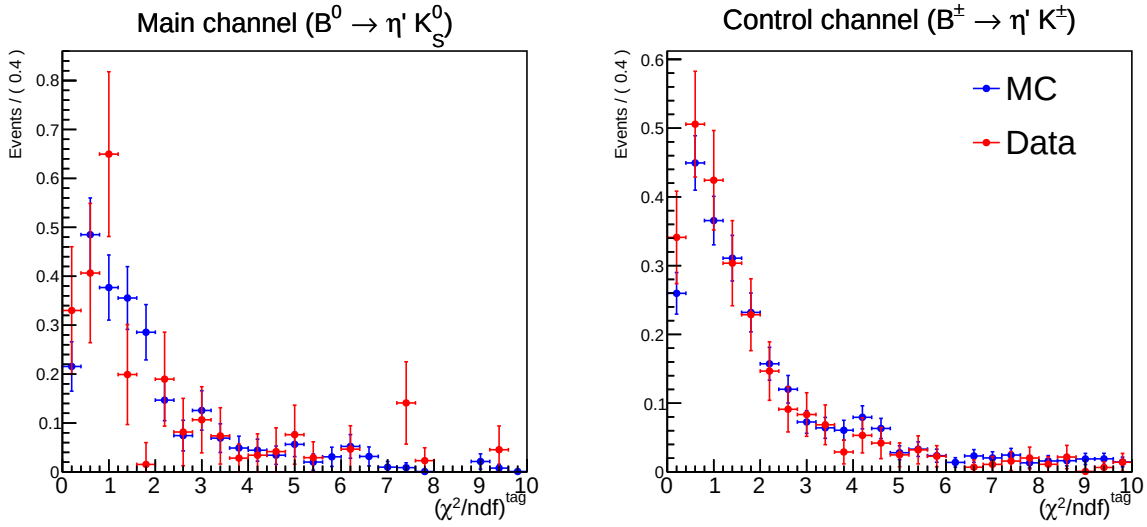


Figure F.4: Comparison of $(\chi^2/ndf)^{tag}$ between MC and Data. (Signal + SxF) components are drawn.

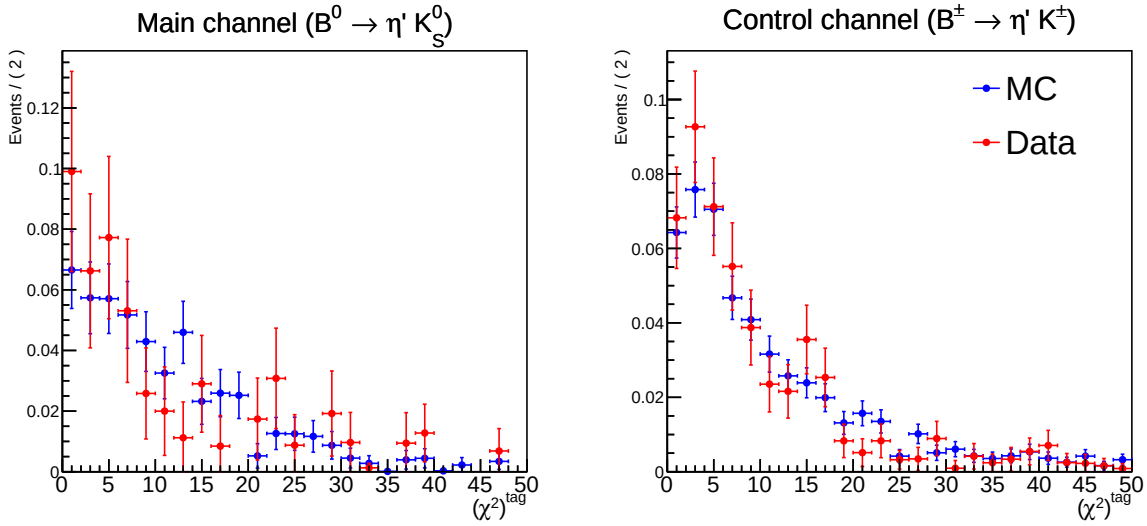


Figure F.5: Comparison of $(\chi^2)^{tag}$ between MC and Data. (Signal + SxF) components are drawn.

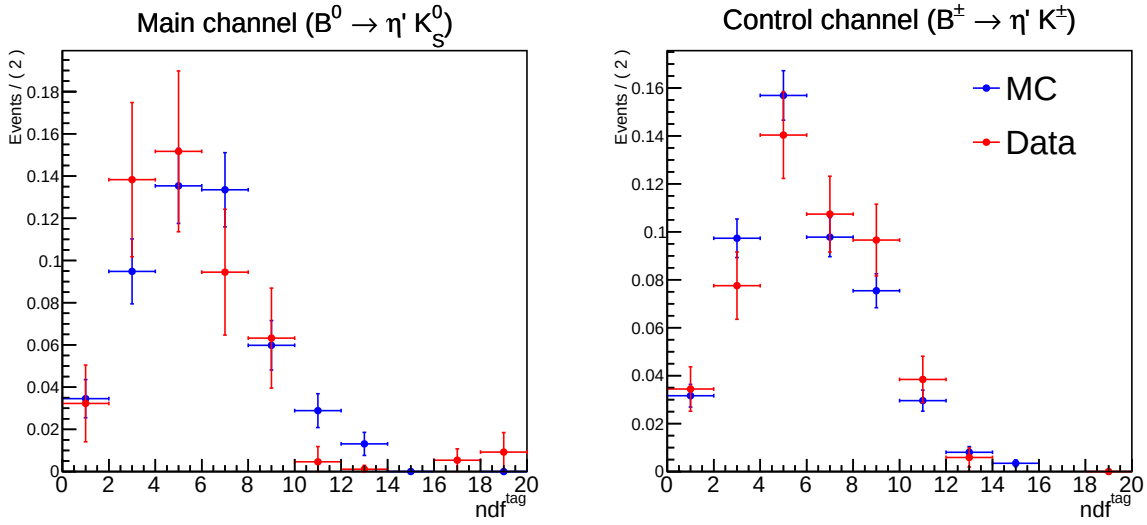


Figure F.6: Comparison of $(ndf)^{tag}$ between MC and Data. (Signal + SxF) components are drawn.

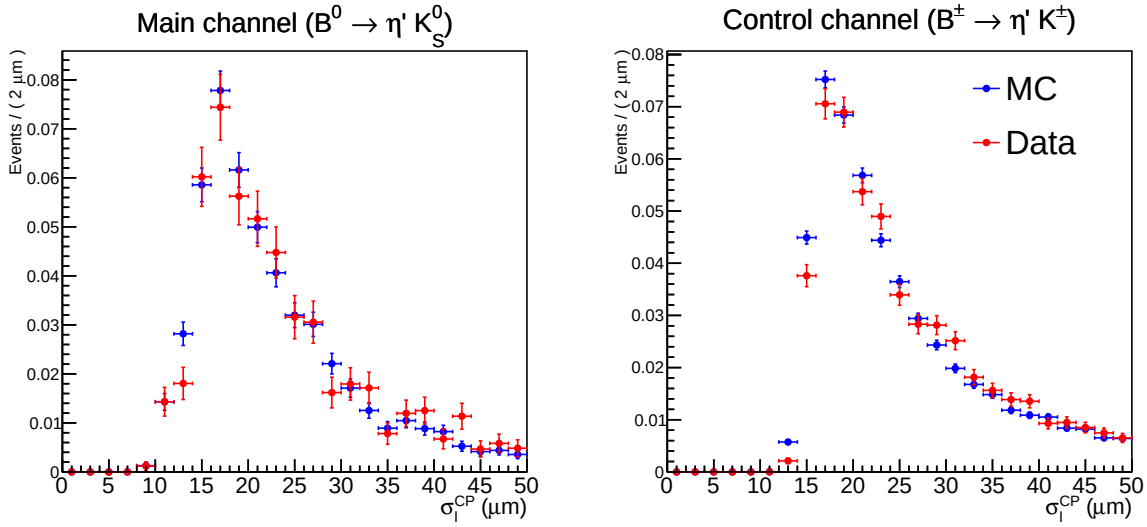


Figure F.7: Comparison of σ_ℓ^{CP} between MC and Data. Background components are drawn.

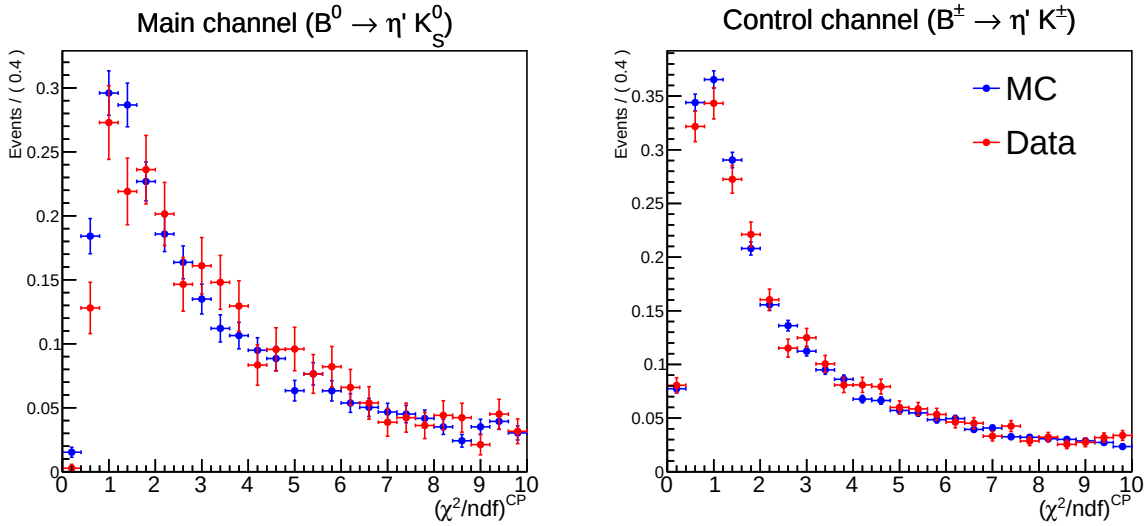


Figure F.8: Comparison of $(\chi^2/ndf)^{CP}$ between MC and Data. Background components are drawn.

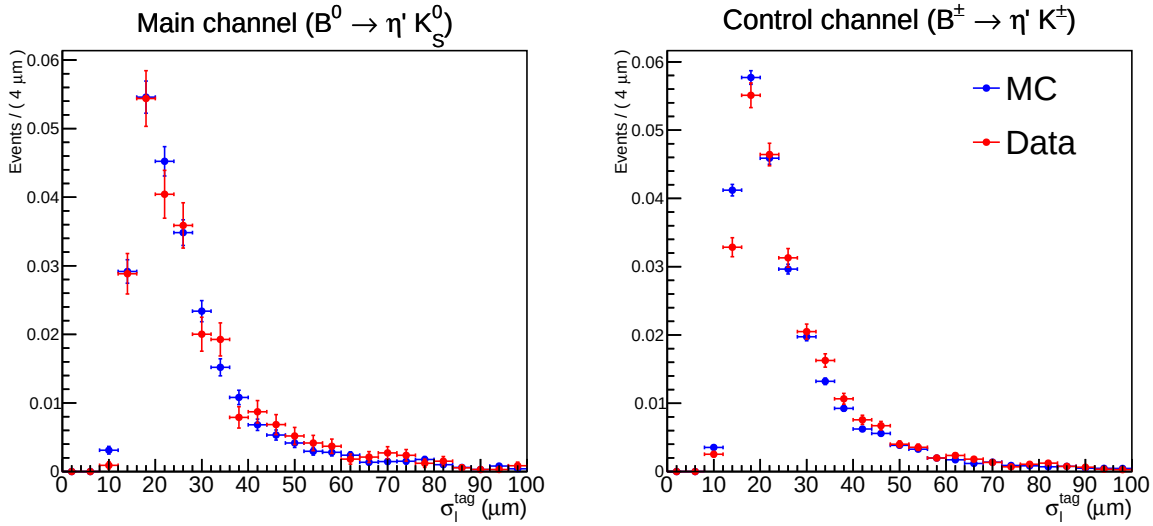


Figure F.9: Comparison of σ_ℓ^{tag} between MC and Data. Background components are drawn.

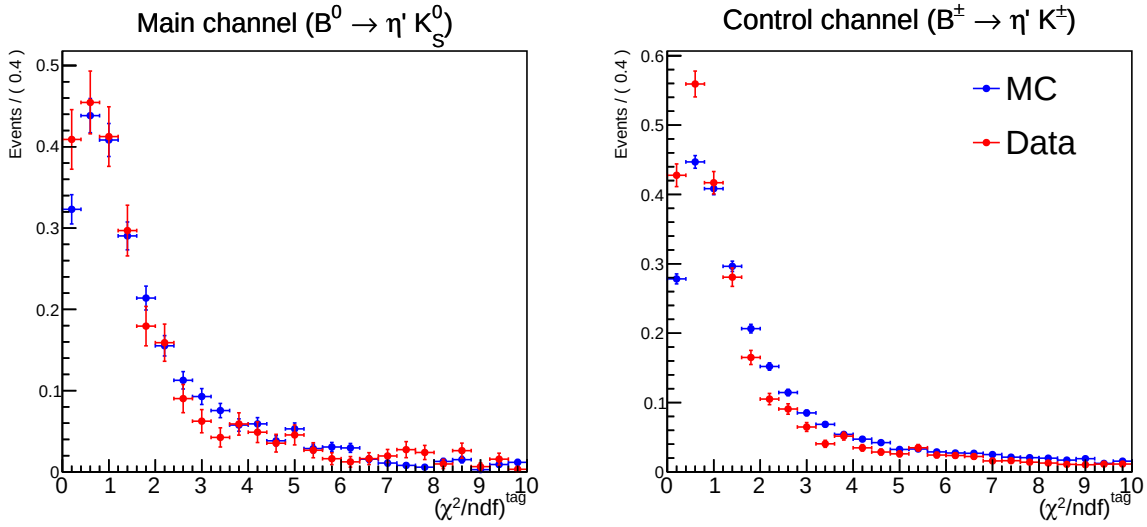


Figure F.10: Comparison of $(\chi^2/ndf)^{tag}$ between MC and Data. Background components are drawn.

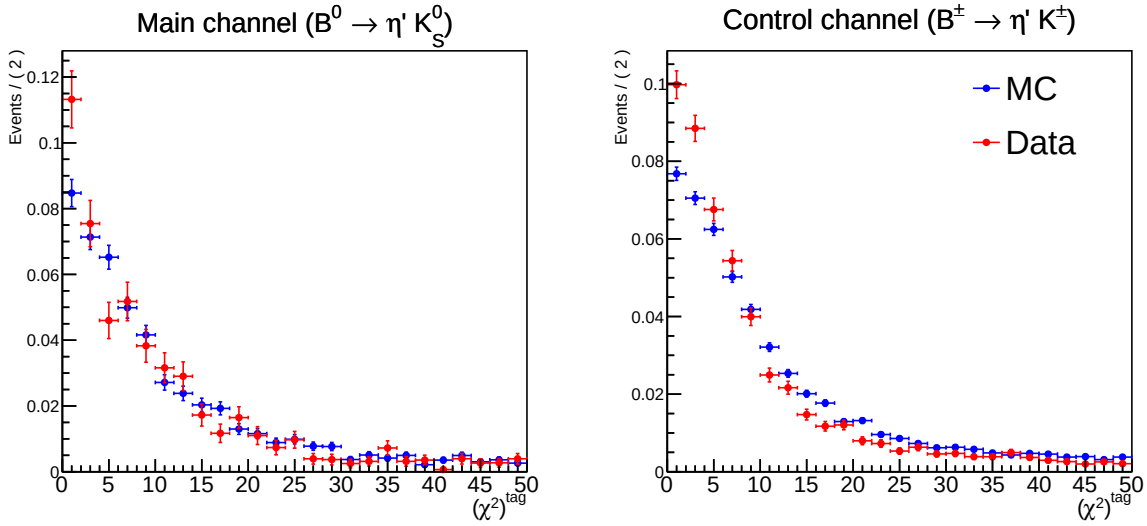


Figure F.11: Comparison of $(\chi^2)^{tag}$ between MC and Data. Background components are drawn.

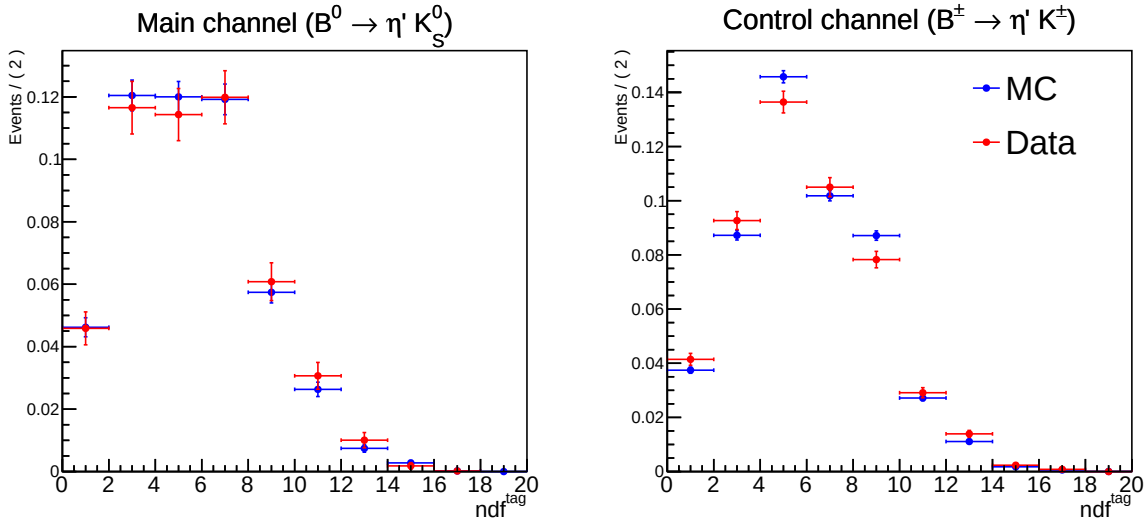


Figure F.12: Comparison of $(ndf)^{tag}$ between MC and Data. Background components are drawn.

F.2 Comparison between two categories in the MC samples

The conditional variables need to be examined to determine whether their shapes are similar between all categories for f_{sig} calculation and Δt PDF. We compared distributions of the variables between (Signal + SxF) and Background categories using MC-truth information. There are minor discrepancies in $\sigma_\ell^{CP,tag}$ between the two categories, but the reduced χ^2 both for CP and tag side shows significant differences. In our CP fitting procedure, we neglected such differences and did not include the Punzi term in the f_{sig} calculation, but we will consider the effect of the ignorance in the Fit Bias systematic study.

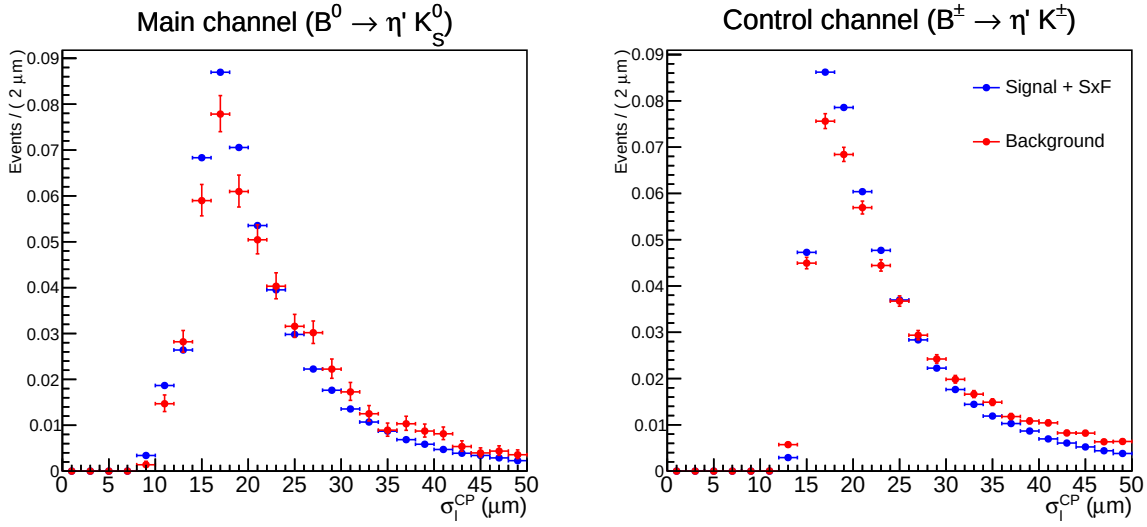


Figure F.13: Comparison of σ_ℓ^{CP} between (Signal + SxF) and Background events.

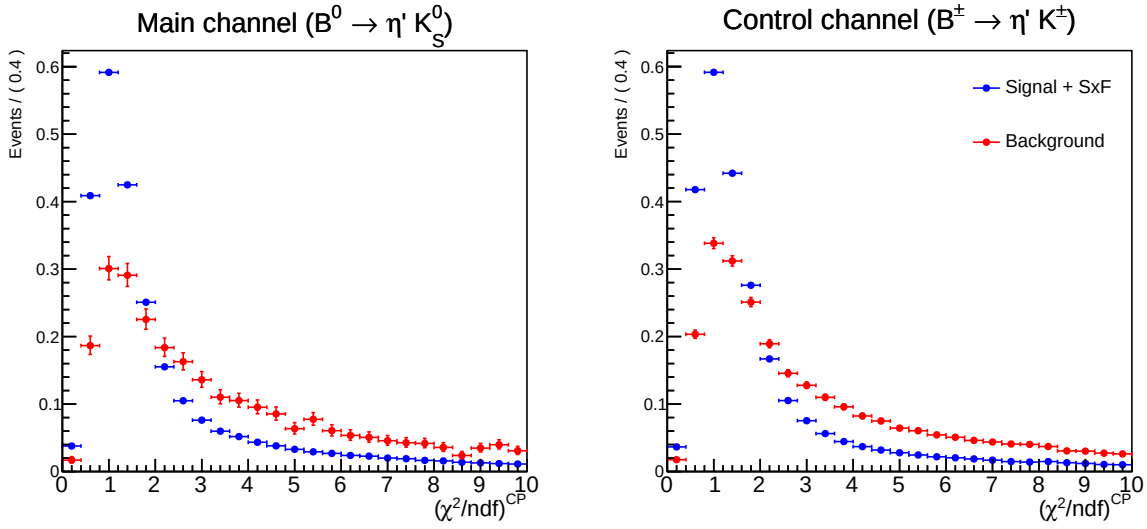


Figure F.14: Comparison of $(\chi^2/ndf)^{CP}$ between (Signal + SxF) and Background events.

