Study of the $X(3872)$ Invariant Mass Distribution Using the $B \to (X(3872) \to D^0\overline{D}^{*0})K$ Decay at the Belle Experiment

Belle実験における $B \to (X(3872) \to D^0\overline{D}^{*0})K$ 崩壊を用いた
$X(3872)$ 不変質量分布の研究

Hikari Hirata

High Energy Physics Laboratory

Graduate School of Science, Nagoya University

February 26, 2023

Abstract

The charmonium-like $X(3872)$ state was discovered as a narrow peak in the vicinity of the $D^0\overline{D}^{*0}$ threshold on the $J/\psi\pi^+\pi^-$ invariant mass distribution in exclusive $B^+ \to J/\psi\pi^+\pi^-K^+$ decays at the Belle experiment. Various interpretations have been proposed, such as a loosely bound $D\overline{D}^*$ molecule state, an admixture of a molecular state and a charmonium, and a cusp at the $D^0\overline{D}^{*0}$ threshold. The structure of the state remains uncertain.

To reveal the structure, the $X(3872)$ lineshape, which reflects the information of the structure, is examined. In this study, we use $B \to (X(3872) \to D^0\overline{D}^{*0})K$ decays in a data sample of 772×10^6 $B\overline{B}$ pairs collected with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. An important performance for this analysis is mass resolution, and this $X(3872) \to D^0\overline{D}^{*0}$ decay mode is superior in it. The lineshape is evaluated using the Flatté lineshape. This lineshape model can describe the lineshape distortion due to the $D\overline{D}^*$ coupled-channel effect, which cannot be represented by the Breit-Wigner lineshape commonly used for normal hadrons. Compared with the previous analysis of the Flatté lineshape using $X(3872) \to J/\psi\pi^+\pi^-$ decays at the LHCb experiment, this study is inferior in data size; however, it has two advantages. The first is the aforementioned better resolution. The second is including the information on the $X(3872) \to D^0\overline{D}^{*0}$ branching fraction, which is important for determining the coupling strength to the $D\overline{D}^*$ channel. Therefore, we aim to provide more information on the lineshape, and in particular on the coupling strength of $X(3872) \to D^0\overline{D}^{*0}$ by analyzing the $D^0\overline{D}^{*0}$ decay.

First, we evaluate with the relativistic Breit-Wigner lineshape to ensure this analysis method. The branching fraction is found to be

$$\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}) = (0.97^{+0.21}_{-0.18}(\text{stat}) \pm 0.10(\text{syst})) \times 10^{-4}.$$

The signal from B^0 decays is firstly observed with 5.2σ significance. The relative branching fraction between $B^0 \to X(3872)K^0$ and $B^+ \to X(3872)K^+$ is measured to be

$$\mathcal{B}(B^0 \to X(3872)K^0)/\mathcal{B}(B^+ \to X(3872)K^+) = 1.34^{+0.47}_{-0.40}(\text{stat})^{+0.10}_{-0.12}(\text{syst}).$$

The mass and width parameters are determined to be

$$m = 3873.71^{+0.56}_{-0.50}(\text{stat}) \pm 0.13(\text{syst})\ \text{MeV}/c^2, \quad \Gamma_0 = 5.2^{+2.2}_{-1.5}(\text{stat}) \pm 0.4(\text{syst})\ \text{MeV}.$$

In this study, the precision of the lineshape measurement is relatively improved by at least 22% compared with the previous studies using $D^0\overline{D}^{*0}$ decays. The improvement is contributed by including additional D^0 decay modes used in the reconstruction and studying detailed mass dependence of the detector response.

Second, we establish a method to measure the undetermined parameter in the previous study at LHCb, the $D\overline{D}^*$ coupling constant g. It is measured to be

$$g > 0.075 \text{ at } 95\% \text{ credibility}.$$

Because the lower limit is more stringent than the previous study at LHCb, it suggests that analysis using the $D^0\overline{D}^{*0}$ decay mode can indeed complement the study of the $J/\psi\pi^+\pi^-$ decay mode in this framework. The limit shows that the coupling strength to the $D\overline{D}^*$ channel is not small.

For future improvements, because uncertainty due to the limited statistics is dominant in all of the results, it is important to use higher statistical data, e.g. the data at the

successor experiment, Belle II. For the measurement of the Flatté lineshape, the sensitivity can be improved further by performing a simultaneous fit between the $D^0\overline{D}^{*0}$ and $J/\psi\pi^+\pi^-$ decays. Such an analysis could fully determine the lineshape in the coupled-channel framework, and greatly contribute to determining the internal structure.

Contents

Chapter 1

Introduction

The matter has a hierarchical structure. For example, the molecules forming matter are composed of atoms, and an atom is composed of electrons and a nucleus, which is a bound system of protons and neutrons. The particles forming the proton and neutron are quarks, currently known as the most fundamental particles. In our nature, quarks are not observed alone, and thus they are confined to composite particles called hadrons. Hadrons observed to date can be well classified into mesons and baryons by the number of quarks in the hadron. For example, a meson and a baryon are respectively composed of $q\overline{q}$ and qqq, where q represents a quark and $\overline{q}$ represents an antiquark. On the other hand, the quantum chromodynamics (QCD), the fundamental theory of the strong interactions between quarks, allows structures other than the normal mesons and baryons. An example is multiquark states, composed of four or more quarks. For many years, such "exotic" hadrons have been explored.

Since the search for hadrons containing heavy quarks began, the candidates of exotic hadrons have been discovered. They are called "X, Y, Z, and Pc" and are found to have quantum numbers that the normal mesons and baryons do not have. These states are expected to be hadronic molecular states, multiquark states, virtual states, and so on. However, their true nature is still uncertain.