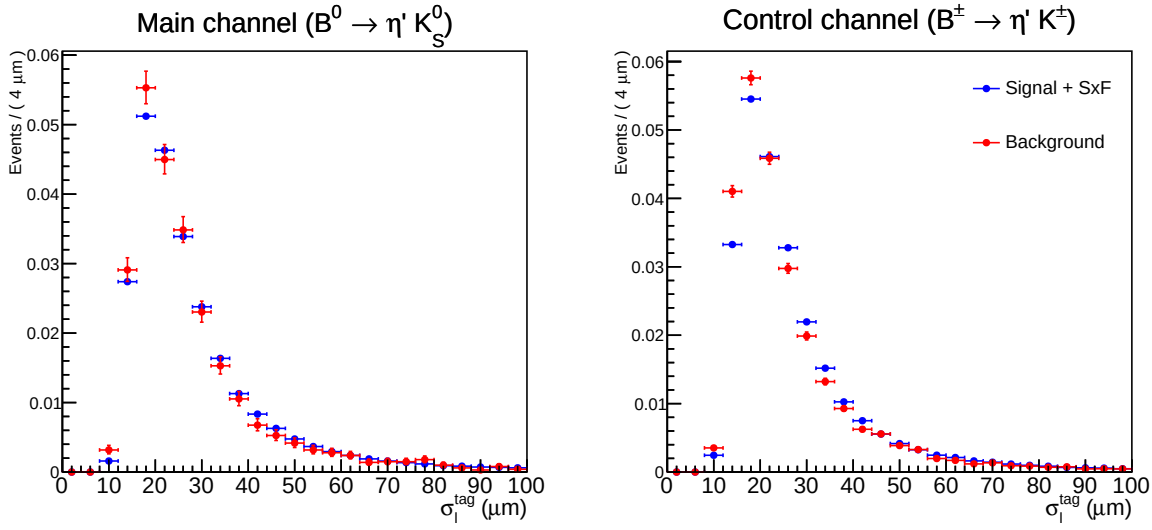


Figure F.15: Comparison of σ_ℓ^{tag} between (Signal + SxF) and Background events.

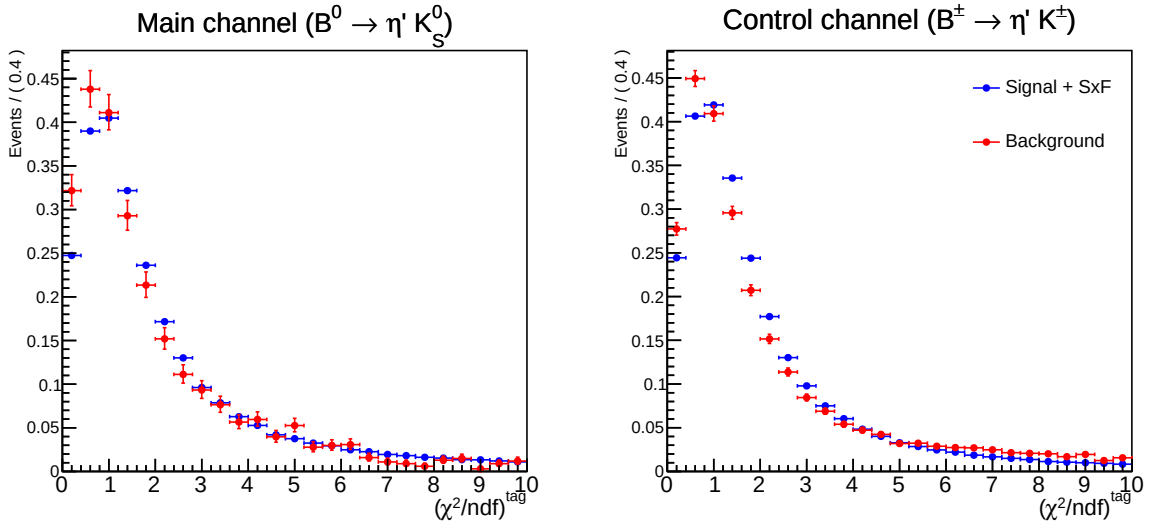


Figure F.16: Comparison of $(\chi^2/ndf)^{tag}$ between (Signal + SxF) and Background events.

Appendix G

The Confidence Region using Feldman-Cousins Approach

We estimate the confidence region of the *CP* fitting using the Feldman-Cousins approach[58] since we confirmed a bias and underestimation of uncertainty on A_{CP}^{fit} and S_{CP}^{fit} from the ensemble test for the *CP* fitting, which is described in Sec. 6.2.4. Determination of the confidence region using the Feldman-Cousins method requires the precise modeling of response from our *CP* fitter by the various input for *CP* asymmetries of A_{CP}^{input} and S_{CP}^{input}. In this appendix, we discuss the details of the modeling, and evaluation of the Feldman-Cousins confidence level.

We perform ensemble tests at 317 points of $(A_{CP}^{input}, S_{CP}^{input})$ generated in a mesh shape with a spacing of 0.1 within the physical region $(\sqrt{(A_{CP}^{input})^2 + (S_{CP}^{input})^2} \leq 1)$, which is given in Fig. G.1. We generate 10000 Toy MC datasets for each input point with N_{sig} and N_{bckg} calculate using F_{sig} and N_{total} given in Tab. 4.7.

We fitted the distribution of $(A_{CP}^{fit}, S_{CP}^{fit})$ with conditional variables of A_{CP}^{input} and S_{CP}^{input} to a response function of our *CP* fitter that consists of a summation of three 2D-Gaussian functions, which is given by the following equation:

$$\begin{aligned}
&P_{FC}(A_{CP}^{fit}, S_{CP}^{fit} | A_{CP}^{input}, S_{CP}^{input}) \\
&\quad = f_{major}\mathrm{Gaus}(A_{CP}^{fit} | \mu_{A,\,major}, \sigma_{A,\,major})\mathrm{Gaus}(S_{CP}^{fit} | \mu_{S,\,major}, \sigma_{S,\,major}) \\
&\qquad + f_{minor}(1 - f_{major})\mathrm{Gaus}(A_{CP}^{fit} | \mu_{A,\,minor}, \sigma_{A,\,minor})\mathrm{Gaus}(S_{CP}^{fit} | \mu_{S,\,minor}, \sigma_{S,\,minor}) \\
&\qquad + (1 - f_{minor})(1 - f_{major})\mathrm{Gaus}(A_{CP}^{fit} | \mu_{A,\,tail}, \sigma_{A,\,tail})\mathrm{Gaus}(S_{CP}^{fit} | \mu_{S,\,tail}, \sigma_{S,\,tail})
\end{aligned} \tag{G.1}$$

where Gaus() is a Gaussian function given in Eq. 4.16, and μ of σ are defined as

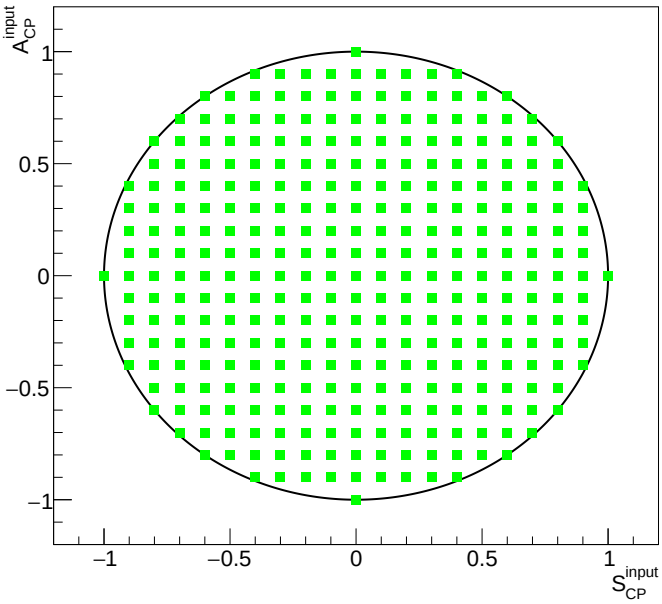


Figure G.1: Input points of $(A_{CP}^{input}, S_{CP}^{input})$ for the ensemble tests and a boundary (A black solid line) for the physical region.

$$
\begin{aligned}
f_{major} &= \max(0, \min(1, f_0^{major} + f_2^{A,\,major}(A_{CP}^{input})^2 + f_2^{S,\,major}(S_{CP}^{input})^2)), && \text{(G.2)}\\
f_{minor} &= \max(0, \min(1, f_0^{minor} + f_2^{A,\,minor}(A_{CP}^{input})^2 + f_2^{S,\,minor}(S_{CP}^{input})^2)), && \text{(G.3)}\\
\sigma_{A,major} &= \sigma_0^{A,\,major} + \sigma_2^{A,\,major}(A_{CP}^{input})^2 + \sigma_4^{A,\,major}(A_{CP}^{input})^4, && \text{(G.4)}\\
\sigma_{A,\,minor} &= \sigma^{A,\,minor}, && \text{(G.5)}\\
\sigma_{A,\,tail} &= \sigma^{A,\,tail}, && \text{(G.6)}\\
\sigma_{S,\,major} &= \sigma_0^{S,\,major} + \sigma_2^{S,\,major}(S_{CP}^{input})^2 + \sigma_2^{corrA,\,major}(A_{CP}^{input})^2, && \text{(G.7)}\\
\sigma_{S,\,minor} &= \sigma^{S,\,minor}, && \text{(G.8)}\\
\sigma_{S,\,tail} &= \sigma^{S,\,tail}, && \text{(G.9)}\\
\mu_{A,\,major} &= \mu_1^{A,\,major} A_{CP}^{input} + \mu_3^{A,\,major}(A_{CP}^{input})^3 + \mu_5^{A,\,major}(A_{CP}^{input})^5, && \text{(G.10)}\\
\mu_{A,\,minor} &= \mu_1^{A,\,minor} A_{CP}^{input} + \mu_3^{A,\,minor}(A_{CP}^{input})^3, && \text{(G.11)}\\
\mu_{A,\,tail} &= \mu_1^{A,\,tail} A_{CP}^{input}, && \text{(G.12)}\\
\mu_{S,\,major} &= \mu_1^{S,\,major} S_{CP}^{input} + \mu_3^{S,\,major}(S_{CP}^{input})^3, && \text{(G.13)}\\
\mu_{S,\,minor} &= \mu_1^{S,\,minor} S_{CP}^{input}, \text{ and} && \text{(G.14)}\\
\mu_{S,\,tail} &= \mu_1^{S,\,tail} S_{CP}^{input}. && \text{(G.15)}
\end{aligned}
$$

Thus, we have 26 parameters of the PDF in total, and they are determined in the fitting using results from the ensemble tests, and the resultant parameters are given in Tab. G.1. Figures G.2, G.3, G.4, G.5, G.6, G.7, G.8, and G.9 show a projection on A_{CP}^{fit}

or S_{CP}^{fit} of fitting results for P_{FC} and the distribution of $(A_{CP}^{fit}, S_{CP}^{fit})$ from the ensemble tests.

Parameter	Determined value
f_0^{major}	0.84840 ± 0.00184
f_0^{minor}	0.96396 ± 0.00130
$f_2^{A,\,major}$	-0.22332 ± 0.00309
$\mu_1^{A,\,major}$	0.98911 ± 0.00176
$\mu_3^{A,\,major}$	-0.04509 ± 0.00757
$\mu_5^{A,\,major}$	0.04463 ± 0.00746
$\sigma_0^{A,\,major}$	0.31248 ± 0.00034
$\sigma_2^{A,\,major}$	-0.08325 ± 0.00210
$\sigma_4^{A,\,major}$	0.03117 ± 0.00273
$f_2^{A,\,minor}$	-0.04447 ± 0.00228
$\mu_1^{A,\,minor}$	1.20112 ± 0.00418
$\mu_3^{A,\,minor}$	0.13358 ± 0.00640
$\sigma^{A,\,minor}$	0.45638 ± 0.00100
$\mu_1^{A,\,tail}$	2.33638 ± 0.01313
$\sigma^{A,\,tail}$	0.85259 ± 0.00439
$\sigma_2^{corrA,\,major}$	-0.04965 ± 0.00148
$\sigma_0^{S,\,major}$	0.45750 ± 0.00069
$\sigma_2^{S,\,major}$	-0.03056 ± 0.00141
$f_2^{S,\,major}$	-0.13520 ± 0.00274
$\mu_1^{S,\,major}$	0.94874 ± 0.00168
$\mu_3^{S,\,major}$	0.03503 ± 0.00294
$f_2^{S,\,minor}$	-0.02881 ± 0.00231
$\mu_1^{S,\,minor}$	1.61890 ± 0.00500
$\sigma^{S,\,minor}$	0.80803 ± 0.00205
$\mu_1^{S,\,tail}$	3.30646 ± 0.02531
$\sigma^{S,\,tail}$	1.65648 ± 0.00865

Table G.1: List of parameters of the response function for the calculation of the Feldman-Cousins confidence level.

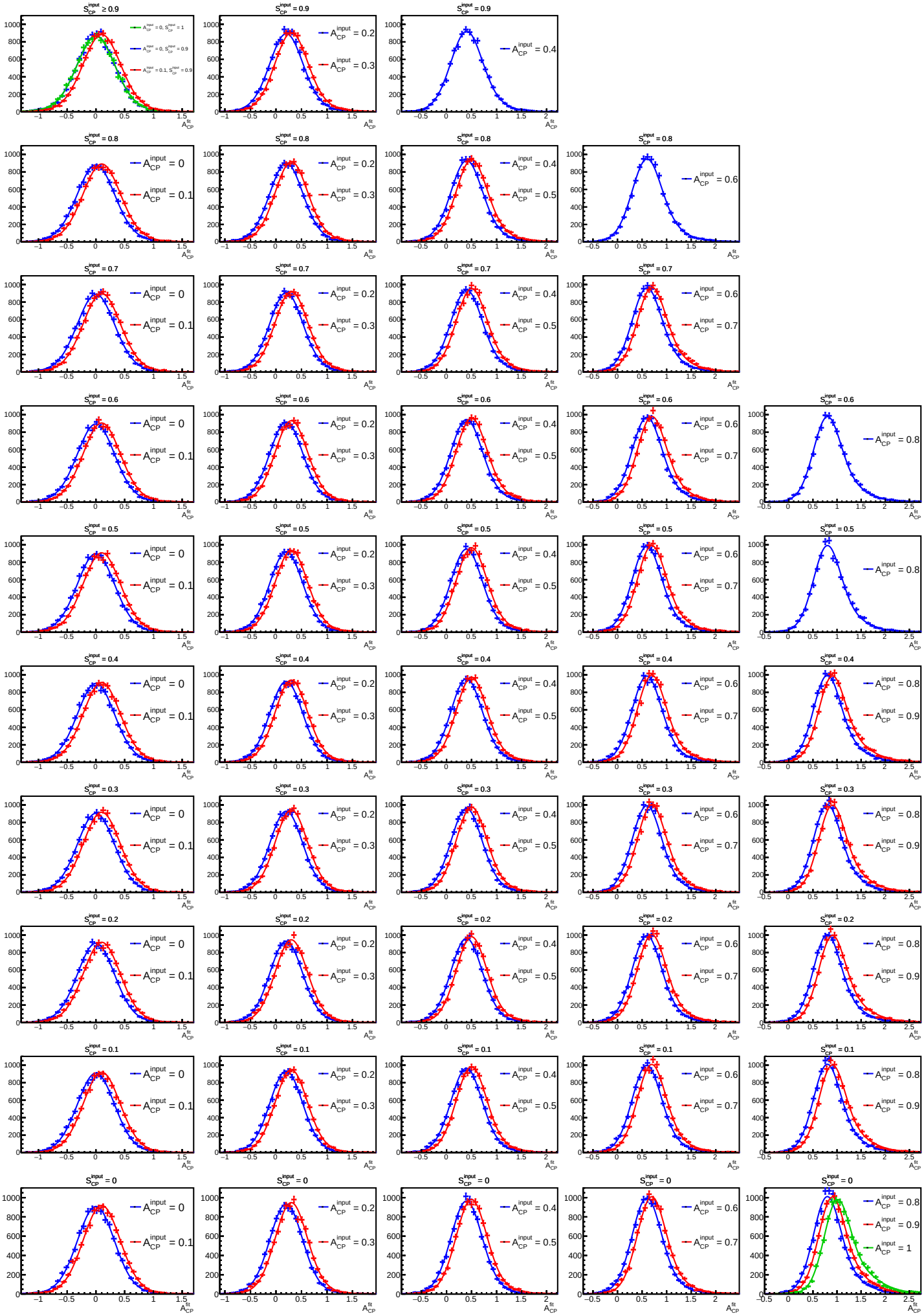


Figure G.2: Projection on A_{CP}^{fit} of the fitted P_{FC} and distribution of $(A_{CP}^{fit}, S_{CP}^{fit})$ from the various $(A_{CP}^{input}, S_{CP}^{input})$. Cases of $A_{CP}^{input} \geq 0$ and $S_{CP}^{input} \geq 0$ are drawn.

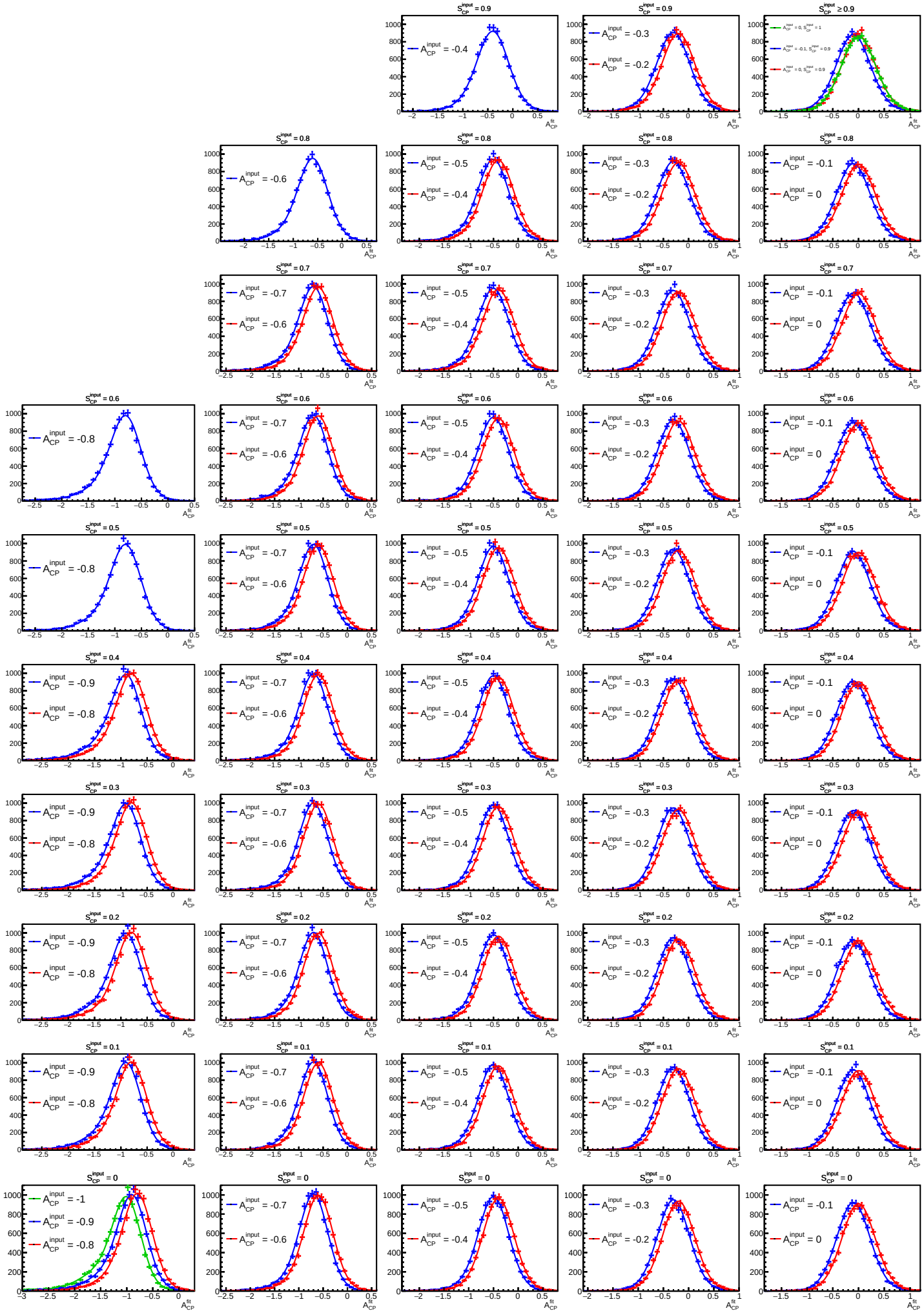


Figure G.3: Projection on A_{CP}^{fit} of the fitted P_{FC} and distribution of $(A_{CP}^{fit}, S_{CP}^{fit})$ from the various $(A_{CP}^{input}, S_{CP}^{input})$. Cases of $A_{CP}^{input} \leq 0$ and $S_{CP}^{input} \geq 0$ are drawn.