I focus on a candidate of exotic hadrons $X(3872)$, discovered in the vicinity of the $D^0\overline{D}^{*0}$ threshold at Belle [1], and aim to elucidate its internal structure. The shape of the invariant mass distribution of the $X(3872)$ signal, hereafter referred to as *lineshape*, is one of the key properties to achieve that goal. In this study, two lineshape measurements in the $D^0\overline{D}^{*0}$ decay are performed using an analysis method improved from the previous studies at Belle and *BABAR* [2, 3]. The first is the relativistic Breit-Wigner lineshape, commonly used for a resonance state. Through the evaluation with this model, we ensure the analysis method and examine the tendency to yield the higher mass, the larger width, and the higher relative branching fraction between $B^0 \to X(3872)K^0$ and $B^+ \to X(3872)K^+$ in the $X(3872) \to D^0\overline{D}^{*0}$ decay mode only. The second is the Flatté lineshape, a more sophisticated model considering the coupled-channel effect. It can describe both a resonance state, a bound state, and a virtual state. However, due to the model property, it is difficult to determine all lineshape parameters with the analysis in the $J/\psi\pi^+\pi^-$ decay, as reported in the previous study at LHCb [4]. For a countermeasure, the analysis in the $D^0\overline{D}^{*0}$ decay is essential. When all parameters are completely determined, it provides us with the pole location of the scattering amplitude, which is one of the pieces of information needed to classify hadrons.

Therefore, in this thesis, a study of the $X(3872)$ lineshape in the $B \to X(3872)K \to$

$D^0\overline{D}^{*0}K$ decay is described. This analysis is based on a data sample of 772×10^6 $B\overline{B}$ pairs collected at the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. The thesis is organized as follows. Chapter 2 describes the motivation for this study, focusing on our target of an exotic hadron candidate $X(3872)$. Chapter 3 provides an overview of the Belle experiment, where the used data were collected. In Chapter 4, the reconstruction and selection for the signal events are described. In Chapter 5, the fitting procedure and its validation are presented. The results of fitting the data and systematic uncertainties are summarized in Chapter 6. The discussion based on the results and the prospects are given in Chapter 7. Finally, chapter 8 contains the conclusions of the thesis.

Chapter 2

Candidate of Exotic States $X(3872)$

My research theme is a study of the $X(3872)$ lineshape using $X(3872) \to D^0\overline{D}^{*0}$ candidates produced from exclusive decays $B \to D^0\overline{D}^{*0}K$. This chapter describes the implications of studying hadron physics with a focus on exotic hadrons, the background of the target hadron $X(3872)$, and the research issues.

2.1 Hadrons and Quarks

Several hundred hadrons have been observed to date. In the quark model, hadrons can be well classified by considering hadrons to be composed of fermions with a spin 1/2 called quarks [5, 6]. Hadrons composed of a quark and an antiquark are mesons, and those composed of three quarks are baryons. To date, six flavors of quarks have been observed: down (d), up (u), strange (s), charm (c), bottom (b), and top (t), as shown in Table 2.1. Each quark has quantum numbers related to its flavor, e.g., z-component of isospin (I_z), strangeness (S), charm (C), bottomness (B), topness (T). The following relationship holds for electric charge (Q/e), flavor quantum numbers, and baryon number ($\mathcal{B}$), which is known as the Gell-Mann-Nishijima formula [7]

$$\frac{Q}{e} = I_z + \frac{1}{2} \cdot (\mathcal{B} + S + C + B + T). \tag{2.1}$$

In QCD, a new degree of freedom of color is introduced for quarks, and it is described by a field theory based on $SU(3)$ gauge symmetry of three colors, e.g., red, green, and

Table 2.1: Quarks and their quantum number.

	d	u	s	c	b	t
Electric charge (Q/e)	$-\frac{1}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$
Isospin (I)	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	0
Isospin z-component (I_z)	$\frac{1}{2}$	$-\frac{1}{2}$	0	0	0	0
Strangeness (S)	0	0	-1	0	0	0
Charm (C)	0	0	0	1	0	0
Bottomness (B)	0	0	0	0	-1	0
Topness (T)	0	0	0	0	0	1
Baryon number ($\mathcal{B}$)	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

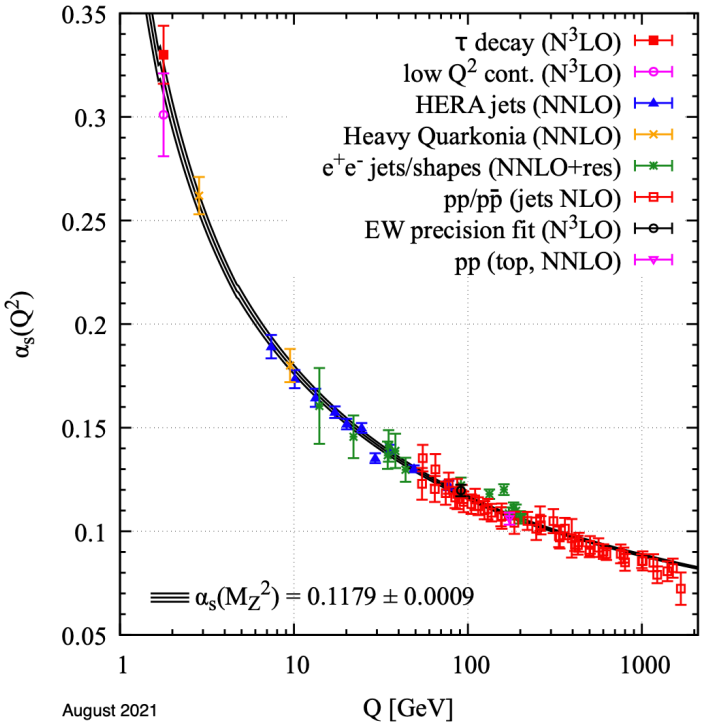


Figure 2.1: Summary of measurements of the coupling constant of the strong interaction α_s as a function of the energy scale Q [18].

blue. The strong interaction is mediated by a massless particle called a gluon. One of the QCD characteristics is the asymptotic freedom, where the coupling constant of the strong interaction is very small in a high-energy region. It was confirmed from both the theoretical and experimental sides, as shown in the relation between the coupling constant and the energy scale (Fig. 2.1). On the other hand, the coupling constant becomes large in the low energy region, and the phenomena are non-perturbative. One of the most important phenomena is the quark confinement. In other words, quarks have never been observed alone.

Hadrons are observed as only color-charge-neutral states, i.e., white. The fact predicts the existence of a tetraquark ($qq\overline{q}\overline{q}$) [8–10], a pentaquark ($qqqq\overline{q}$) [11], a H-dibaryon ($uuddss$) [12] a hadronic molecule [13], a glueball (gg) [14–16], and a $q\overline{q}$-pair with an excited gluon (a hybrid, $q\overline{q}g$) [17], where q ($\overline{q}$) denotes a quark (anti-quark) and g denotes a valance gluon. Exploring these states and investigating their internal structure must answer the question of how valance quarks and gluons are combined to form hadrons.

2.2 Charmonium Spectroscopy

Particles composed of $c\overline{c}$ are collectively known as charmonia. The constituent quark mass of the c quark is large, about 1500 MeV[*1], and thus the quarks behave non-relativistically. Therefore, the mass spectrum can be predicted by solving the Schrödinger equation with the Coulombic potential and the linear potential as a bound state of a two-body system [19, 20]. The expected spectrum is shown in Fig. 2.2 (dashed lines). Below the $D\overline{D}$

[*1] Note that it is different from the current mass in QCD.

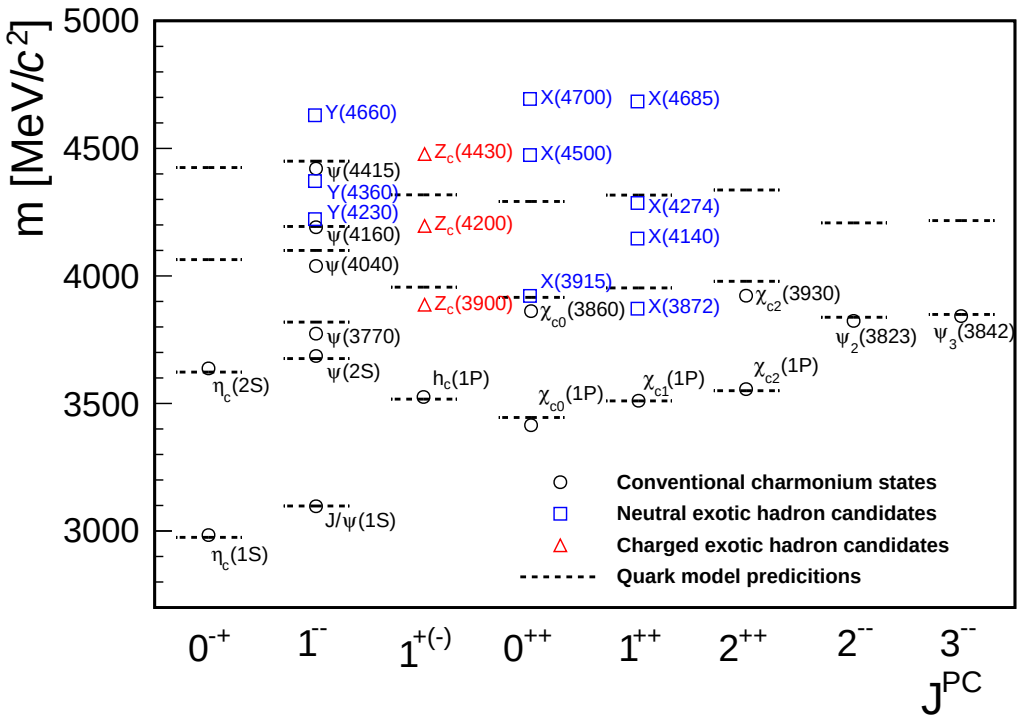


Figure 2.2: The current status of the charmonium-like spectrum. The dashed black lines indicate the theoretical predictions based on the Godfrey-Isgur relativized potential model [20]. The open circles are the observed conventional charmonia. The open squares and open triangles are candidates of neutral and charged exotic hadrons, respectively. Here, masses and spin-parity (J^{PC}) quantum numbers of experimentally observed states are taken from Ref. [18].

threshold of about 3740 MeV, the predictions match the observed states (open circles) well. On the other hand, many unexpected states named XYZ[*2] have been found above the threshold. They are exotic hadron candidates because they have properties that normal charmonium should not have, e.g., the non-zero charge and the violation of isospin. Although these hadron properties have been well investigated, there is no conclusive evidence to determine their internal structure. In this study, I focus on $X(3872)$ among the exotic hadron candidates because its data size are relatively rich. The other states are described in detail elsewhere [22, 23].