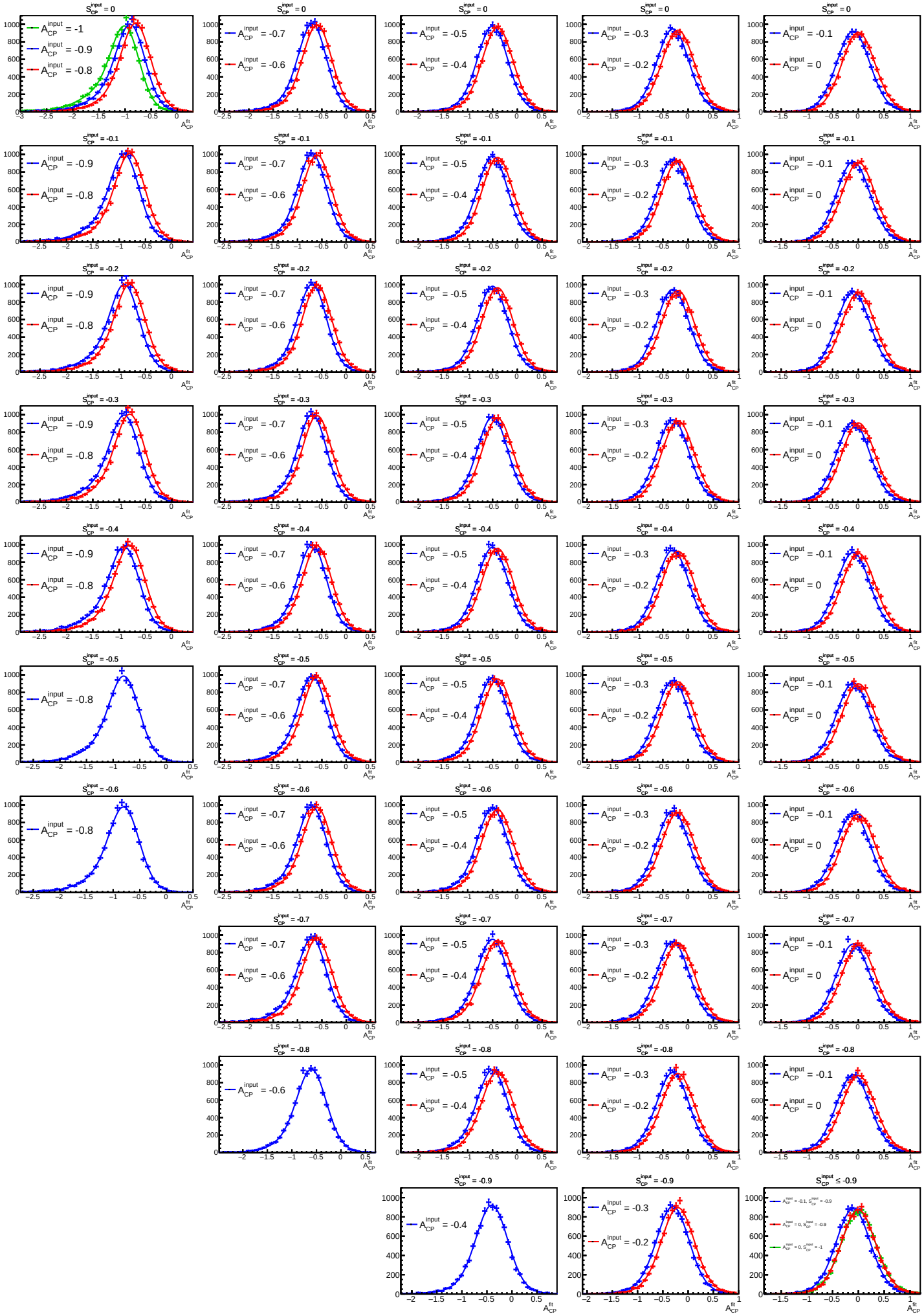


Figure G.4: Projection on A_{CP}^{fit} of the fitted P_{FC} and distribution of $(A_{CP}^{fit}, S_{CP}^{fit})$ from the various $(A_{CP}^{input}, S_{CP}^{input})$. Cases of $A_{CP}^{input} \leq 0$ and $S_{CP}^{input} \leq 0$ are drawn.

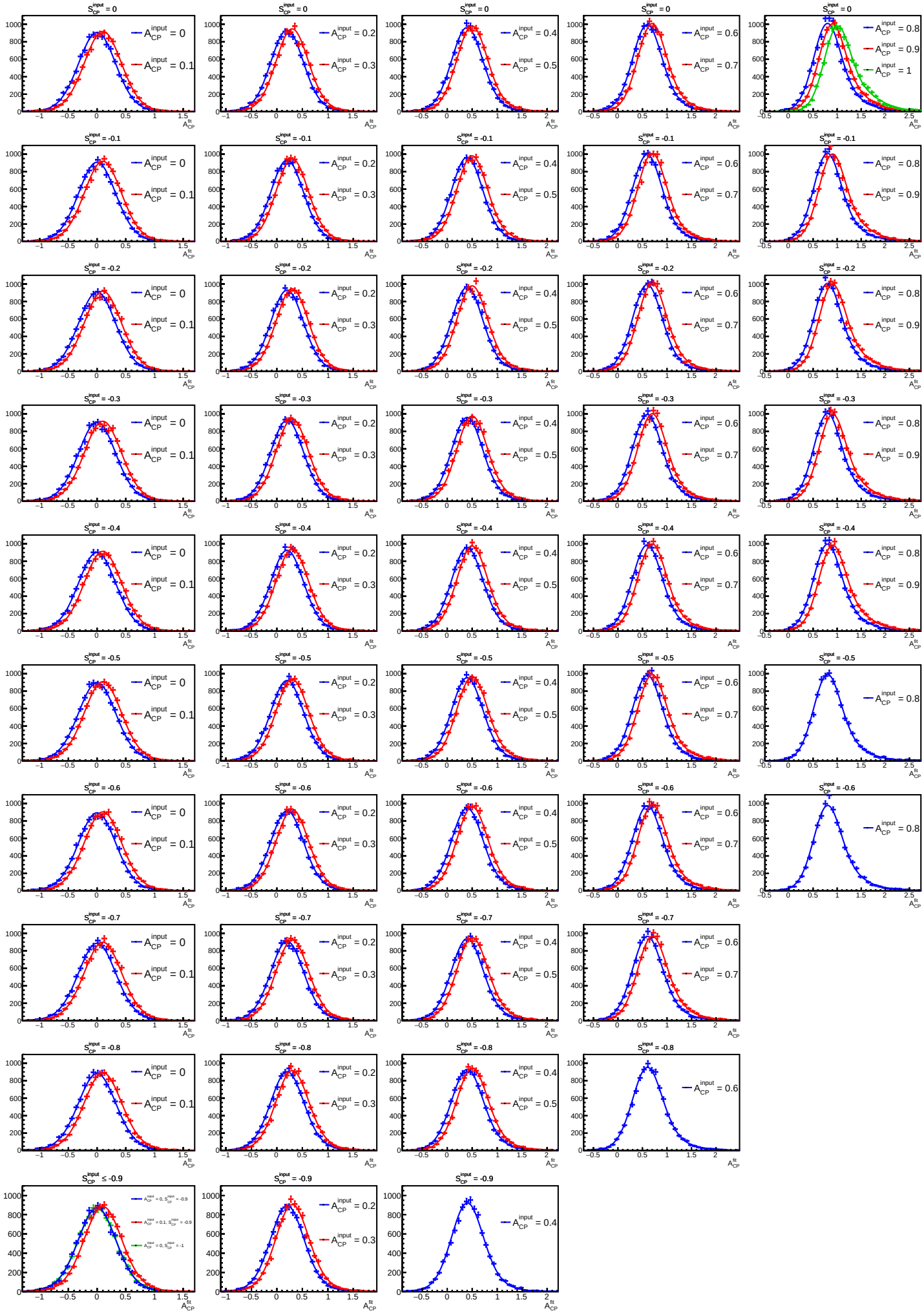


Figure G.5: Projection on A_{CP}^{fit} of the fitted P_{FC} and distribution of $(A_{CP}^{fit}, S_{CP}^{fit})$ from the various $(A_{CP}^{input}, S_{CP}^{input})$. Cases of $A_{CP}^{input} \geq 0$ and $S_{CP}^{input} \leq 0$ are drawn.

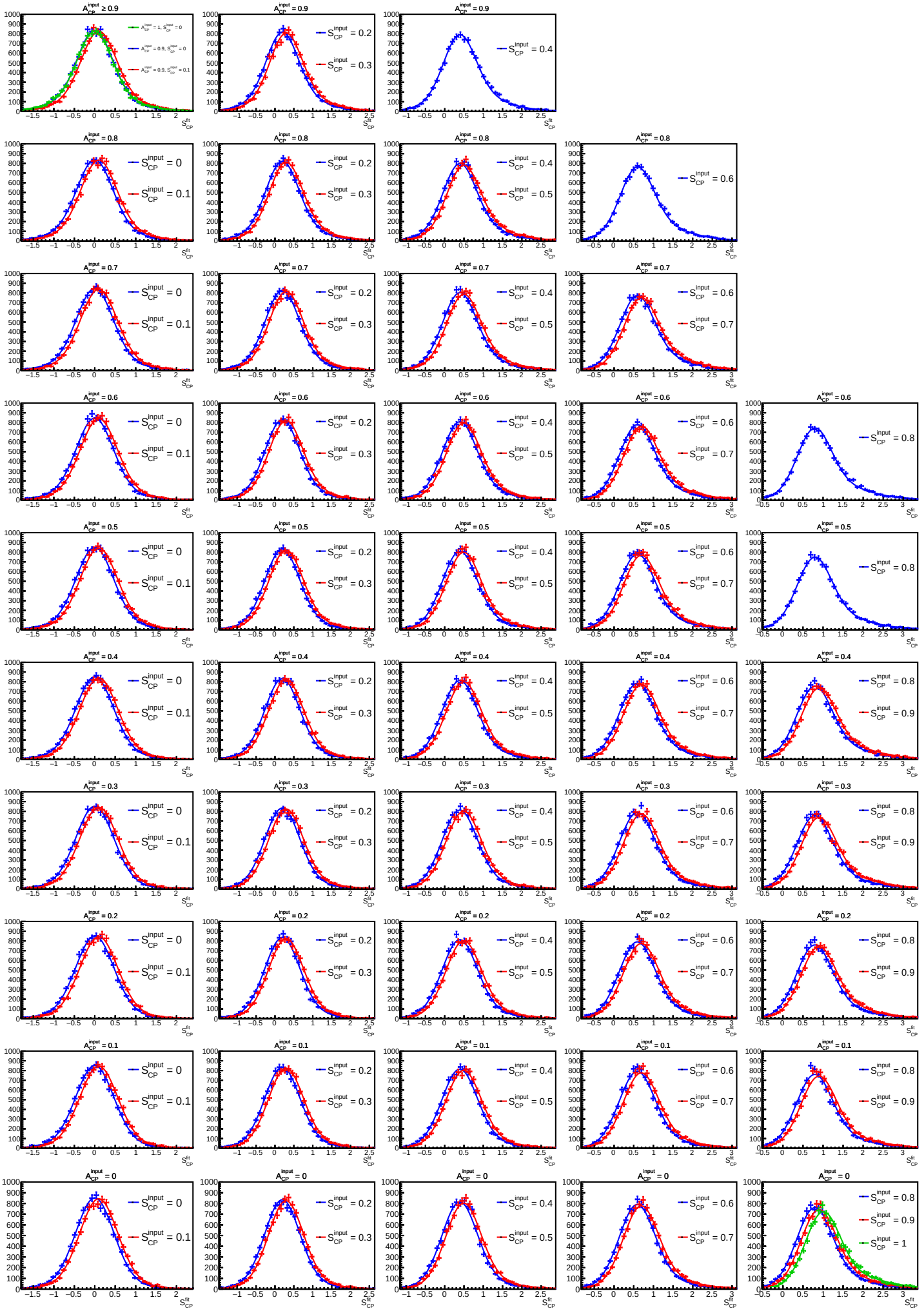


Figure G.6: Projection on S_{CP}^{fit} of the fitted P_{FC} and distribution of $(A_{CP}^{fit}, S_{CP}^{fit})$ from the various $(A_{CP}^{input}, S_{CP}^{input})$. Cases of $A_{CP}^{input} \geq 0$ and $S_{CP}^{input} \geq 0$ are drawn.

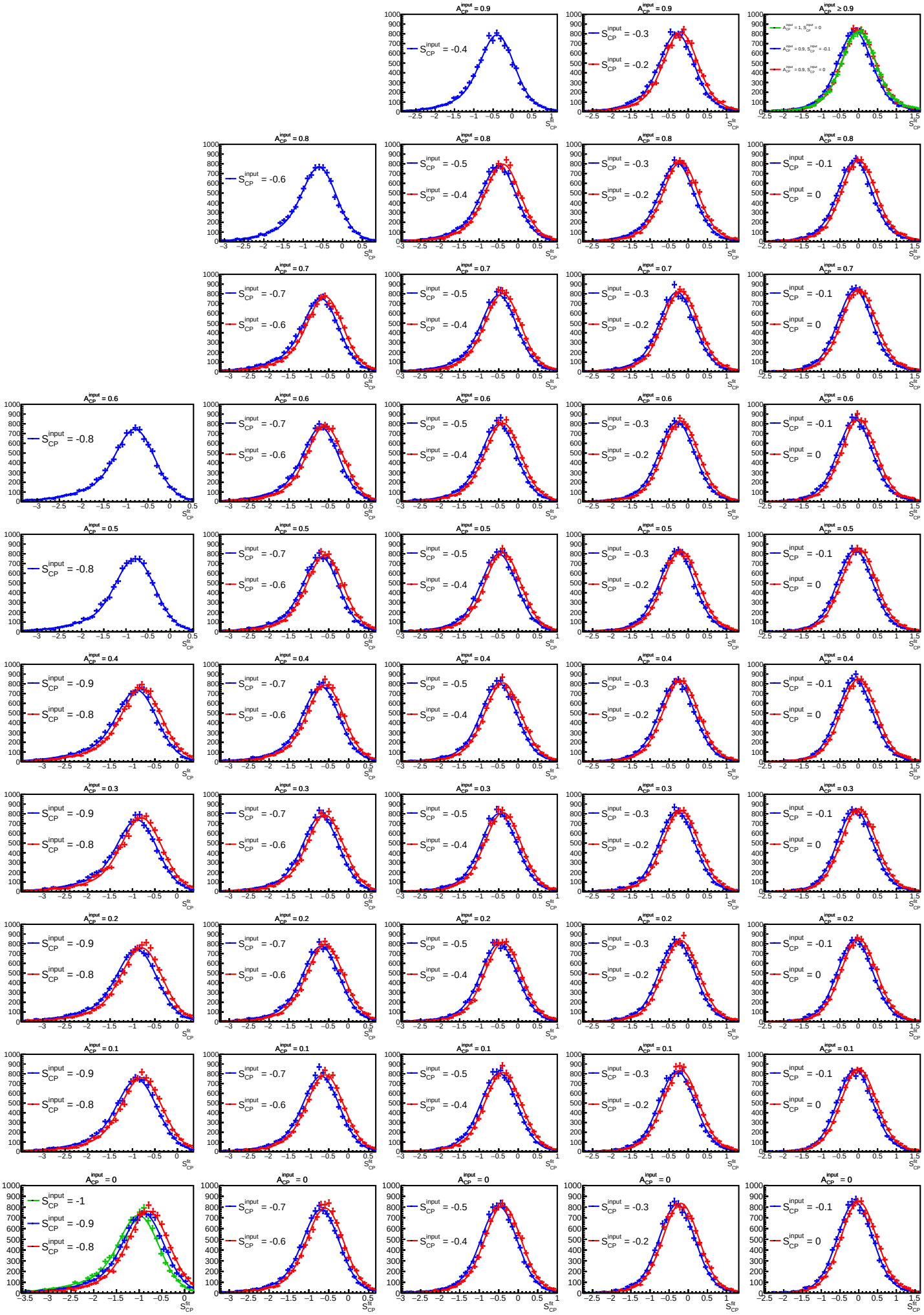


Figure G.7: Projection on S_{CP}^{fit} of the fitted P_{FC} and distribution of $(A_{CP}^{fit}, S_{CP}^{fit})$ from the various $(A_{CP}^{input}, S_{CP}^{input})$. Cases of $A_{CP}^{input} \leq 0$ and $S_{CP}^{input} \geq 0$ are drawn.

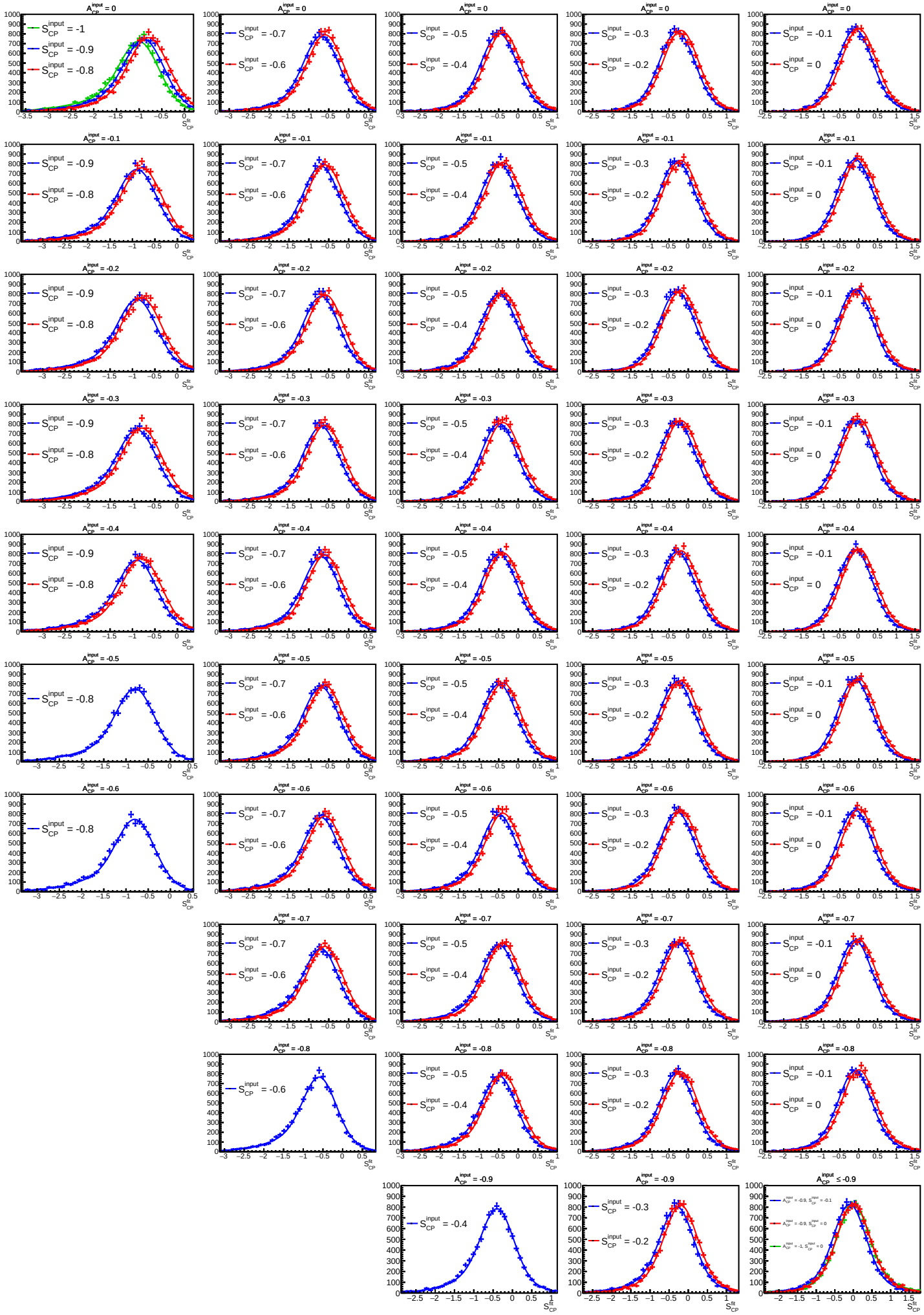


Figure G.8: Projection on S_{CP}^{fit} of the fitted P_{FC} and distribution of $(A_{CP}^{fit}, S_{CP}^{fit})$ from the various $(A_{CP}^{input}, S_{CP}^{input})$. Cases of $A_{CP}^{input} \leq 0$ and $S_{CP}^{input} \leq 0$ are drawn.