2.3 Charmonium-like $X(3872)$ State

The charmonium-like $X(3872)$ state, also known as $\chi_{c1}(3872)$, was discovered by the Belle experiment as a narrow peak in the vicinity of the $D^0\overline{D}^{*0}$ threshold on the $J/\psi\pi^+\pi^-$ invariant mass distribution in exclusive $B^+ \to J/\psi\pi^+\pi^-K^+$ decays [1]. Its existence has been confirmed by multiple experiments: D0 [24], *BABAR* [25], CDF [26], LHCb [27], and BESIII [28]. In addition to the $J/\psi\pi^+\pi^-$ decay, other decays such as $J/\psi\omega$, $J/\psi\gamma$,

[*2]The new naming scheme was proposed by the LHCb collaboration [21], and it is currently being discussed in experiments other than the LHCb experiment. This thesis follows the conventional naming scheme.

$\psi(2S)\gamma$, $D^0\overline{D}^{*0}$, $D^0\overline{D}^0\pi^0$, and $\pi^0\chi_{c0}$ have been observed [18]. The $X(3872)$ peak has been analyzed with the Breit-Wigner lineshape. The mass is 3871.65 ± 0.06 MeV/c^2, and the width is 1.19 ± 0.21 MeV on the world average based on the analyses of the decays including J/ψ [18]. The quantum number J^{PC} was determined to be 1^{++} [29, 30]. Based on the experimental results, it seems to be a candidate of exotic hadrons rather than a pure charmonium. There are three reasons.

- Mass inconsistent with the quark model predictions:
  Comparing the measured mass and J^{PC} with the quark model predictions (Fig. 2.2), this state is close to the $\chi_{c1}(2P)$ state, but the mass is inconsistent with the prediction.

- Narrow width:
  Since this state has been found to a decay via a strong interaction, $D^0\overline{D}^{*0}$ decay, the width is expected to be much larger than the measured width of the $\chi_{c1}(1P)$ state, 0.84 ± 0.04 MeV [18], if it is a normal charmonium. However, the measured $X(3872)$ width is only slightly larger.

- Isospin violation:
  For normal charmonia, the isospin should be zero. For $X(3872)$, the $\pi^+\pi^-$ invariant mass distribution of the $X(3872) \to J/\psi\pi^+\pi^-$ decay is consistent with a prediction for the $\rho \to \pi^+\pi^-$ decay with a non-zero isospin, $I = 1$ [31, 32]. In addition, the decay mode with the final state of $I = 0$, $X(3872) \to J/\psi\omega$, was discovered, too [33–35]. This fact concludes that $X(3872)$ has components other than $c\bar{c}$ and large isospin mixture.

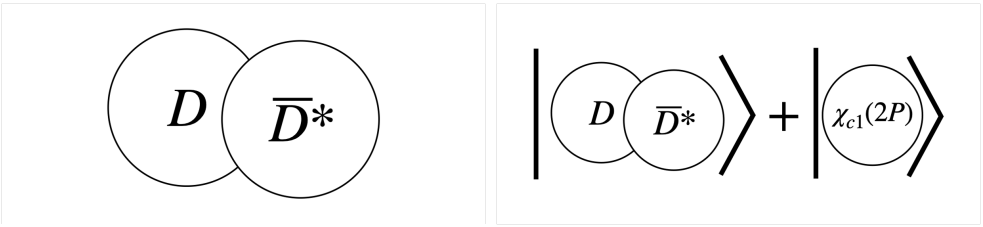
Because the observed mass coincides with the $D^0\overline{D}^{*0}$ threshold of 3871.69 ± 0.10 MeV/c^2 and the J^{PC} state can couple to the $D^0\overline{D}^{*0}$ channel in S-wave, some $D\overline{D}^*$ component contribution is expected in the $X(3872)$ nature. Various interpretations such as a loosely bound state of $D\overline{D}^*$ [36–39], an admixture of a molecular state and a pure charmonium [40], and a cusp at the $D^0\overline{D}^{*0}$ threshold [41–43] have been proposed (Fig. 2.3), and the structure of the state remains uncertain. Measurement of the lineshape in various decay modes helps to discriminate between different options for the structure, because the lineshape reflects poles of the scattering amplitude corresponding to the complex energy eigenstates [44].

2.4 Lineshape Models of $X(3872)$

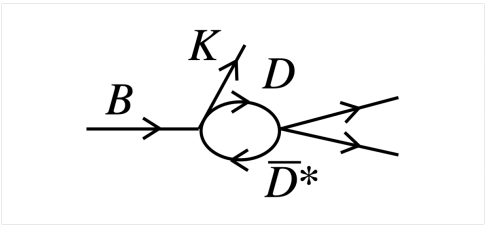
In this study, issues regarding studies of two lineshapes using $D^0\overline{D}^{*0}$ decay are investigated: the Breit-Wigner lineshape and the Flatté-inspired parametrization. This section summarizes the definition, important properties, and a current issue of each line-shape model.

2.4.1 Relativistic Breit-Wigner Lineshape

The differential branching fraction as a function of the observed invariant mass has been analyzed with the Breit-Wigner lineshape, commonly used for a resonance state.



(a) The loosely bound state of $D\overline{D}^*$ (b) The admixture of a molecular state and a pure charmonium



(c) The cusp at the $D^0\overline{D}^{*0}$ threshold

Figure 2.3: Schematic diagrams of the $X(3872)$ interpretations.