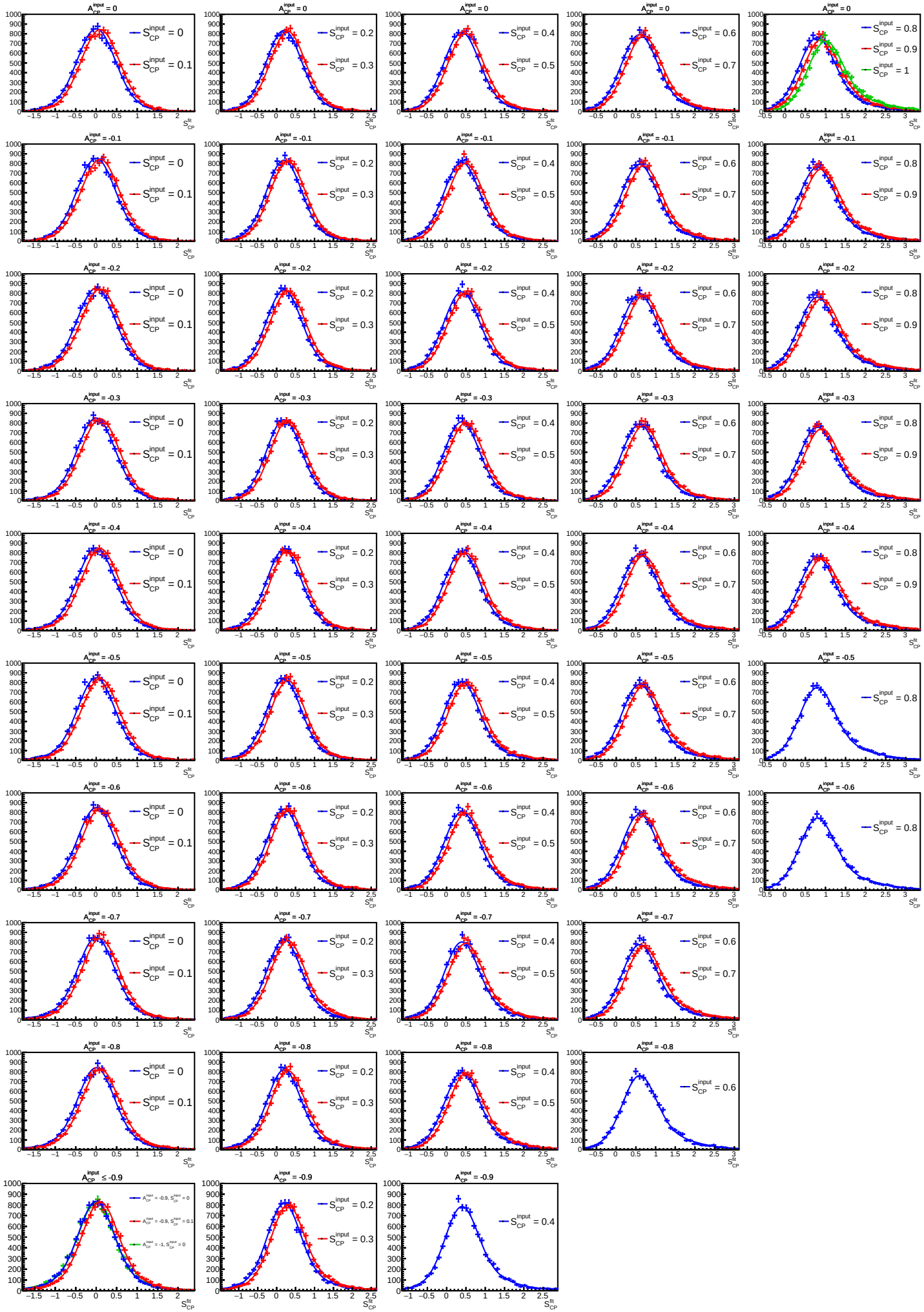


Figure G.9: Projection on S_{CP}^{fit} of the fitted P_{FC} and distribution of $(A_{CP}^{fit}, S_{CP}^{fit})$ from the various $(A_{CP}^{input}, S_{CP}^{input})$. Cases of $A_{CP}^{input} \geq 0$ and $S_{CP}^{input} \leq 0$ are drawn.

As the Feldman-Cousins method is a frequentist approach for the calculation of the confidence level, we should integrate the response function P_{FC} to calculate the confidence level to reject a specific set of true CP asymmetries of $(A_{CP}^{true}, S_{CP}^{true})$ when we measured the CP asymmetries by $(A_{CP}^{meas}, S_{CP}^{meas})$. The Feldman-Cousins method defines an integral region using the likelihood ratio. The likelihood ratio is defined as follows:

$$LR(A_{CP}^{fit}, S_{CP}^{fit}|A_{CP}^{true}, S_{CP}^{true}) = \frac{P_{FC}(A_{CP}^{fit}, S_{CP}^{fit}|A_{CP}^{true}, S_{CP}^{true})}{P_{FC}(A_{CP}^{fit}, S_{CP}^{fit}|A_{CP}^{best}, S_{CP}^{best})} \tag{G.16}$$

where $(A_{CP}^{best}, S_{CP}^{best})$ is a value of $(A_{CP}^{true}, S_{CP}^{true})$ which maximizes the $P_{FC}(A_{CP}^{fit}, S_{CP}^{fit}|A_{CP}^{true}, S_{CP}^{true})$ with a given $(A_{CP}^{fit}, S_{CP}^{fit})$ value.

We calculate a confidence level $CL(A_{CP}^{true}, S_{CP}^{true})$ to reject a set of $(A_{CP}^{true}, S_{CP}^{true})$ by integrating P_{FC} in a region R of $(A_{CP}^{fit}, S_{CP}^{fit})$ that satisfies $LR(A_{CP}^{meas}, S_{CP}^{meas}|A_{CP}^{true}, S_{CP}^{true}) \geq LR(A_{CP}^{fit}, S_{CP}^{fit}|A_{CP}^{true}, S_{CP}^{true})$ with a given result of the measurement $(A_{CP}^{meas}, S_{CP}^{meas})$ as follows:

$$CL(A_{CP}^{true}, S_{CP}^{true}) = \iint_R P(A_{CP}^{fit}, S_{CP}^{fit}|A_{CP}^{true}, S_{CP}^{true})\ \mathrm{d}A_{CP}^{fit}\ \mathrm{d}S_{CP}^{fit}. \tag{G.17}$$

We determined $(A_{CP}^{best}, S_{CP}^{best})$ using MIGRAD to maximize $\ln(P_{FC}(A_{CP}^{fit}, S_{CP}^{fit}|A_{CP}^{best}, S_{CP}^{best}))$. The distribution of A_{CP}^{best}, S_{CP}^{best}, and value of P_{FC} by each point of $(A_{CP}^{fit}, S_{CP}^{fit})$ is shown in Figs. G.10 and G.11. The integration of $CL(A_{CP}^{true}, S_{CP}^{true})$ is performed numerically. Figure G.12 shows an example confidence region calculated using the method and the response function described in this appendix chapter when the measured result of CP asymmetries is $A_{CP}^{meas} = 0.05$ and $S_{CP}^{meas} = 0.63$.

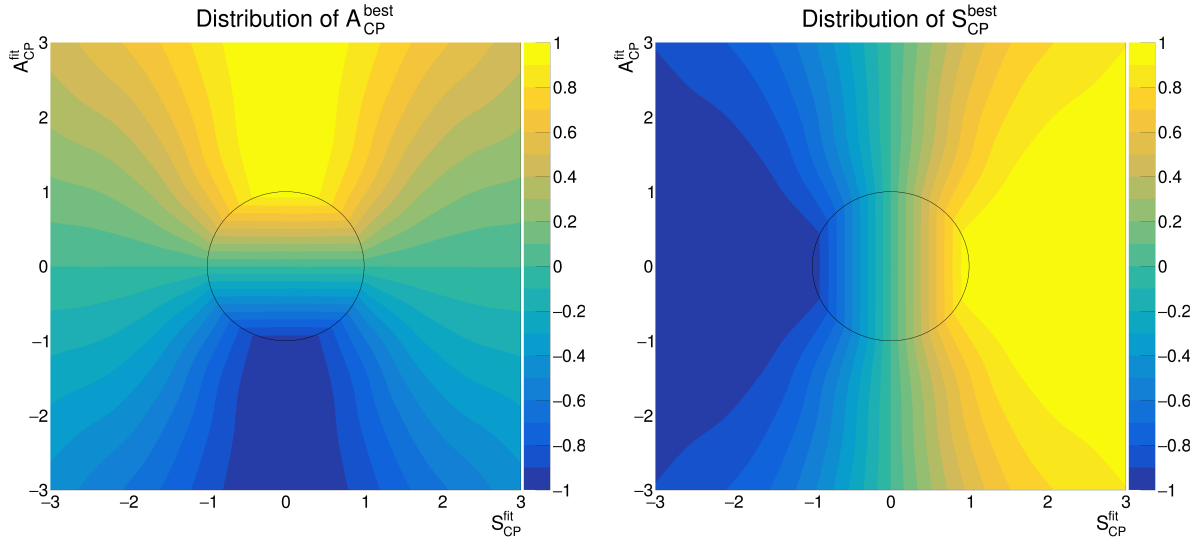


Figure G.10: Distribution of A_{CP}^{best} and S_{CP}^{best}. The black circle shows the physical boundary for A_{CP} and S_{CP}.

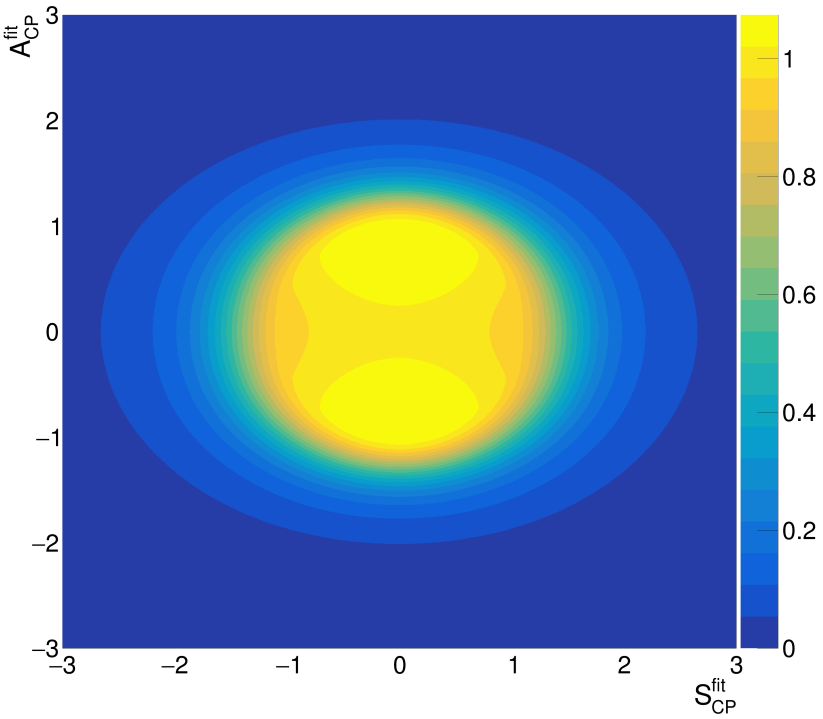


Figure G.11: Distribution of $P_{FC}(A_{CP}^{fit}, S_{CP}^{fit}|A_{CP}^{best}, S_{CP}^{best})$.

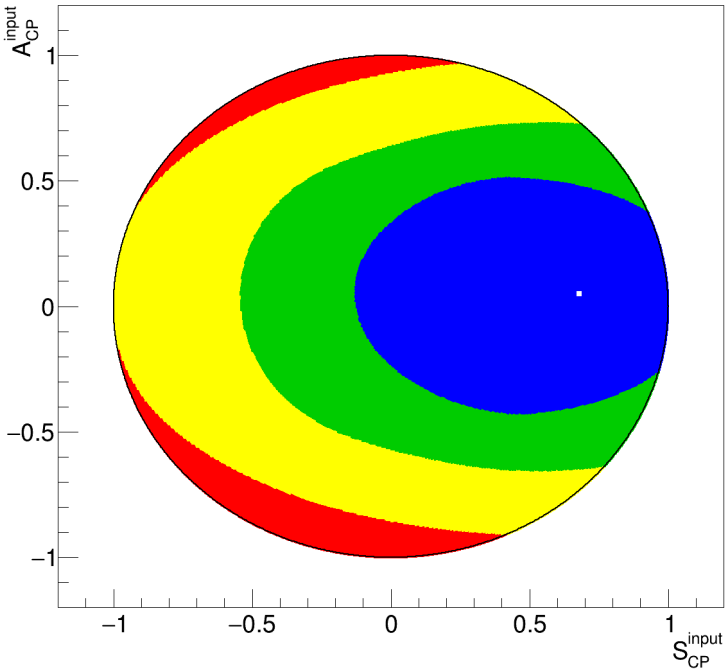


Figure G.12: An example confidence region calculated the Feldman-Cousins approach with response function described above. We assume the measured *CP* asymmetries are $A_{CP}^{meas} = 0.05$ and $S_{CP}^{meas} = 0.63$, of which a white marker indicates the point of the measurement. Legend for the color of each region is the same as Fig. 7.1.

Appendix H

Input Variables to FastBDT for Continuum Suppression and Correlation

We calculated Pearson coefficients between M_{bc}, ΔE, and input variables to FastBDT for the continuum suppression. The coefficients for the main channel are shown in Figures H.1(Signal + SxF) and H.2 (Background), and those for the control channel are shown in Figures H.3 (Signal + SxF) and H.4 (Background). The Figures do not suggest a significant correlation between variables for the signal extraction (M_{bc} and ΔE) and the input variables. In addition, we compared the input variables between (Signal + SxF) and Background categories. The results are shown in Figures H.5 and H.6.

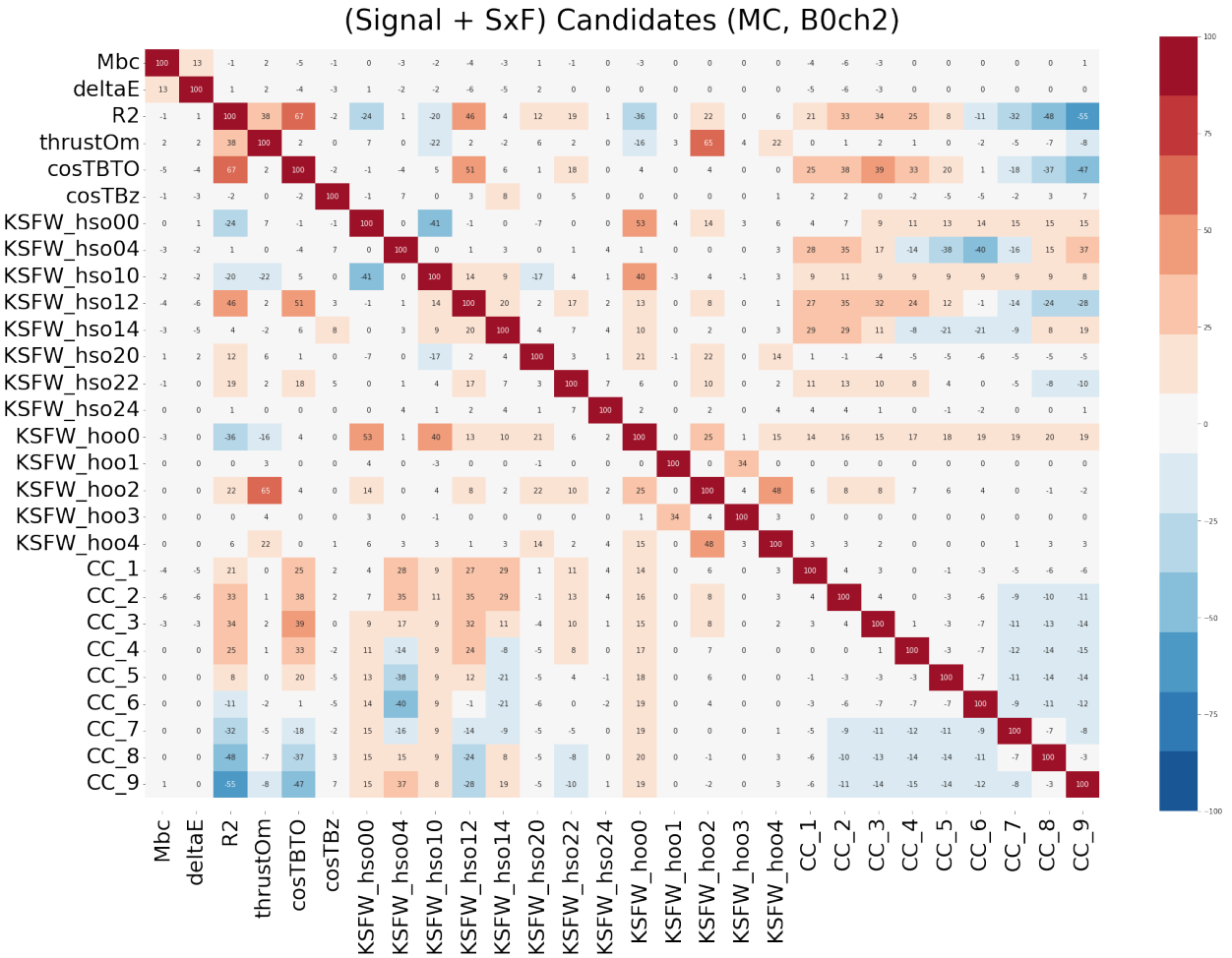


Figure H.1: Pearson correlation coefficients for (Signal + SxF) events in the main channel, between M_{bc}, ΔE, and all input variables for continuum suppression.

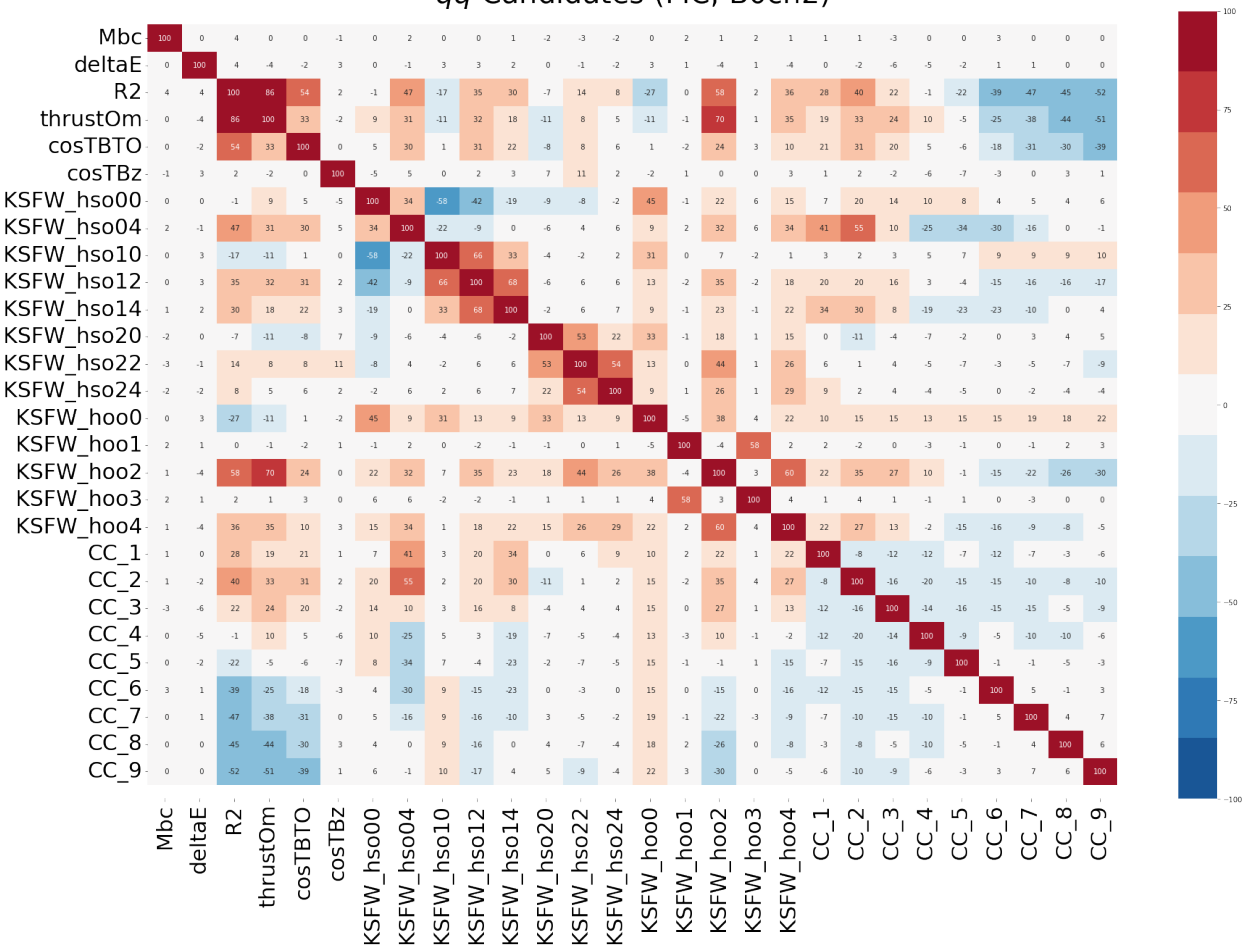


Figure H.2: Pearson correlation coefficients for Background events in the main channel, between M_{bc}, ΔE, and all input variables for continuum suppression.

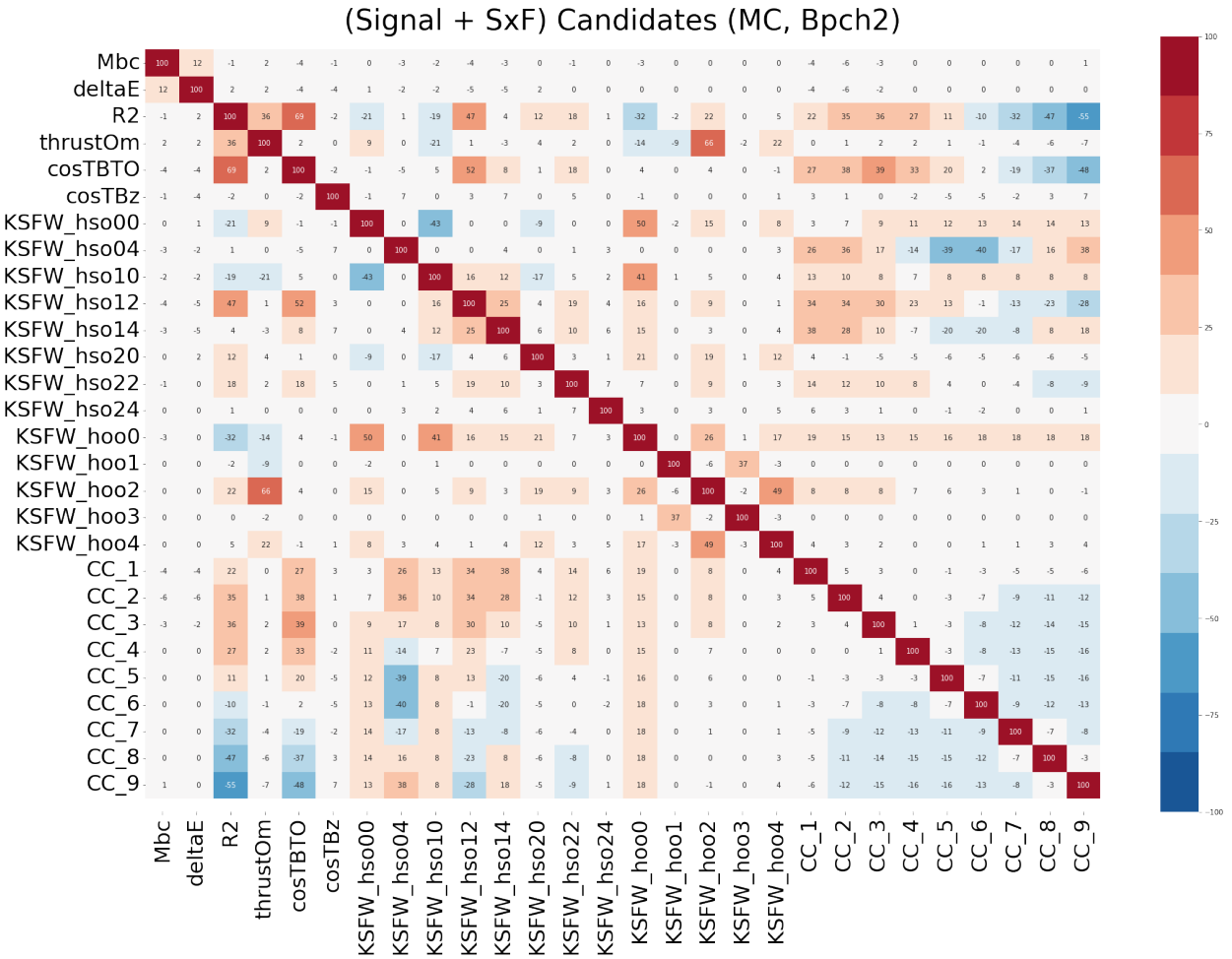


Figure H.3: Pearson correlation coefficients for (Signal + SxF) events in the control channel, between M_{bc}, ΔE, and all input variables for continuum suppression.

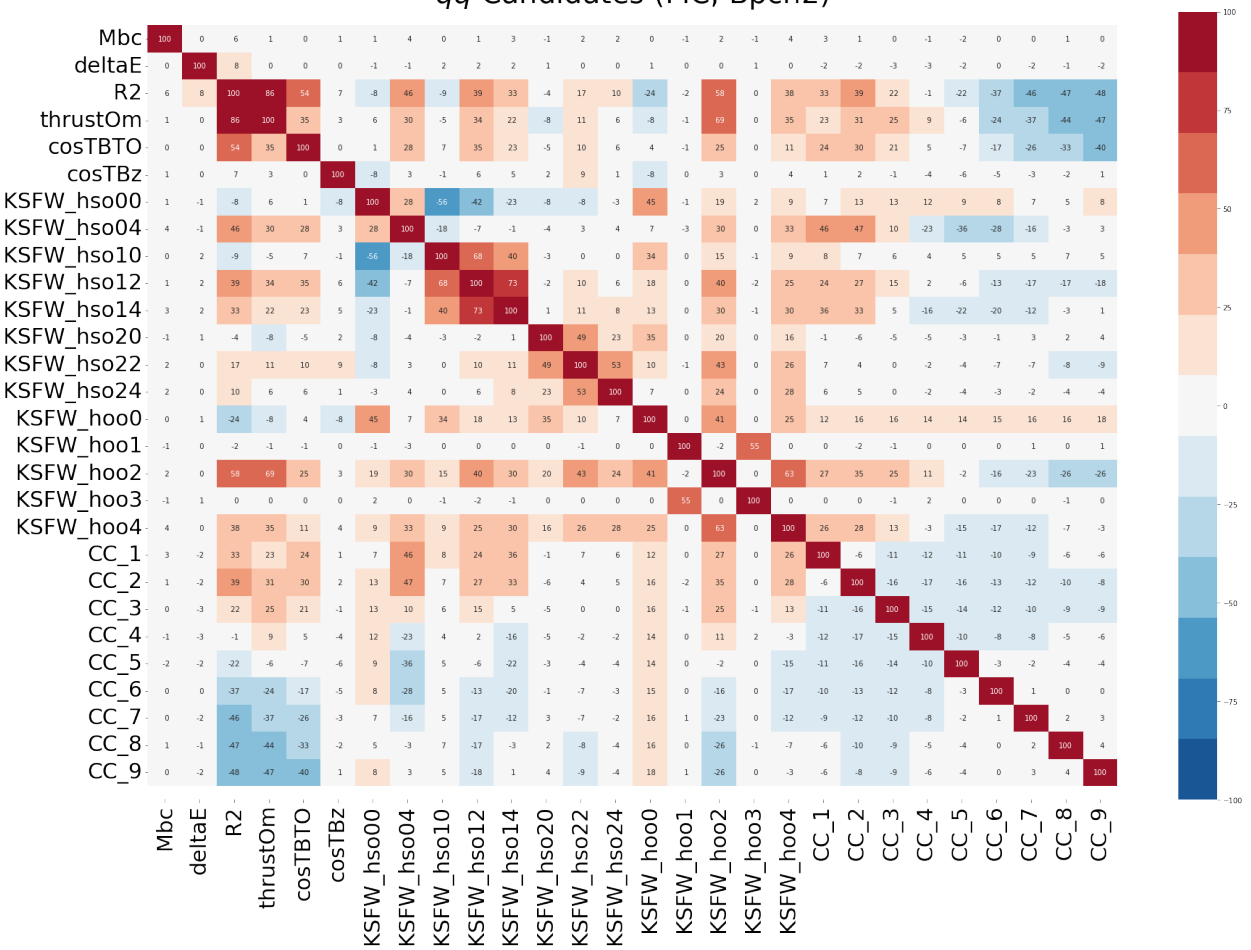


Figure H.4: Pearson correlation coefficients for Background events in the control channel, between M_{bc}, ΔE, and all input variables for continuum suppression.

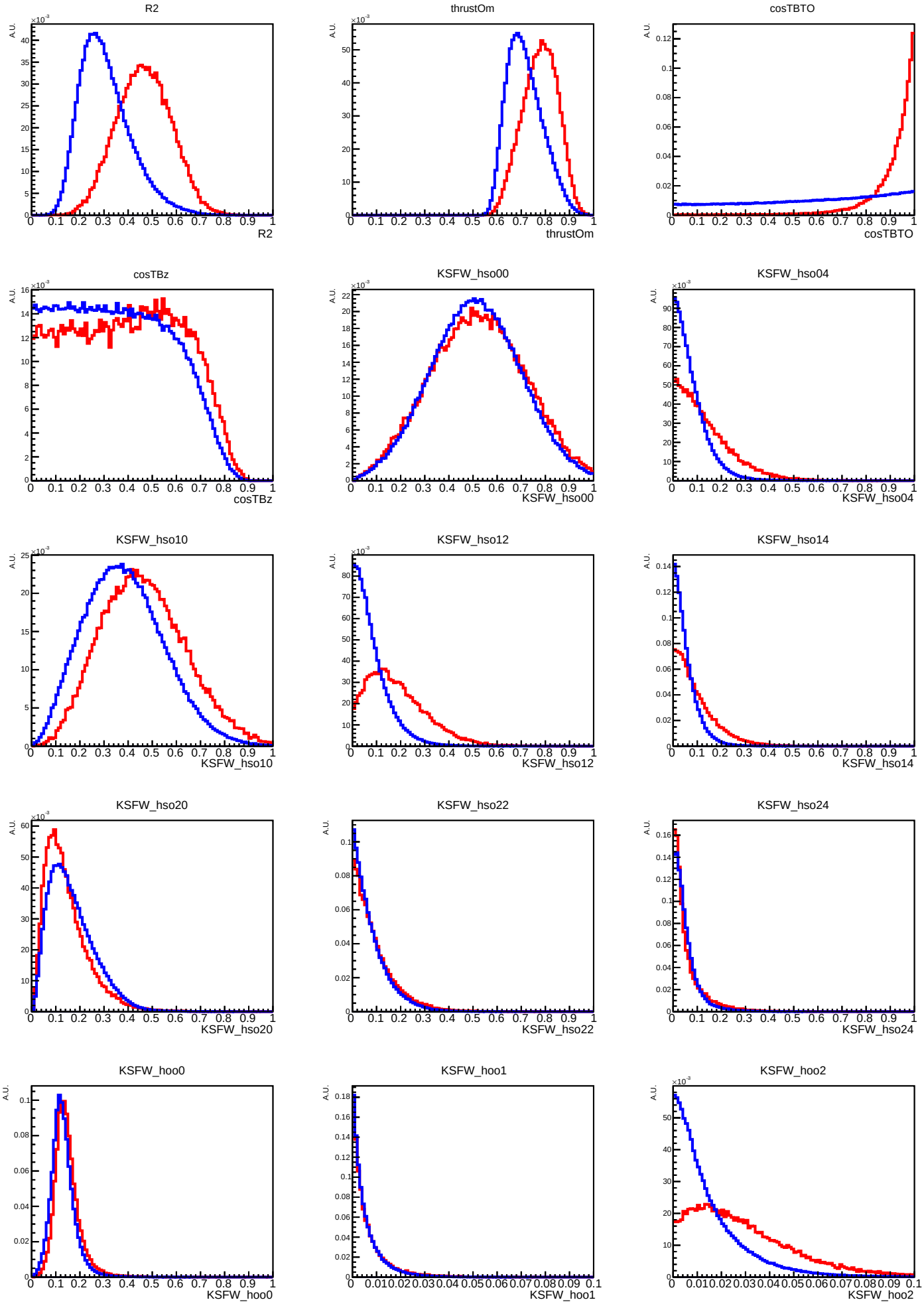


Figure H.5: Comparison of the shape for input variables between (Signal + SxF) and Background categories in the main channel. The blue line is for the (Signal + SxF) category, and the red line is for the Background category where $q\bar{q}$ events dominate.

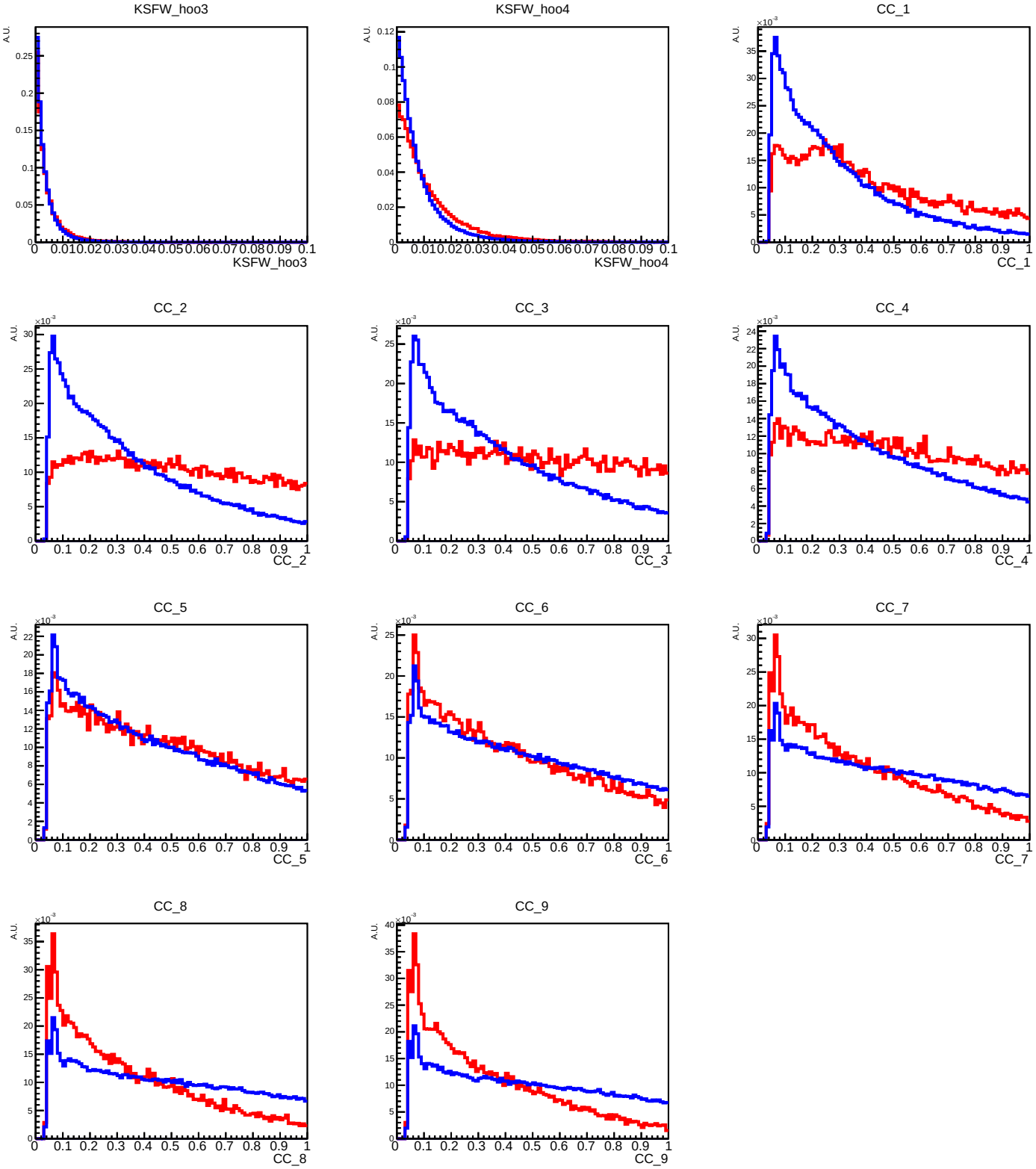


Figure H.6: Comparison of the shape for input variables between (Signal + SxF) and Background categories in the main channel. The blue line is for the (Signal + SxF) category, and the red line is for the Background category where $q\bar{q}$ events dominate.

Appendix I

Ensemble and Linearity Tests with Various Scenarios

We performed various tests using the Toy MC dataset. The following sections will describe the details of each ensemble or linearity test.

I.1 Ensemble tests

We performed ensemble tests with the following configurations:

- Ensemble test with f_{sig} calculated using 1D (M_{bc}, ΔE, and $\mathcal{O}_{CS}$) signal extraction.
- Ensemble test with f_{sig} calculated using 2D ($M_{bc} \cdot \Delta E$) signal extraction.
- Ensemble test with an equivalent statistic to Belle without the number of CDC hits selection to charged tracks.

The following subsections describe the details of each test.

I.1.1 f_{sig} with 1D signal extraction of M_{bc}, ΔE, or $\mathcal{O}_{CS}$

In this scenario, we calculate f_{sig} with $\vec{x} = (M_{bc}, E^{*}_{beam}, \cos(\theta_B^{Boost}), |q_{tag} \cdot r|)$ using Eq. 4.9 for the *CP* fitting. Also, $P_x^{\Delta E}(\Delta E)$ and $P_x^{\mathcal{O}_{CS}}(\mathcal{O}_{CS})$ in Eqs. 4.7 and 4.8 are replaced to 1. An example distribution for f_{sig} only with M_{bc} is shown in Sub-figure I.1. We repeated an ensemble test by changing the number of events in the Background category by 0%, 25%, 50%, 75%, and 100% of the Background yield from the experimental data. We did not change the number of events in the (Signal + SxF) category over all the tests. A MINOS error distribution from ensemble tests the f_{sig} is shown in Fig. I.2. In this case, we can see a significant shift in the distribution by the number of Background events.

Also, we performed similar studies with different f_{sig} configurations with PDFs for the 1D signal extraction of $\vec{x} = (\Delta E, \cos(\theta_B^{Boost}), |q_{tag} \cdot r|)$ for ΔE and $\vec{x} = (\mathcal{O}_{CS}, \cos(\theta_B^{Boost}), |q_{tag} \cdot r|)$

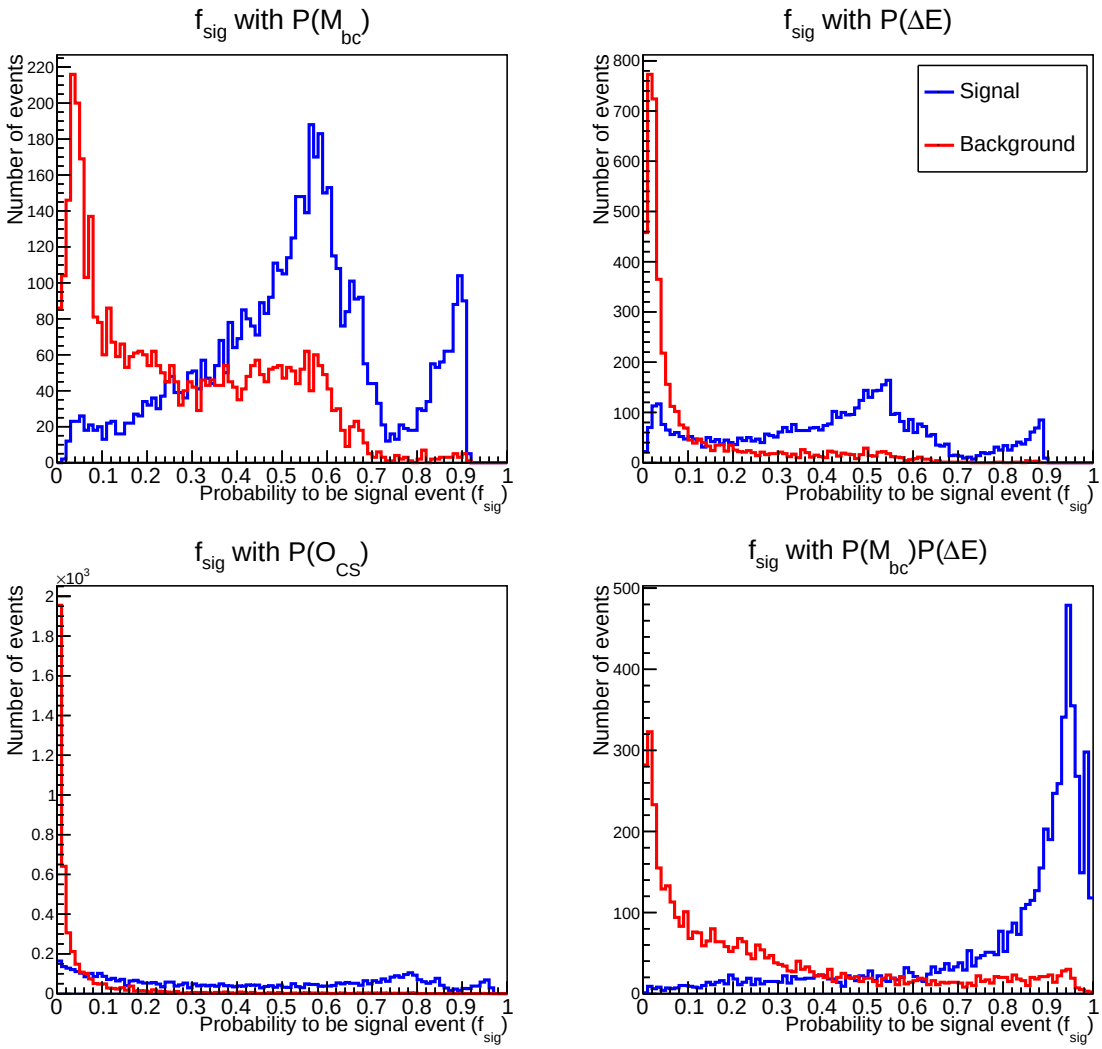


Figure I.1: Example distributions for f_{sig} with various configurations (1D of M_{bc}, ΔE, and $\mathcal{O}_{CS}$ and 2D of $M_{bc} \cdot \Delta E$) for the signal extraction. Events in the signal region are used for the plotting, and the multi-modal shapes are due to the Punzi terms for $\cos(\theta_B^{Boost})$ and $|q_{tag} \cdot r|$. Those plots are drawn with $N_{total} = 81600$ and $F_{sig} = 0.075$.

for $\mathcal{O}_{CS}$. Example distributions of f_{sig} for each case of ΔE and $\mathcal{O}_{CS}$ are shown in Sub-figs. I.1. MINOS error distributions from the ensemble tests with those f_{sig}s are shown in Figs. I.3 and I.4. In those cases, the distributions for the uncertainty show significant shifts by the number of events in the Background category.

I.1.2 f_{sig} with 2D signal extraction of $M_{bc} \cdot \Delta E$

We constructed a 2D PDF of $P^{M_{bc}}(M_{bc}) \cdot P^{\Delta E}(\Delta E)$ to calculate f_{sig} and performed ensemble tests with the f_{sig} and an example distribution for f_{sig} from the generator for the Toy MC dataset is shown in Sub-figure I.1. The methodology is the same as the one stated in the subsection for Toy MC tests, and the result is shown in Fig. I.5. The uncertainty does not shift much even if we adopt f_{sig} with the 2D PDF for the signal extraction, of which separation power is weaker than f_{sig} calculated with a full 3D PDF.