Definition

The relativistic Breit-Wigner lineshape is defined as follows [18]:

$$f_{\mathrm{BW}}(M) = \frac{mM\Gamma(M)}{(M^2 - m^2)^2 + m^2\Gamma(M)^2}, \tag{2.2}$$

where M is the observed invariant mass, and m is the mass of the resonance. The mass-dependent width $\Gamma(M)$ is defined as

$$\Gamma(M) = \Gamma_0 \frac{m}{M} \left(\frac{p(M)}{p(m)} \right)^{2L+1}, \tag{2.3}$$

where Γ_0 and L are the width of resonance and the orbital momentum, respectively. Taking account of the closeness to the threshold, the decay is assumed to be S-wave ($L = 0$) with no D-wave ($L = 2$) admixture. The momentum of one of the daughters in the rest frame of $X(3872)$, $p(M)$, can be calculated as

$$p(M) = \sqrt{\frac{(M^2 - (m_{D^0} + m_{D^{*0}})^2 (M^2 - (m_{D^0} - m_{D^{*0}})^2}{4M^2}}, \tag{2.4}$$

where m_{D^0} and $m_{D^{*0}}$ are the nominal masses of D^0 and D^{*0} (Table 2.2), respectively [18].

Relation between Lineshape and Parameters

In general, m and Γ_0 are consistent with a peak and a full width at half maximum (FWHM), respectively. If m is close to the threshold, however, the denominator of Eq. (2.2) increases monotonically with M. This often makes a dump near the threshold, i.e., m and Γ_0 are not always consistent with the peak position and the FWHM, respectively.

Table 2.2: Summary of the mass and width values in the lineshape definitions.

Parameter	Value
m_{D^0}	1864.84 ± 0.05 MeV/c^2
$m_{D^{*0}}$	2006.85 ± 0.05 MeV/c^2
m_{D^+}	1869.66 ± 0.05 MeV/c^2
$m_{D^{*-}}$	2010.26 ± 0.05 MeV/c^2
$m_{J/\psi}$	3096.900 ± 0.006 MeV/c^2
m_ρ	775.26 ± 0.34 MeV/c^2
Γ_ρ	149.1 ± 0.8 MeV
m_ω	782.66 ± 0.13 MeV/c^2
Γ_ω	8.68 ± 0.13 MeV

Table 2.3: Results of previous studies.

	BABAR 2008 [3]	Belle 2010 [2]
Amount of data [fb^{-1}]	347	605
m [MeV/c^2]	$3875.1^{+0.7}_{-0.5} \pm 0.5$	$3872.9^{+0.6+0.4}_{-0.4-0.5}$
Γ_0 [MeV]	$3.0^{+1.9}_{-1.4} \pm 0.9$	$3.9^{+2.8+0.2}_{-1.4-1.1}$
$\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0})$ [10^{-4}]	$1.67 \pm 0.36 \pm 0.47$	$0.77 \pm 0.16 \pm 0.10$
$\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0})$ [10^{-4}]	$2.22 \pm 1.05 \pm 0.42$	$0.97 \pm 0.46 \pm 0.13$
$\mathcal{B}(B^0 \to X(3872)K^0)/\mathcal{B}(B^+ \to X(3872)K^+)$	$1.33 \pm 0.69 \pm 0.43$	$1.26 \pm 0.65 \pm 0.06$

Issue of Inconsistent Results Depending on the Decay Mode

Based on the analyses of the decays including J/ψ, the mass is 3871.65 ± 0.06 MeV/c^2, and the width is 1.19 ± 0.21 MeV on the world average [18]. Analyses of the decay to $D^0\overline{D}^{*0}$ based on the Breit-Wigner lineshape tend to yield a higher mass and a larger width, with the width measurement subject to large uncertainties, as shown in Table 2.3 [2, 3]. Therefore, a more precise lineshape measurement in the $D^0\overline{D}^{*0}$ decay is needed.

In addition, although the relative branching fraction between $B^0 \to X(3872)K^0$ and $B^+ \to X(3872)K^+$ is expected to be independent of the $X(3872)$ decay mode, that in the $D^0\overline{D}^{*0}$ decay tended to be shifted by $+1.2\sigma$ from the average of the measurements in the $J/\psi\pi^+\pi^-$ decay at Belle and *BABAR*, 0.48 ± 0.13 [45]. To clarify the disagreement, we need to observe a significant signal from $B^0 \to X(3872)K^0$.

2.4.2 Flatté Lineshape

Discrepancies in the lineshape between the decays to the $J/\psi\pi^+\pi^-$ and $D^0\overline{D}^{*0}$ final state can arise near the threshold due to coupling to the $D\overline{D}^*$ channel, so-called coupled-channel effects. One of the models to account for coupled-channel effects is the Flatté-inspired parametrization[*3], a Breit-Wigner model with an explicit expression for the energy-dependent partial width. Furthermore, this model has the advantage of describing both resonant states, bound states, and virtual states.

[*3]The Flatté model was originally introduced to describe the $\pi\eta$ and $K\overline{K}$ invariant mass distributions near the $K\overline{K}$ threshold for the scalar-isovector meson $a_0(980)$ [46]. Experimentally, it is used to represent lineshapes of the near-threshold states, such as the light scalar mesons of $a_0(980)$ [47] and $f_0(980)$ [48–50], and the nucleon resonance $N(1535)$ [51]. An extension of this model for $X(3872)$ is proposed in Refs [41, 52].

Definition

The Flatté-inspired parametrization is defined as follows using the energy from the $D^0\overline{D}^{*0}$ threshold, $E = M - (m_{D^0} + m_{D^{*0}})$ [41, 52]:

$$f_{\text{Flatte}}(E) = \frac{g k_{D^0\overline{D}^{*0}}}{|D(E)|^2}, \tag{2.5}$$

$$D(E) = \begin{cases} E - E_f - \frac{1}{2} g \kappa_{D^+D^{*-}} + \frac{i}{2}[g k_{D^0\overline{D}^{*0}} + \Gamma(E)] & \text{for } 0 < E < \delta, \\ E - E_f + \frac{i}{2}[g(k_{D^0\overline{D}^{*0}} + k_{D^+D^{*-}}) + \Gamma(E)] & \text{for } E > \delta, \end{cases} \tag{2.6}$$

where $E_f = m_0 - (m_{D^0} + m_{D^{*0}})$ is the mass difference of this state (m_0) from the threshold, and g is a coupling constant to the $D\overline{D}^*$ channels[*4]; we assume the coupling constants for the $D^0\overline{D}^{*0}$ and D^+D^{*-} channels are the same due to isospin symmetry. The momenta k_a and κ_a for the channel a are measured in the rest frame of the $X(3872)$. They are expressed using the reduced mass μ as

$$\begin{aligned} k_{D^0\overline{D}^{*0}} &= \sqrt{2\mu_{D^0\overline{D}^{*0}} E}, \\ k_{D^+D^{*-}} &= \sqrt{2\mu_{D^+D^{*-}}(E - \delta)}, \\ \kappa_{D^+D^{*-}} &= \sqrt{2\mu_{D^+D^{*-}}(\delta - E)}, \\ \delta &= (m_{D^+} + m_{D^{*-}}) - (m_{D^0} + m_{D^{*0}}). \end{aligned} \tag{2.7}$$

The energy-dependent width $\Gamma(E)$ is defined by

$$\Gamma(E) = \Gamma_{J/\psi\rho}(E) + \Gamma_{J/\psi\omega}(E) + \Gamma_0, \tag{2.8}$$

where Γ is the partial width of the channel indicated by the subscript. For the $J/\psi\rho$ and $J/\psi\omega$ channels, the dependence on E is defined using the phase space and an effective coupling constant, f_ρ or f_ω:

$$\Gamma_{J/\psi\rho}(E) = f_\rho \int_{2m_\pi}^{M(E)-m_{J/\psi}} \frac{dm'}{2\pi} \frac{q(m', E)\Gamma_\rho}{(m' - m_\rho)^2 + \Gamma_\rho^2/4}, \tag{2.9}$$

$$\Gamma_{J/\psi\omega}(E) = f_\omega \int_{3m_\pi}^{M(E)-m_{J/\psi}} \frac{dm'}{2\pi} \frac{q(m', E)\Gamma_\omega}{(m' - m_\omega)^2 + \Gamma_\omega^2/4}, \tag{2.10}$$

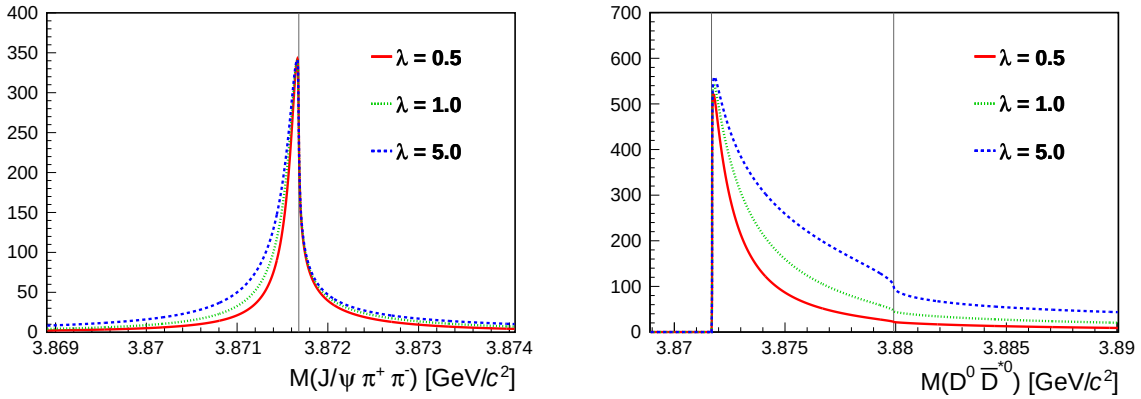
where Γ_ρ and Γ_ω are the total widths for the ρ and ω resonances, respectively. The mass and width values in the definition are summarized in Table 2.2. The upper limit of the integral is set by the difference between

$$M(E) = E + (m_{D^0} + m_{D^{*0}}) \tag{2.11}$$

and $m_{J/\psi}$. In each case, $q(m', E)$ is the momentum of the two- or three-pion system in the rest frame of the $X(3872)$:

$$q(m', E) = \frac{1}{2M(E)} \sqrt{(M^2(E) - (m' + m_{J/\psi})^2)(M^2(E) - (m' - m_{J/\psi})^2)}. \tag{2.12}$$

[*4]The coupling constant of this Flatté model is dimensionless so that it differs from the conventional constant. This coupling constant g is related to the matrix element $\mathcal{M}$ by $g = \frac{|\mathcal{M}|^2}{8\pi m_0^2}$ according to the Fermi's golden rules, in which the partial width for $X(3872) \to D^0\overline{D}^{*0}$ ($\Gamma_{D^0\overline{D}^{*0}}$) is derived by $\Gamma_{D^0\overline{D}^{*0}} = \frac{k_{D^0\overline{D}^{*0}}}{8\pi m_0^2}|\mathcal{M}|^2 = g k_{D^0\overline{D}^{*0}}$.



(a) The $J/\psi\pi^+\pi^-$ invariant mass. (b) The $D^0\overline{D}^{*0}$ invariant mass.

Figure 2.4: The Flatte lineshape when all the free parameters variables (E_f, g, f_ρ, f_ω, and Γ_0) are multiplied by one constant, λ. The dotted green line is the lineshape measured at LHCb [4]. The solid red and dashed blue lines are cases of $\lambda = 0.5$ and $\lambda = 5.0$, respectively.