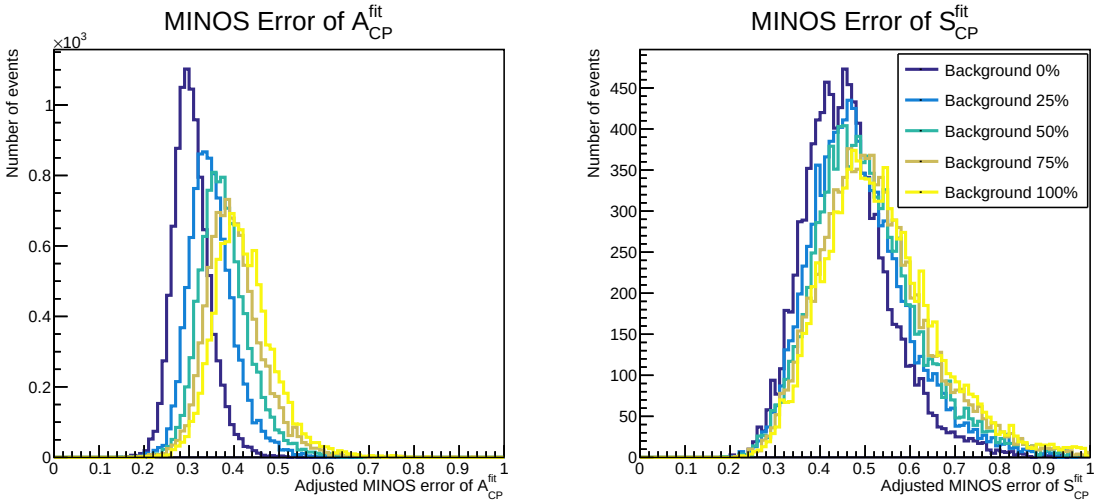


Figure I.2: MINOS error distributions for A_{CP}^{fit} and S_{CP}^{fit} from the ensemble tests with f_{sig} calculated with $P^{M_{bc}}(M_{bc})$. The number of events in the Background category varies from 0% to 100% with 25% spacing.

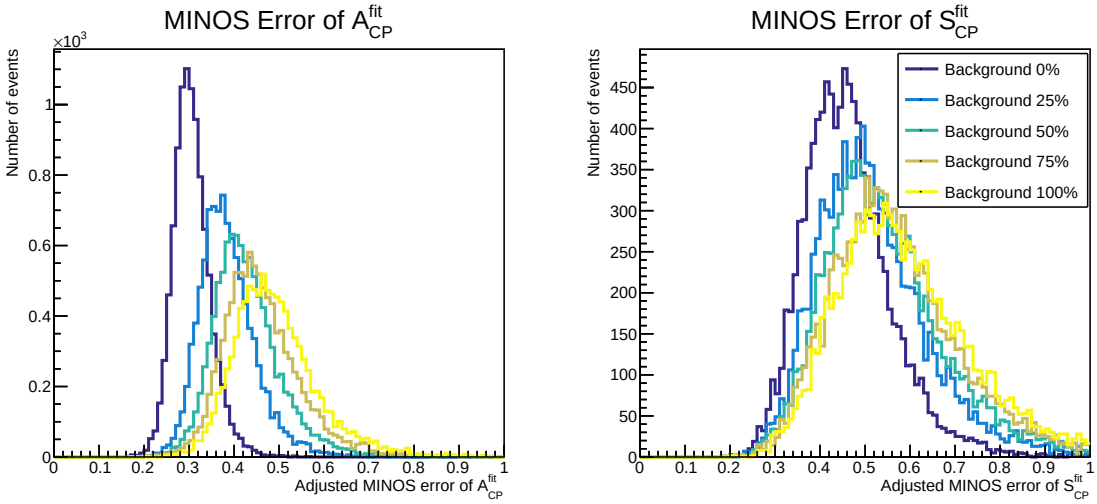


Figure I.3: MINOS error distributions for A_{CP}^{fit} and S_{CP}^{fit} from the ensemble tests with f_{sig} calculated with $P^{\Delta E}(\Delta E)$. The number of events in the Background category is given in the plot.

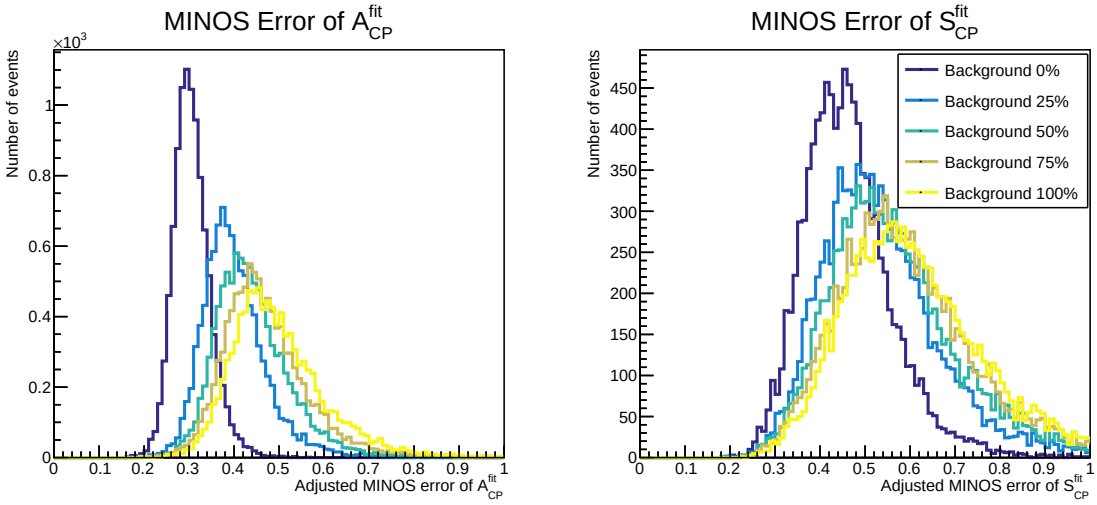


Figure I.4: MINOS error distributions for A_{CP}^{fit} and S_{CP}^{fit} from the ensemble tests with f_{sig} calculated with $P^{\mathcal{O}_{CS}}(\mathcal{O}_{CS})$. The number of events in the Background category is given in the plot.

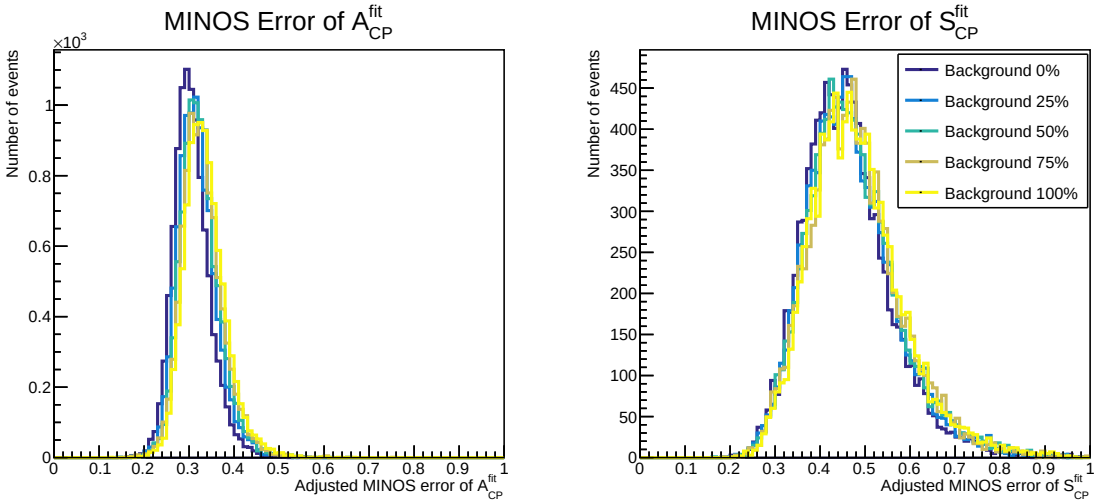


Figure I.5: MINOS error distributions for A_{CP}^{fit} and S_{CP}^{fit} from the ensemble tests with f_{sig} calculated with $P^{M_{bc}}(M_{bc}) \cdot P^{\Delta E}(\Delta E)$.

I.1.3 An equivalent statistic to Belle without the CDC criterion

We compared the sensitivity of this analysis with the result from Belle by performing an ensemble test. The size of a dataset used for the latest analysis of *CP* asymmetry measurement in the $B^0 \to \eta' K_S^0$ channel from Belle is 710 fb. Therefore, the number of events in the (Signal + SxF) category is calculated by multiplying the expected value given in Tab. 4.8 by a factor of 710/362. Also, since the signal efficiency is improved to 1.3 times if we remove the selection criterion, we multiply the number of (Signal + SxF) events by 1.3. The number of Background events is determined using the result of the signal extraction fit, and we also multiply this number by 1.3 to simulate the increasing background events due to removing the selection.

We perform the ensemble test by generating 10000 samples with events of the (Signal + SxF) and Background categories as many as the calculated numbers for each category and check the distribution of MINOS error for A_{CP}^{fit} and S_{CP}^{fit}. The mean values of MINOS errors for A_{CP}^{fit} and S_{CP}^{fit} are 0.1766 and 0.2448, respectively, and these values are 97.57% and 94.52% of Belle result, which indicate the better sensitivity of this analysis than Belle.

I.2 Linearity tests

We performed linearity tests with the following configurations:

- Linearity test with non-zero μ.
- Linearity test with a large dataset of $20\,\mathrm{ab}^{-1}$.

The following subsections describe the details of each test.

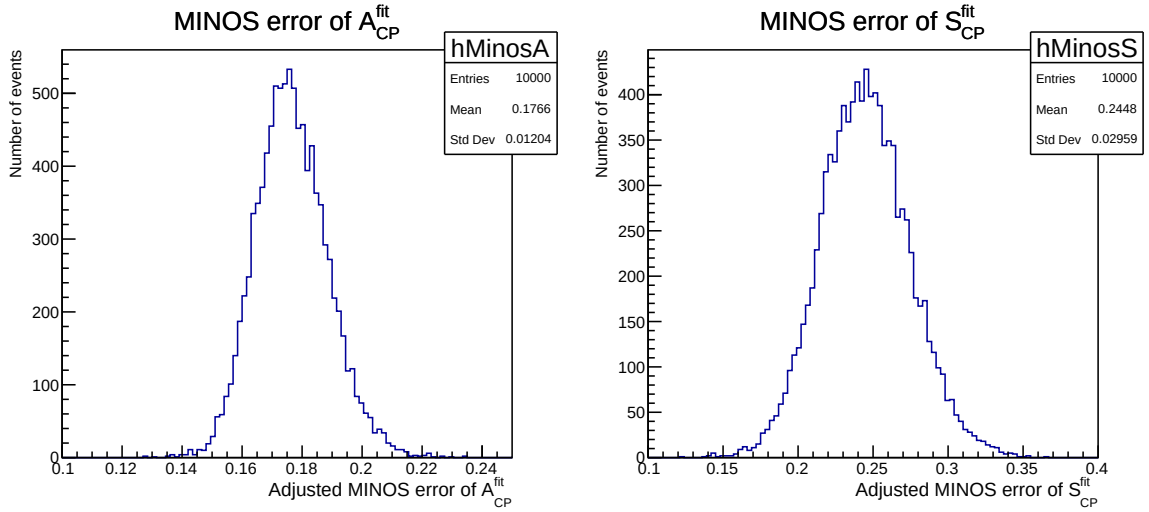


Figure I.6: MINOS error distributions for A_{CP}^{fit} and S_{CP}^{fit} from the ensemble tests with the equivalent statistics to the Belle analysis.

I.2.1 Linearity tests with non-zero μ

We performed a linearity test with a non-zero μ value. We referenced the value from the "Data" column of Table 4.1. In the test, we used configurations for the *CP* fitting from the experimental data and with $10 \cdot N_{total}$ where N_{total} is the total number from the signal extraction fit for the experimental data. The number of samples from the Toy MC generator is 10000 for each $(A_{CP}^{input}, S_{CP}^{input})$ and Fig. I.7. From the linearity test, we can confirm the constant bias in A_{CP}^{fit} over all inputs.

I.2.2 Linearity tests with a large dataset of $20\,\mathrm{ab}^{-1}$

We performed a linearity test with a large dataset of $20\,\mathrm{ab}^{-1}$. The number of samples from the Toy MC generator is 10000 for each $(A_{CP}^{input}, S_{CP}^{input})$ and Fig. I.8. From the linearity test, we can confirm the bias and the bias and underestimated uncertainties that emerge in tests with the small statistics, such as shown in Fig. 6.12 are largely improved.

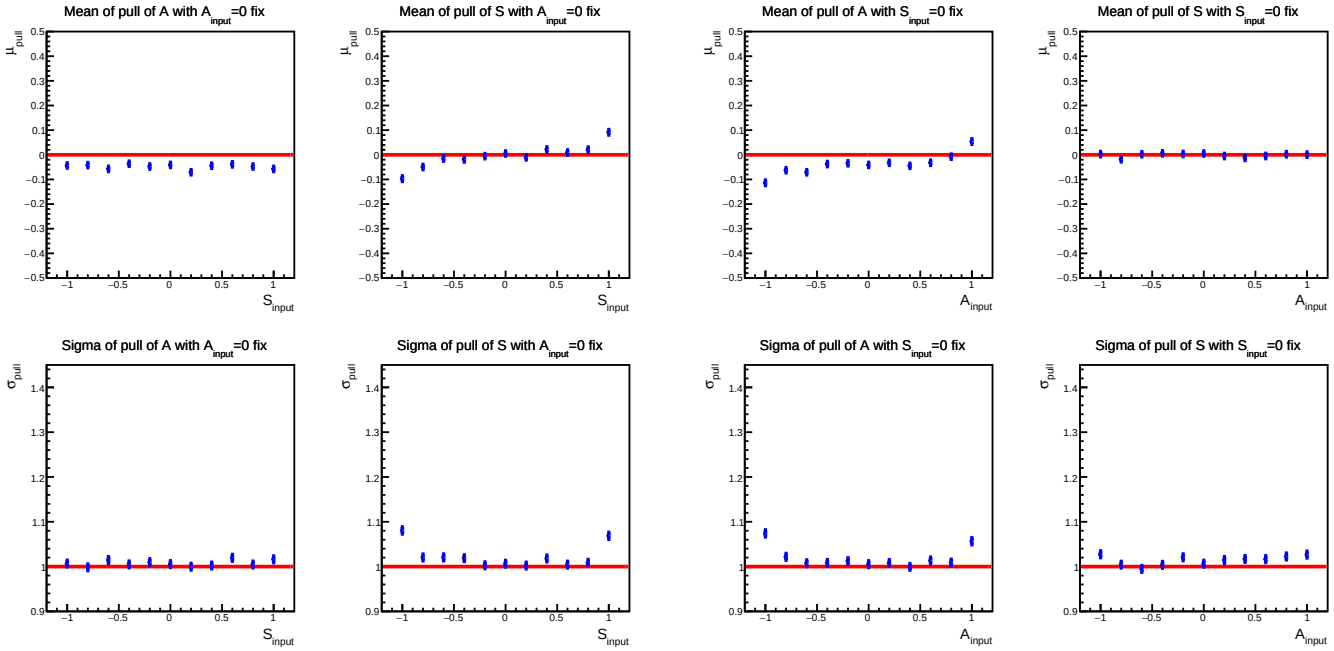


Figure I.7: Result of the linearity test with non-zero μ value. Blue bars indicate fitted μ and σ for the pull distributions for $(A_{CP}^{fit}, S_{CP}^{fit})$ and their error.

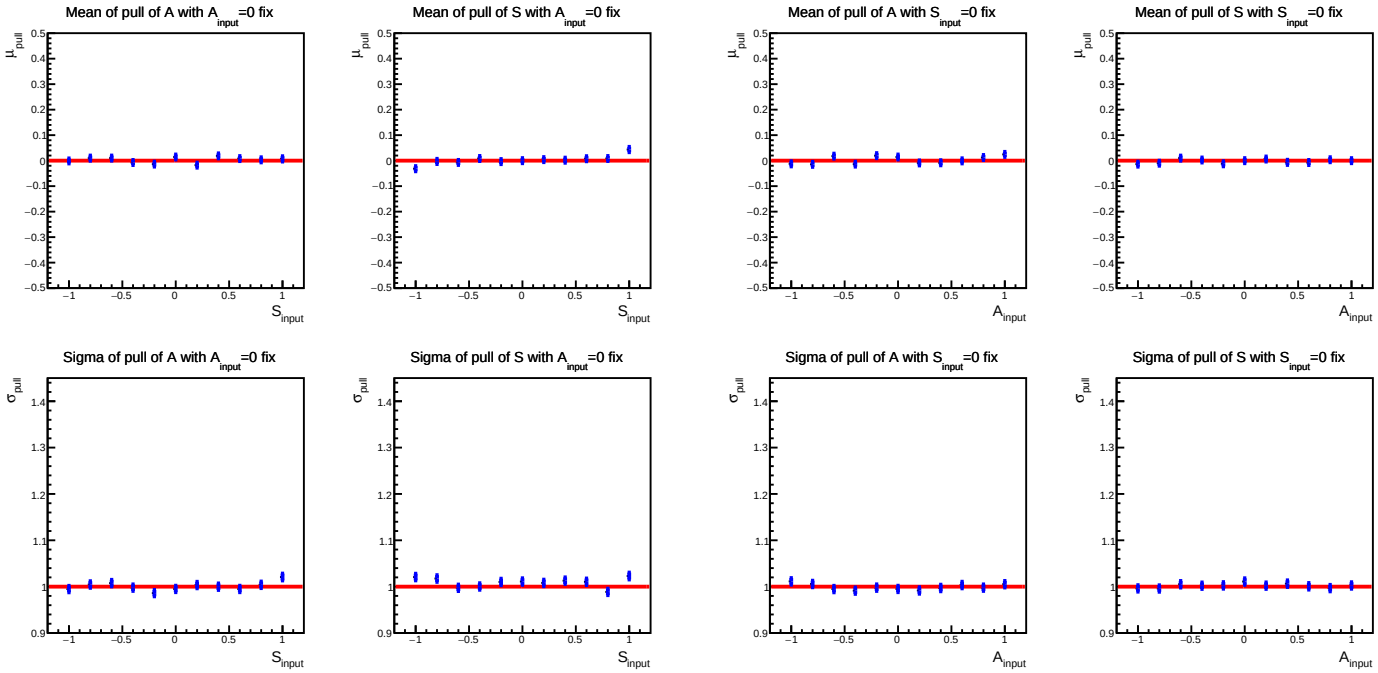


Figure I.8: Result of the linearity test with a large dataset of $20\,\mathrm{ab}^{-1}$. Blue bars indicate fitted μ and σ for the pull distributions for $(A_{CP}^{fit}, S_{CP}^{fit})$ and their error.

Appendix J

Resolution Function and Background Δt for Control Channel

We analyzed the control channel to validate our analysis procedure in the main channel. To achieve that purpose, we should unify analysis configurations for the control channel with that for the main channel as much as possible.

In the control channel, five charged tracks directly come from the vertex of $B^{\pm}_{CP}$, so the resolution for *CP*-side vertex fitting will be enhanced than that in the main channel. To mimic configurations for the *CP*-side vertex fitting in the main channel, we multiplied 1000 times the uncertainty of track parameters of d_0 and z_0 of $K^{\pm}$ so the particle does not anticipate the vertex fitting. Also, to emulate the reduced χ^2 value in the main channel, the reduced χ^2 value for the vertex fitting on the *CP* side in the control channel was transformed as $(\chi^2/ndf)^{CP} \to (\chi^2/(ndf-2))^{CP}$, where the $ndf-2$ is an effective ndf value due to removing track parameters of d_0 and z_0 of $K^{\pm}$.

We compared variables that indicate the quality of the vertex fitting for the *CP* and tag side in Figure J.1 and J.2, respectively. Reduced χ^2 distribution with the modified formula for the control channel is similar to that without modification in the main channel, but there is a bump in the distribution of ℓ uncertainty from the *CP*-side vertex fitting (σ_ℓ^{CP}) around $10-13\mu$m in the main channel which does not exist in that in the control channel. The reason for the bump is events of which K^0_S has PXD hits. This conclusion can be justified with facts of input particles having PXD hits improve the uncertainty from the vertex fitting, and if we apply a cut of the flight length of K^0_S is longer than 4cm so that K^0_S cannot have the PXD hits, then the bump disappears and distributions for σ_ℓ^{CP} for main and control channels become similar. Since the conditional variables for the *CP*-side resolution function were different, we consider the possible difference of R_{CP} between the control and main channel in the f^{CP}_{core} parameter. The other four parameters for R_{CP} are the same for the main channel. Determined parameters are shown in Table 5.1.

The distributions of variables for the vertex quality of B^0_{tag} show some discrepancy

between the main and control channels, but it is insignificant, so we did not put any special treatments for R_{tag} in the control channel in σ or the fraction parameters. However, for the $B^{\pm}_{tag}$, the lifetime of the long-lived intermediate state is different from that in the main channel. Thus, we consider differences by determining $\tau^{tag}_{\mathrm{exp},const}, \tau^{tag}_{\mathrm{exp},slope}$ newly, and other parameters will be shared with that for the main channel. Table 5.2 shows the parameters.

For the background Δt PDF, we estimated it with the same methodology as the main channel. Fitting parameters for Δt PDF are given in Table J.1 and curves with data points are shown in Figures J.3 for MC and J.4 for Data.