The parameter Γ_0 is the sum of the partial widths of other channels, such as radiative decays. In total, this model has five free parameters, E_f, g, f_ρ, f_ω, and Γ_0.

The lineshapes in other decays can be similarly described in the model. They are defined by rewriting the numerator of Eq. (2.5) to the partial width of the decay to draw. For example, in the case of $J/\psi\pi^+\pi^-$, it is defined by

$$f_{\text{Flatte}}(E) = \frac{\Gamma_{J/\psi\rho}(E)}{|D(E)|^2}. \tag{2.13}$$

Lineshape Features

In this model, a peak is easy to make in the vicinity of the $D^0\overline{D}^{*0}$ threshold on $M(J/\psi\pi^+\pi^-)$ despite the parameter relating to the mass E_f, because the derivative of the denominator $|D(E)|$ becomes discontinuous at the $D\overline{D}^*$ thresholds for $g > 0$. On the other hand, in the $D^0\overline{D}^{*0}$ channel, such a sharp peak is suppressed because of the phase space. A wider peak is predicted on $M(D^0\overline{D}^{*0})$. Thus, the model can explain the inconsistency of the decay width between the $J/\psi\pi^+\pi^-$ and $D^0\overline{D}^{*0}$ decays in the past measurements. The lineshape for each decay is shown in the dotted green lines of Fig. 2.4.

One of the important properties is that the area under the lineshape is proportional to the branching fraction. For example, it can be derived for the $J/\psi\pi^+\pi^-$ channel by

$$\begin{aligned}\int f_{\text{Flatte}}(M(J/\psi\pi^+\pi^-))dM(J/\psi\pi^+\pi^-) &\propto \int \frac{d\mathcal{B}(X(3872)\to J/\psi\pi^+\pi^-)}{dE}dE \\ &= \mathcal{B}(X(3872)\to J/\psi\pi^+\pi^-).\end{aligned} \tag{2.14}$$

Using this property, the lineshape parameters can be constrained. For example, f_ω can be constrained so that the branching fraction of the $J/\psi\pi^+\pi^-$ mode and that of the $J/\psi\omega$ mode are equal according to experimental results to date. Details are given in Sec. 5.4.

Relation with Internal Structures

In this sub-subsection, we show that the Flatté lineshape is related to the internal structures such as a resonance, a bound state of $D\overline{D}^*$, and a virtual state[*5].

Examples of the lineshape in the $J/\psi\pi^+\pi^-$ decay for different structures are shown in Fig. 2.5; here, g is set to zero for the resonance state, i.e., the lineshape is consistent with the Breit-Wigner lineshape, and the parameter sets 4 and 3 in Table 2 of Ref. [52] are used for the cusp and bound state, respectively. The figure indicates that the lineshape is distorted for the bound and virtual states compared to the resonance expressed in the Breit-Wigner lineshape. Thus, the lineshape is sensitive to the internal structure.

Strictly, we need to search poles on the scattering amplitude to identify the internal structure from the measured lineshape. The amplitude $F_{\rm Flatte}(E)$ is defined as

$$F_{\rm Flatte}(E) = -\frac{1}{2k_{D^0\overline{D}^{*0}}}\frac{gk_{D^0\overline{D}^{*0}}}{D(E)}. \tag{2.15}$$

The pole is given by solving $D(E) = 0$. The energy is a multivalued function because it is defined as $E = k^2_{D^0\overline{D}^{*0}}/(2\mu_{D^0\overline{D}^{*0}})$. Therefore, the pole position is represented by Riemann sheets; the basic theory on this subject is described in Refs [18, 53]. In this case, four Riemann sheets associated with the $D^0\overline{D}^{*0}$ channel are treated because the thresholds of the other channels, i.e., $J/\psi\pi^+\pi^-$, $J/\psi\pi^+\pi^-\pi^0$, and radiative decays, are much far from the $D^0\overline{D}^{*0}$ channel. According to Ref. [54], the Riemann sheet is labeled as follows:

Sheet I: $\mathrm{Im}k_{D^0\overline{D}^{*0}} > 0$ and $\mathrm{Im}k_{\rm other} > 0$;

Sheet II: $\mathrm{Im}k_{D^0\overline{D}^{*0}} > 0$ and $\mathrm{Im}k_{\rm other} < 0$;

Sheet III: $\mathrm{Im}k_{D^0\overline{D}^{*0}} < 0$ and $\mathrm{Im}k_{\rm other} < 0$; and

Sheet VI: $\mathrm{Im}k_{D^0\overline{D}^{*0}} < 0$ and $\mathrm{Im}k_{\rm other} > 0$.

Here, $k_{\rm other}$ represents the momentum of the channel with the next lower threshold than the $D^0\overline{D}^{*0}$ channel in this model. Sheets I and II (III and IV) correspond to a physical (unphysical) sheet for the $D^0\overline{D}^{*0}$ channel. Sheets I and II are labeled separately due to the presence of the other channels. The closeness to the real axis depends on the energy, as shown in Fig. 2.6. Above the $D^0\overline{D}^{*0}$ threshold, sheet III is close to the real axis, and between the $D^0\overline{D}^{*0}$ threshold and the lower threshold, sheet II is close to the axis. The relation between the pole position and the hadronic state is as follows:

Resonance: Pole on the sheet III above the $D^0\overline{D}^{*0}$ threshold;

Virtual state: Pole on sheet II above the $D^0\overline{D}^{*0}$ threshold,
on sheet III below the $D^0\overline{D}^{*0}$ channel, or on the sheet IV;

Bound state: Pole on the real axis of sheet I; and

Quasi-bound state: Pole on the sheet II below the $D^0\overline{D}^{*0}$ threshold.