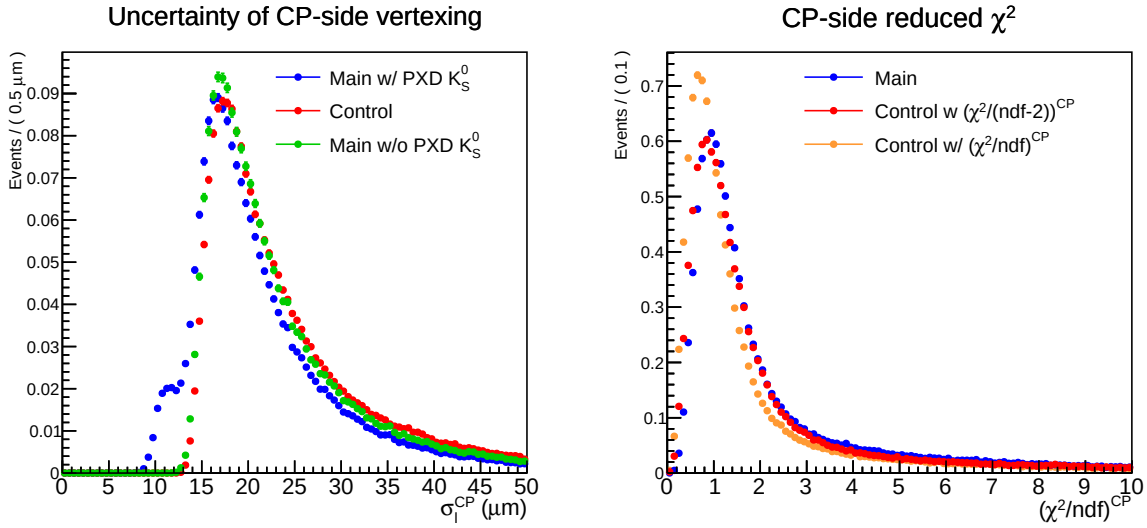


Figure J.1: Comparison of variables for the quality of *CP*-side vertex fitting.

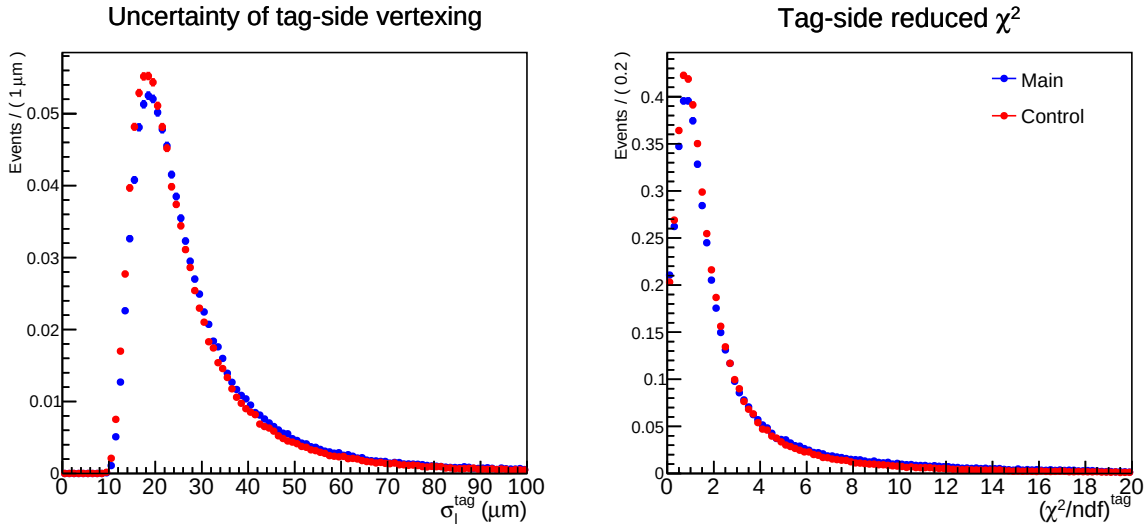


Figure J.2: Comparison of variables for the quality of tag-side vertex fitting.

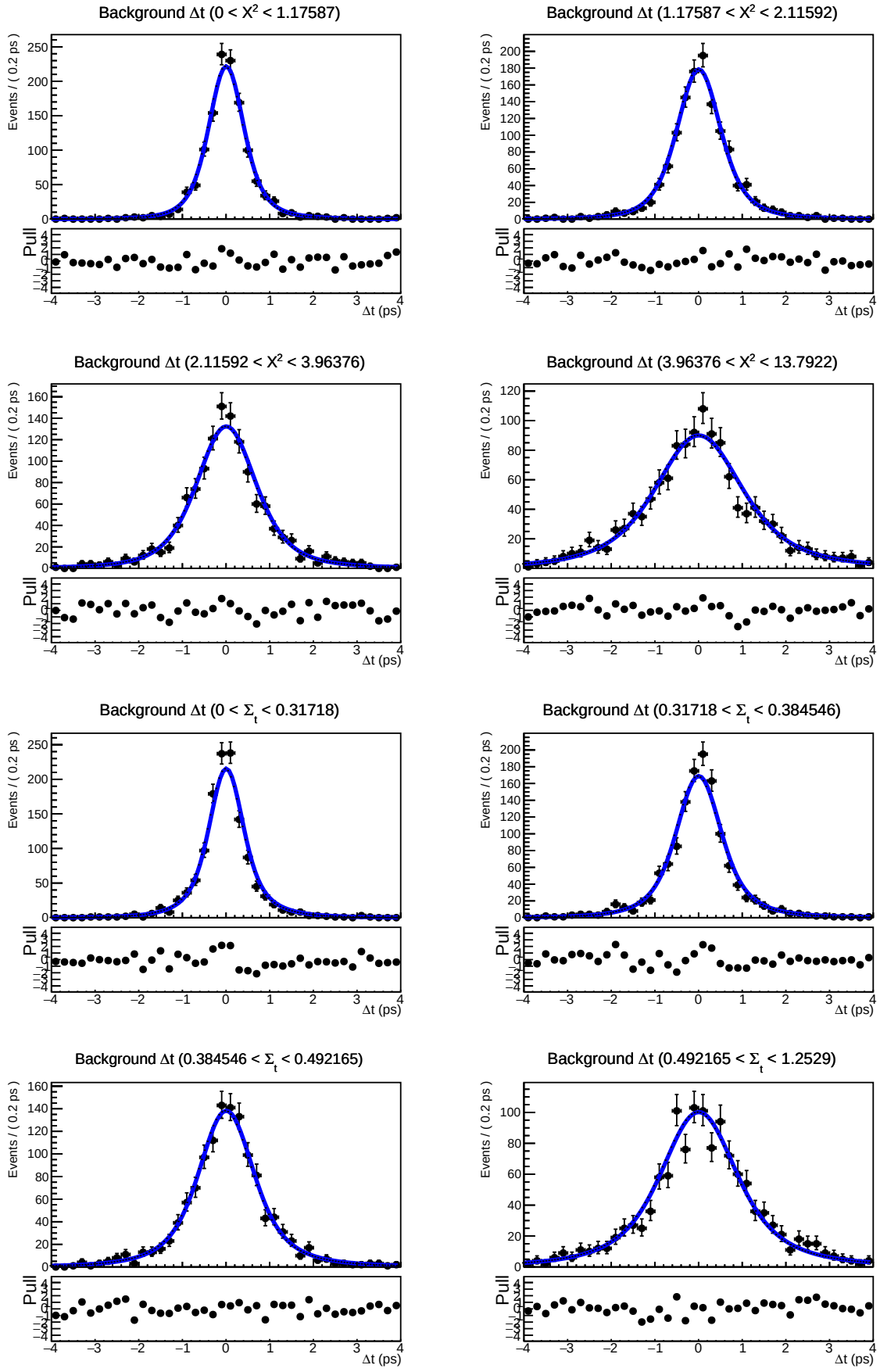


Figure J.3: Δt distribution of background events from 400/fb Generic MC and fitted curve of Background Δt PDF for the control channel. The upper four plots are sliced by X, and Σ_t slices the below four. The unit of Σ_t is ps, and each slicing contains 25% of total events.

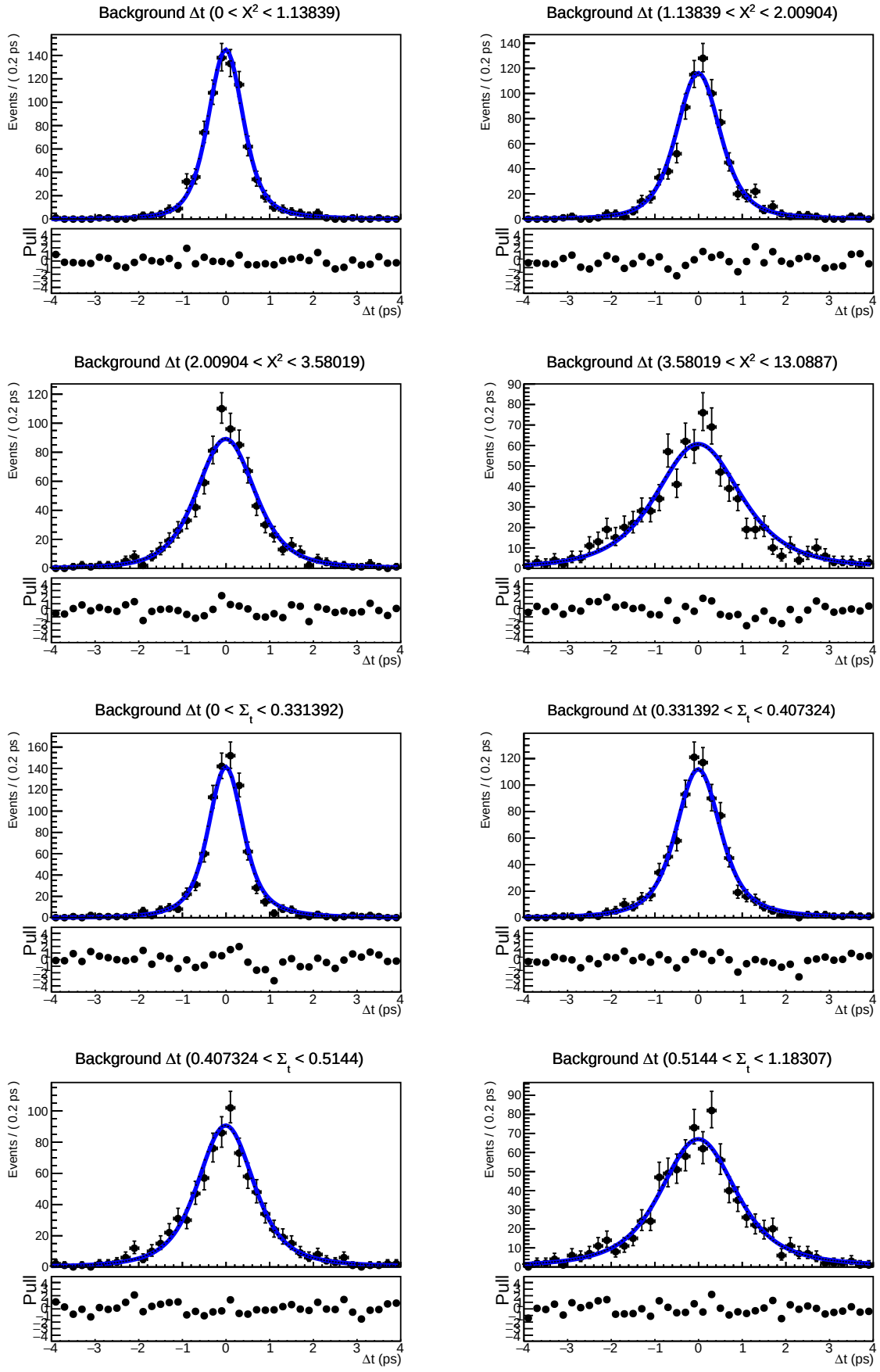


Figure J.4: Δt distribution of sideband events from experimental data and fitted curve of Background Δt PDF for the control channel. The upper four plots are sliced by X, and Σ_t slices the below four. The unit of Σ_t is ps, and each slicing contains 25% of total events.

Channel	Parameters	Fitted value using	
		MC backgrounds	Sideband in experimental data
Control	f_{core}^{bckg}	$0.88300^{+0.01116}_{-0.01213}$	$0.89595^{+0.01570}_{-0.01765}$
	μ_{core}^{bckg}	$0.00379^{+0.00570}_{-0.00570}$	$-0.00843^{+0.01031}_{-0.01032}$
	c_{core}^{bckg}	$0.85498^{+0.01906}_{-0.01903}$	$0.78349^{+0.03464}_{-0.03451}$
	s_{core}^{bckg}	$0.32691^{+0.00947}_{-0.00958}$	$0.33146^{+0.01622}_{-0.01610}$
	μ_{tail}^{bckg}	$0.13555^{+0.04462}_{-0.04300}$	$0.24142^{+0.10304}_{-0.09602}$
	c_{tail}^{bckg}	$2.17520^{+0.16020}_{-0.14332}$	$2.59629^{+0.30699}_{-0.26581}$
	s_{tail}^{bckg}	$0.86435^{+0.05952}_{-0.05725}$	$0.79387^{+0.11772}_{-0.10448}$

Table J.1: List of parameters of Background Δt PDFs from MC and the Data.

Appendix K

Supplemental Studies for Lifetime Fitting

In this chapter, we will discuss the supplemental studies for *CP* fittings.

K.1 Determination of lifetime using intermediate MC information

We performed additional three fittings to determine the lifetime of B^0 and $B^\pm$ with the intermediate true information from the MC generator. The variable and PDFs for the fitting for each variable are as follows:

- `mcDeltaT` $\left(\frac{\ell_{CP}^{true}-\ell_{tag}^{true}}{\beta\gamma c}\right)$: Kinematically-smeared differences of decay times between B^0_{CP} and B^0_{tag}. This fitting uses a PDF of E_k and provides the lifetime value of B^0 without any effects from the Belle II detector.
- $\frac{\ell_{CP}-\ell_{tag}^{true}}{\beta\gamma c}$ and $\frac{\ell_{CP}^{true}-\ell_{tag}}{\beta\gamma c}$: $\ell_x^{true}/\beta\gamma c$ means decay time of B^0_x without effects from x-side vertex fitting. In this fitting, we use $E_k \otimes R_{CP}$ and $E_k \otimes R_{tag}$, and it checks the validity of the resolution function and its parameters for the *CP* and tag-side, respectively.

where E_k is given in Equation 5.7.

The results of the lifetime determination with the above three variables are given in Table K.1 and Figures K.1, K.2, and K.3. The determined lifetimes fit well with the reference lifetime that is determined using `mcDeltaTau` variable, and the fitted curves show good agreements with the data points of each variable.

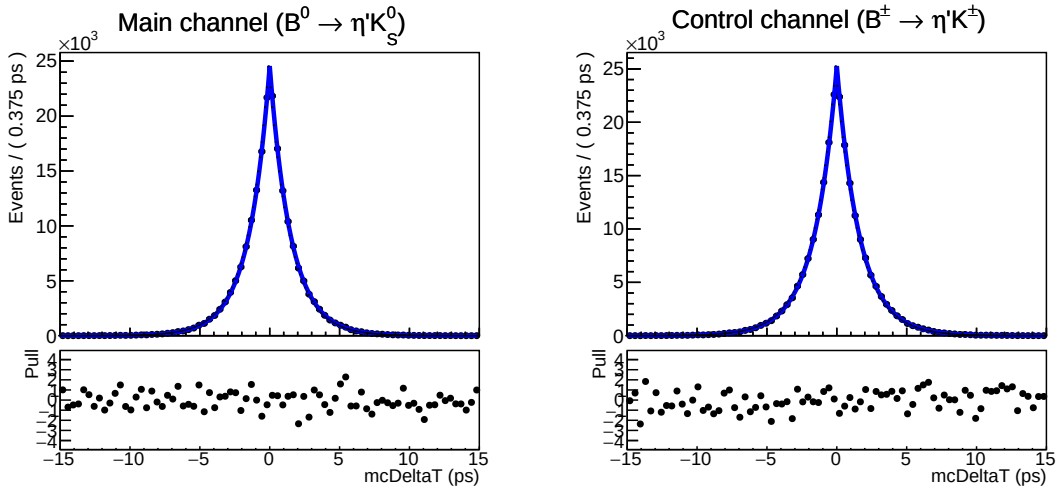


Figure K.1: mcDeltaT distribution and fitted curve of E_k for main and control channel.

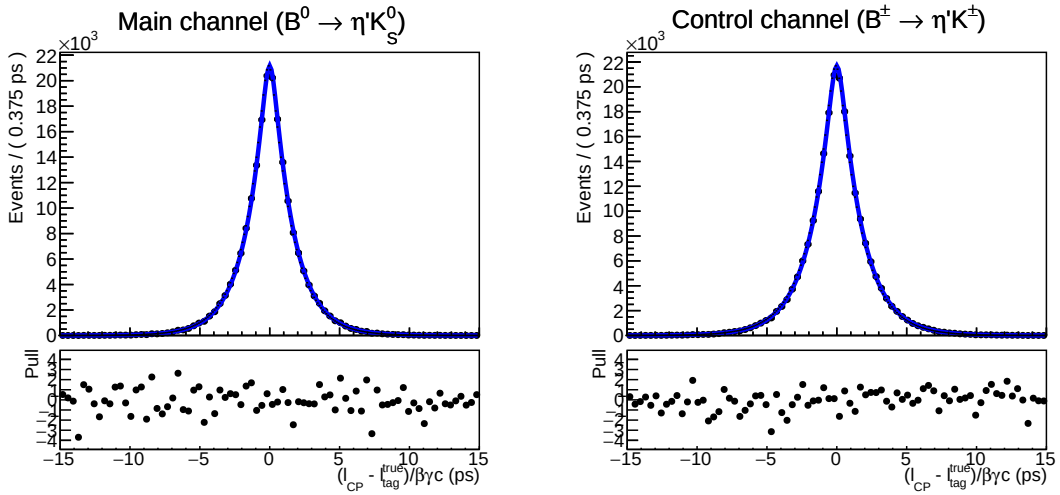


Figure K.2: $\frac{\ell_{CP}-\ell^{true}_{tag}}{\beta\gamma c}$ distribution and fitted curve of $E_k \otimes R_{CP}$ for main and control channel.

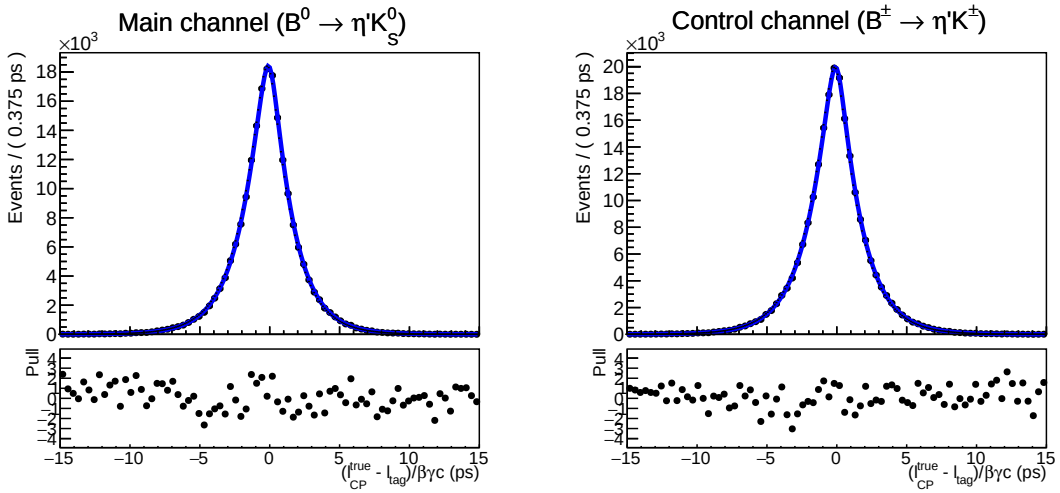


Figure K.3: $\frac{\ell^{true}_{CP}-\ell_{tag}}{\beta\gamma c}$ distribution and fitted curve of $E_k \otimes R_{tag}$ for main and control channel.

Variables for fitting	Fitted τ_{B^0}	Fitted $\tau_{B^\pm}$
`mcDeltaTau`	1.520(3)ps	1.641(4)ps
`mcDeltaT`	1.521(3)ps	1.642(4)ps
$\frac{\ell_{CP}-\ell_{tag}^{true}}{\beta\gamma c}$	1.523(3)ps	1.642(4)ps
$\frac{\ell_{CP}^{true}-\ell_{tag}}{\beta\gamma c}$	1.518(4)ps	1.632(4)ps

Table K.1: Determined lifetimes of B^0 and $B^\pm$ from each step. The result from mcDeltaTau is given as a reference value.

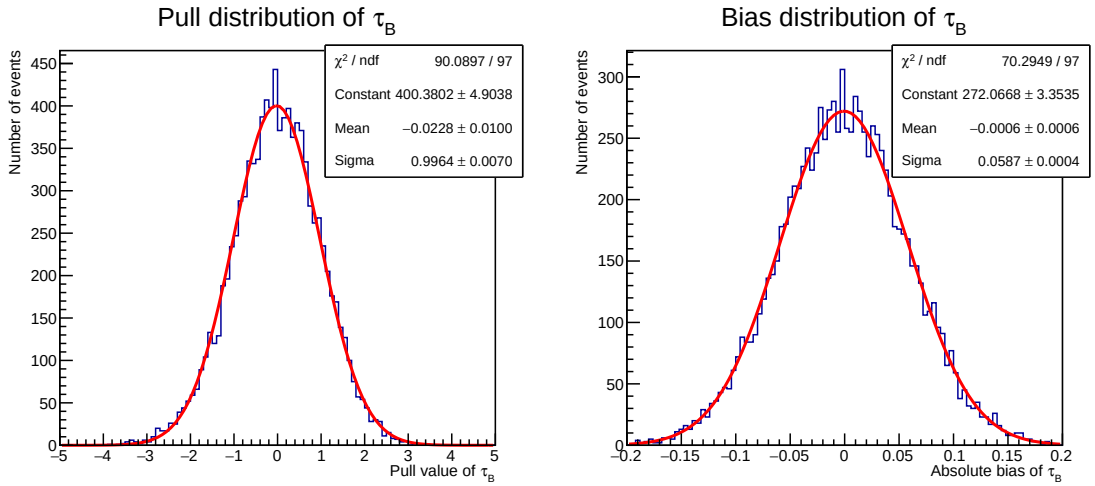


Figure K.4: Pull and absolute bias distribution from the ensemble test for the B^0 lifetime fitting.

Appendix L

The Effects of Background Events on Sensitivity

We repeated ensemble tests by changing the number of events in the Background category with 0%, 25%, 50%, 75%, 100%, 1000%, and finally 2000% to confirm the effect of events from the category. Figure L.1 shows the MINOS error distributions from each trial, and we can confirm the error is not depending much on the number of events from Background categories.

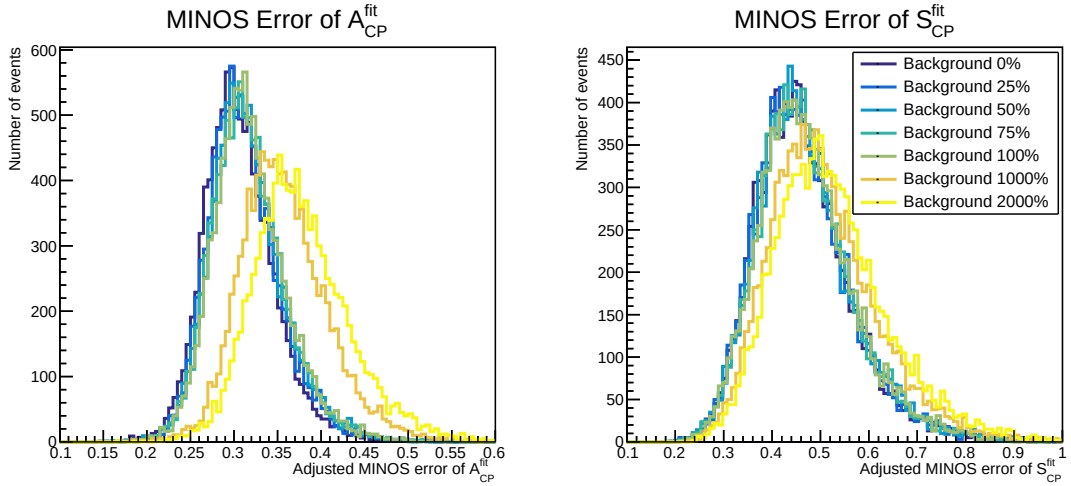


Figure L.1: Distributions of MINOS errors from the ensemble test for the *CP* fitting with configurations from the experimental data.

We consider the reason for such weak dependence of uncertainty for $\mathcal{A}_{CP}^{fit}$ and $\mathcal{S}_{CP}^{fit}$ on N_{bckg} is thanks to the good separation power of f_{sig} given in Equation 4.5. Figure L.2 shows the distribution of the f_{sig} from a generator of the Toy MC dataset with $N_{total} = 81600$ and $F_{sig} = 0.075$ which is given in Table 4.7. If we weaken the separation power of f_{sig} by calculating the value with 1D or 2D signal extraction and perform the *CP* fitting with the newly calculated f_{sig}, then the error distribution shifts significantly

by the number of events of the Background category. (Appendix I)

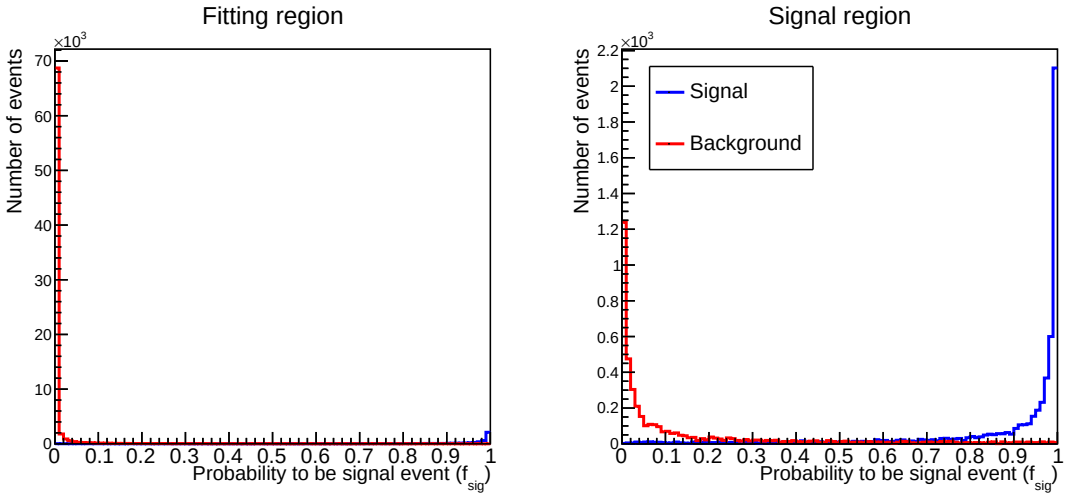


Figure L.2: Distribution of f_{sig} from Toy MC dataset generator. The left plot shows the distribution from the entire fitting region, and the right one shows that from the signal region.

Therefore, we can conclude that the number of Background events is not a big deal, but the number of (Signal + SxF) events is important for the sensitivity for $\mathcal{A}_{CP}$ and $\mathcal{S}_{CP}$ estimation with the current statistics and configurations for the fitting.

Also, we can confirm that $\delta\mathcal{A}_{CP}$ is more affected by N_{bckg} than $\delta\mathcal{S}_{CP}$. We can find the reason from the shape of Background Δt PDF. Thanks to the good resolution for Δt determination of the Background events, events in the Background category are mainly populating in the small-Δt region and will dilute the asymmetry in the region. Because $\mathcal{A}_{CP}$ is a coefficient for cos term, we can conclude that the diluted asymmetry by the Background events mainly affects $\delta\mathcal{A}_{CP}$. Comparison of curves between Background Δt PDF and $\sin^2(\Delta m_d \Delta t)$ or $\cos^2(\Delta m_d \Delta t)$ is shown in Figure L.3.

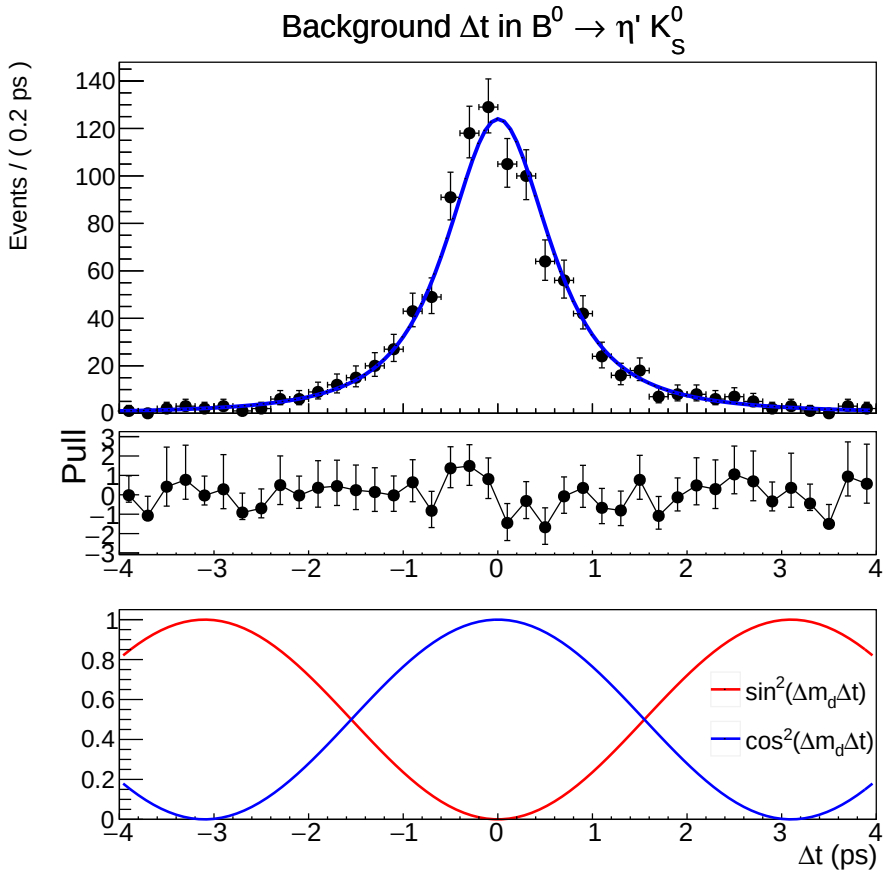


Figure L.3: Comparison of sine or cosine curves and Δt distribution of MC backgrounds with Background Δt PDF. We can see most of the events distributes in $\cos^2(\Delta m_d \Delta t) > \sin^2(\Delta m_d \Delta t)$ region.

Bibliography

[1] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the k_2^0 meson. *Phys. Rev. Lett.*, 13:138–140, Jul 1964. doi: 10.1103/PhysRevLett.13.138. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[2] Makoto Kobayashi and Toshihide Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973. doi: 10.1143/PTP.49.652.

[3] I. I. Bigi and A. I. Sanda. *Quantum mechanics of neutral particles.* Cambridge University Press, 2 edition, 2009. doi: 10.1017/CBO9780511581014.008.

[4] K. Abe et al. Observation of large CP violation in the neutral b meson system. *Physical Review Letters*, 87(9), aug 2001. doi: 10.1103/physrevlett.87.091802. URL: https://doi.org/10.1103%2Fphysrevlett.87.091802.

[5] C. Hagedorn, R. N. Mohapatra, E. Molinaro, C. C. Nishi, and S. T. Petcov. CP violation in the lepton sector and implications for leptogenesis. *International Journal of Modern Physics A*, 33(05n06):1842006, feb 2018. doi: 10.1142/s0217751x1842006x. URL: https://doi.org/10.1142%2Fs0217751x1842006x.

[6] Nicola Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, Jun 1963. doi: 10.1103/PhysRevLett.10.531. URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[7] Lincoln Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945, 1983. doi: 10.1103/PhysRevLett.51.1945.

[8] John S. Hagelin. Mass mixing and cp violation in the b0-b0 system. *Nuclear Physics B*, 193(1):123–149, 1981. ISSN 0550-3213. doi: https://doi.org/10.1016/0550-3213(81)90521-6. URL: https://www.sciencedirect.com/science/article/pii/0550321381905216.

[9] Ulf-G. Meissner and Willibald Plessas, editors. *Lectures on Flavor Physics.* Springer Berlin Heidelberg, 2004. doi: 10.1007/b98411. URL: https://doi.org/10.1007%2Fb98411.

[10] W. Altmannshofer et al. The Belle II Physics Book. *PTEP*, 2019(12): 123C01, 2019. doi: 10.1093/ptep/ptz106,10.1093/ptep/ptaa008. [erratum: PTEP2020,no.2,029201(2020)].

[11] R. L. Workman et al. Review of Particle Physics. *PTEP*, 2022:083C01, 2022. doi: 10.1093/ptep/ptac097.

[12] M. Beneke. Corrections to $\sin(2\beta)$ from cp asymmetries in $b^0 \to (\pi^0, \rho, \eta\eta', \omega, \phi)k_s^0$ decays. *Physics Letters B*, 620(3-4):143–150, aug 2005. doi: 10.1016/j.physletb.2005.06.045. URL: https://doi.org/10.1016%2Fj.physletb.2005.06.045.

[13] Y. Amhis et al. Averages of b-hadron, c-hadron, and τ -lepton properties as of summer 2016. *The European Physical Journal C*, 77(12), dec 2017. doi: 10.1140/epjc/s10052-017-5058-4. URL: https://doi.org/10.1140%2Fepjc%2Fs10052-017-5058-4.

[14] Mikolaj Misiak, Stefan Pokorski, and Janusz Rosiek. Supersymmetry and FCNC effects. *Adv. Ser. Direct. High Energy Phys.*, 15:795–828, 1998. doi: 10.1142/9789812812667_0012.

[15] S. Khalil and E. Kou. On supersymmetric contributions to the CP asymmetry of the $b \to \phi k_s^0$ process. *Physical Review D*, 67(5), March 2003. doi: 10.1103/physrevd.67.055009. URL: https://doi.org/10.1103/physrevd.67.055009.

[16] L. Šantelj et al. Measurement of Time-Dependent CP Violation in $B^0 \to \eta' K^0$ Decays. *JHEP*, 10:165, 2014. doi: 10.1007/JHEP10(2014)165.

[17] Bernard Aubert et al. Measurement of time dependent CP asymmetry parameters in B0 meson decays to omega K0(S), eta-prime K0, and pi0 K0(S). *Phys. Rev. D*, 79:052003, 2009. doi: 10.1103/PhysRevD.79.052003.

[18] Luka Šantelj. Belle Note 1304 – measurement of time-dependent cp asymmetry $B^0 \to \eta' K_S^0$ decays, July 2013. URL: https://belle.kek.jp/secured/belle_note/gn1304/BN1304_v1.pdf.

[19] Yasmine Sara Amhis et al. Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018. *Eur. Phys. J.*, C81:226, 2021. doi: 10.1140/epjc/s10052-020-8156-7. updated results and plots available at https://hflav.web.cern.ch/.

[20] Kazunori Akai, Kazuro Furukawa, and Haruyo Koiso. SuperKEKB Collider. *Nucl. Instrum. Meth.*, A907:188–199, 2018. doi: 10.1016/j.nima.2018.08.017.

[21] T. et al. Abe. Belle II Technical Design Report. 2010. URL: https://arxiv.org/abs/1011.0352.

[22] The Belle II collaboration. Belle-II (Nano beam) IR=±41.5 mrad. (Top view A). URL: https://confluence.desy.de/display/BI/Public+GeneralSlides?preview=/35009238/58917569/BellellTopview415R12.pdf. [Accessed on: 2023-05-22].

[23] L. Zani et al. The silicon vertex detector of the belle II experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1038:166952, sep 2022. doi: 10.1016/j.nima.2022.166952. URL: https://doi.org/10.1016%2Fj.nima.2022.166952.

[24] VXD images [Belle II Internal page]. URL: https://confluence.desy.de/display/BI/VXD+Images. [Accessed on: 2023-05-25].

[25] K. Adamczyk et al. The Design, Construction, Operation and Performance of the Belle II Silicon Vertex Detector. 1 2022.

[26] Henrikas Svidras. 5th Belle II Starter Kit Workshop – The Central Drift Chamber of Belle 2. URL: https://indico.belle2.org/event/1307/sessions/378/attachments/3070/5652/CDC_B2SK.pdf. [Accessed on: 2023-05-25].

[27] Kenkichi Miyabayashi. Belle electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 494(1):298–302, 2002. ISSN 0168-9002. doi: https://doi.org/10.1016/S0168-9002(02)01483-3. URL: https://www.sciencedirect.com/science/article/pii/S0168900202014833. Proceedings of the 8th International Conference on Instrumentatio n for Colliding Beam Physics.

[28] S. Iwata, I. Adachi, K. Hara, T. Iijima, H. Ikeda, H. Kakuno, H. Kawai, T. Kawasaki, S. Korpar, P. Križan, T. Kumita, S. Nishida, S. Ogawa, R. Pestotnik, L. Å antelj, A. Seljak, T. Sumiyoshi, M. Tabata, E. Tahirovic, and Y. Yusa. Particle identification performance of the prototype aerogel RICH counter for the Belle II experiment. *Progress of Theoretical and Experimental Physics*, 2016(3), 03 2016. ISSN 2050-3911. doi: 10.1093/ptep/ptw005. URL: https://doi.org/10.1093/ptep/ptw005.033H01.

[29] M. Yonenaga et al. Performance evaluation of the aerogel RICH counter for the Belle II spectrometer using early beam collision data. *PTEP*, 2020(9):093H01, 2020. doi: 10.1093/ptep/ptaa090.

[30] Martin Bessner. Performance of the Belle II imaging Time-Of-Propagation (iTOP) detector in first collisions. *Nucl. Instrum. Meth. A*, 958:162318, 2020. doi: 10.1016/j.nima.2019.06.059.

[31] T. Kuhr, C. Pulvermacher, M. Ritter, T. Hauth, and N. Braun. The Belle II Core Software. *Comput. Softw. Big Sci.*, 3(1):1, 2019. doi: 10.1007/s41781-018-0017-9.

[32] Mikihiko Nakao et al. Performance of the Unified Readout System of Belle II. *IEEE Trans. Nucl. Sci.*, 68(8):1826–1832, 2021. doi: 10.1109/TNS.2021.3084826.

[33] Belle II Luminosity. URL: `https://confluence.desy.de/display/BI/Belle+II+Luminosity`. [Accessed on: 2023-05-29].

[34] J. F. Krohn et al. Global decay chain vertex fitting at Belle II. *Nucl. Instrum. Meth. A*, 976:164269, 2020. doi: 10.1016/j.nima.2020.164269.

[35] J. Tanaka. Belle Note 194 - Kinematic Fitting, 2000.

[36] H. Tajima et al. Proper time resolution function for measurement of time evolution of B mesons at the KEK B factory. *Nucl. Instrum. Meth. A*, 533:370–386, 2004. doi: 10.1016/j.nima.2004.07.199.

[37] S. Dey and A. Soffer. Beam-Constrained Vertexing for B Physics at the Belle II Experiment. *Springer Proc. Phys.*, 248:411–415, 2020. doi: 10.1007/978-981-15-6292-1_52.

[38] F. Abudinén et. al. B-flavor tagging at belle II. *The European Physical Journal C*, 82(4), apr 2022. doi: 10.1140/epjc/s10052-022-10180-9. URL: `https://doi.org/10.1140%2Fepjc%2Fs10052-022-10180-9`.

[39] Thibaud Humair, Rishabh Mehta, and Petros Stavroulakis. Belle II Note: Calibration of the decay-time resolution and flavor tagger using hadronic B decays reconstructed in Moriond 2023 data. Jan 2023. URL: `https://docs.belle2.org/record/3385`.

[40] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001. doi: 10.1016/S0168-9002(01)00089-4.

[41] Anders Ryd, David J. Lange, Natalia Kuznetsova, Alto Ishikawa, David Kirkby, Marcello Rotondo, Frank Wuerthwein, and S. Henrot Versillé. Evtgen: A monte carlo generator for b-physics. 2005.

[42] S. Jadach, B. F. L. Ward, and Z. Was. The Precision Monte Carlo event generator K K for two fermion final states in e+ e- collisions. *Comput. Phys. Commun.*, 130:260–325, 2000. doi: 10.1016/S0010-4655(00)00048-5.

[43] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015. doi: 10.1016/j.cpc.2015.01.024.

[44] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl.Instrum.Meth.*, A506:250–303, 2003. doi: 10.1016/S0168-9002(03)01368-8.

[45] Hikaru Tanigawa. *Measurement of time-dependent CP asymmetry in $B^0 \to K^0_S K^0_S K^0_S$ decays at the Belle II experiment.* PhD thesis, The University of Tokyo, 2022.

[46] Tobias Schlüter and Markus Prim. Belle II Note: The Belle II V0 Finder. Oct 2015. URL: https://docs.belle2.org/record/307.

[47] Kun Wan. *Time Dependent Charge-Parity Violation in $B^0 \to K^0_S K^0_S K^0_S$ in Belle II early operation.* PhD thesis, The University of Tokyo, 2021.

[48] Geoffrey C. Fox and Stephen Wolfram. Event shapes in e+e — annihilation. *Nuclear Physics B*, 149(3):413–496, 1979. ISSN 0550-3213. doi: https://doi.org/10.1016/0550-3213(79)90003-8. URL: https://www.sciencedirect.com/science/article/pii/0550321379900038.

[49] D. M. et al. Asner. Search for exclusive charmless hadronic b decays. *Phys. Rev. D*, 53:1039–1050, Feb 1996. doi: 10.1103/PhysRevD.53.1039. URL: https://link.aps.org/doi/10.1103/PhysRevD.53.1039.

[50] Thomas Keck. FastBDT: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification. 2016. doi: 10.48550/ARXIV.1609.06119. URL: https://arxiv.org/abs/1609.06119.

[51] Giovanni Punzi. Comments on likelihood fits with variable resolution, 2003.

[52] Wouter Verkerke and David P. Kirkby. The RooFit toolkit for data modeling. *eConf*, C0303241:MOLT007, 2003.

[53] Rene Brun and Fons Rademakers. ROOT an object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(1):81–86, 1997. ISSN 0168-9002. doi: https://doi.org/10.1016/S0168-9002(97)00048-X. URL: https://www.sciencedirect.com/science/article/pii/S016890029700048X. New Computing Techniques in Physics Research V.

[54] F. James and M. Roos. Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations. *Comput. Phys. Commun.*, 10:343–367, 1975. doi: 10.1016/0010-4655(75)90039-9.

[55] Muriel Pivk and Francois R. Le Diberder. SPlot: A Statistical tool to unfold data distributions. *Nucl. Instrum. Meth. A*, 555:356–369, 2005. doi: 10.1016/j.nima.2005.08.106.

[56] Owen Long, Max Baak, Robert N. Cahn, and David P. Kirkby. Impact of tag side interference on time dependent CP asymmetry measurements using coherent B0 anti-B0 pairs. *Phys. Rev. D*, 68:034010, 2003. doi: 10.1103/PhysRevD.68.034010.

[57] Y. Yusa, K. Miyabayashi, and ICPV group. Belle Note 1149 - Measurement of time-dependent CP violation in $B^0 \to (c\bar{c})K^0$ decays with 772 M $B\overline{B}$, 2011.

[58] Gary J. Feldman and Robert D. Cousins. A Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D*, 57:3873–3889, 1998. doi: 10.1103/PhysRevD.57.3873.

[59] A. J. Bevan et al. The physics of the b factories. *The European Physical Journal C*, 74(11), nov 2014. doi: 10.1140/epjc/s10052-014-3026-9. URL: `https://doi.org/10.1140%2Fepjc%2Fs10052-014-3026-9`.

[60] A.Selce T.Koga S.Stengel. Belle II Note: Optimization of pi0 reconstruction selection and first systematic uncertainty evaluation of the efficiencies, Feb 2020. URL: `https://docs.belle2.org/record/1823`.

[61] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

Measurement of Time-Dependent CP Violation in Decays of $B^0 \to \eta' K^0_S$ at the Belle II Experiment

裴　漢　郁