Therefore, the Flatté lineshape measurement is directly related to examining the internal structure of $X(3872)$.

[*5] The threshold cusp effect is enhanced when the virtual state pole is near the threshold.

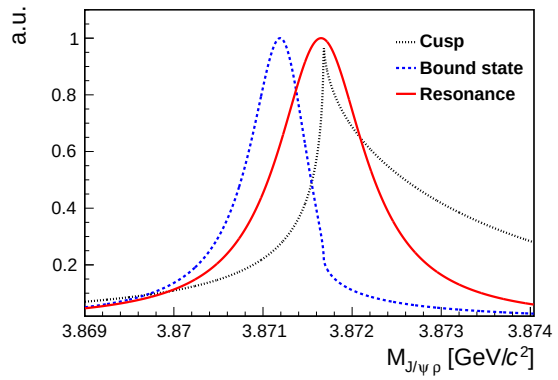


Figure 2.5: The Flatté lineshape of the $J/\psi\pi^+\pi^-$ decays for the cusp state (dotted), the bound state (dashed) and the resonance (solid).

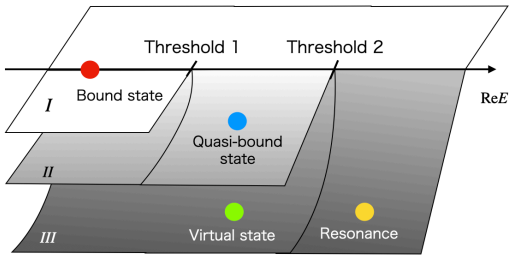


Figure 2.6: Riemann sheets close to the real axis and the classification of the pole for two channels.

Scaling Behavior Issue

This model has a scaling behavior that the lineshape near the threshold does not change when all parameters are scaled by a constant, λ [55]:

$$E_f \to \lambda E_f, \quad g \to \lambda g, \quad f_\rho \to \lambda f_\rho, \quad f_\omega \to \lambda f_\omega, \quad \Gamma_0 \to \lambda \Gamma_0. \tag{2.16}$$

Under the scaling transformation, the lineshape is transformed as follows for an arbitrary channel:

$$f_{\text{Flatte}}(E) = \frac{\Gamma_\forall}{|E/\lambda - E_f + \frac{i}{2}[g(k_{D^0\overline{D}^{*0}} + k_{D^+D^{*-}}) + \Gamma_{J/\psi\rho}(E) + \Gamma_{J/\psi\omega}(E) + \Gamma_0]|^2}, \tag{2.17}$$

where $\Gamma_\forall$ denotes the partial width for the arbitrary channel. This formula indicates that the lineshape is scale-invariant at the threshold, i.e., $E = 0$. Figure 2.4 (a) shows the Flatté lineshape of $J/\psi\pi^+\pi^-$ channel for $\lambda = 0.5$, 1.0 or 5.0. The narrow peak at the threshold is not sensitive to the scaling parameter λ. On the other hand, the lineshape is a little more sensitive to the transformation for the larger $|E|$. However, for $\lambda \gg 0$, the term E/λ becomes quite small, and the lineshape converges to

$$\lim_{\lambda\to\infty} f_{\text{Flatte}}(E) = \frac{\Gamma_\forall}{|-E_f + \frac{i}{2}[g(k_{D^0D^{*0}} + k_{D^+\overline{D}^{*-}}) + \Gamma_{J/\psi\rho}(E) + \Gamma_{J/\psi\omega}(E) + \Gamma_0]|^2}. \tag{2.18}$$

It causes a poor sensitivity to a large λ.

At LHCb, the analysis of the Flatté lineshape was performed using the $J/\psi\pi^+\pi^-$ decay [4]. Due to the scaling behavior issue and the poor mass resolution of 2.4–3.0 MeV compared to the lineshape, it was difficult to determine all of the parameters. For example, the mass difference from the threshold value was widely allowed to be -270 MeV $< E_f < -2.0$ MeV at 90% credibility. Instead, the ratio of a coupling constant to E_f was determined precisely,

$$\frac{dg}{dE_f} = (-15.11 \pm 0.16)\ \text{GeV}^{-1}. \tag{2.19}$$

Fixing E_f to -7.2 MeV/c^2, the other parameters were determined:

$$\begin{cases} g = 0.108 \pm 0.003^{+0.005}_{-0.006}, \\ f_\rho = (1.8 \pm 0.6^{+0.7}_{-0.6}) \times 10^{-3}, \\ \Gamma_0 = 1.4 \pm 0.4 \pm 0.6 \text{ MeV}, \end{cases} \tag{2.20}$$

where $E_f = -7.2$ MeV/c^2 is an assumption for the scattering amplitude analysis at LHCb.

As proposed in Ref. [52], there is possibility of pinning down the scaling behavior using the $D^0\overline{D}^{*0}$ decay. The reason is that the $D^0\overline{D}^{*0}$ decay causes the following two changes even under scaling transformations; (1) the lineshape changes at higher mass, and (2) the area associated with the $X(3872) \to D^0\overline{D}^{*0}$ branching fraction changes. In addition, the $D^0\overline{D}^{*0}$ decay has one more advantage of the good mass resolution, which is about 100 keV near the mass peak. Therefore, we attempt to experimentally pin down the scaling behavior observed at LHCb using the $D^0\overline{D}^{*0}$ decay in this study.

2.5 Previous Studies Using $X(3872) \to D^0\overline{D}^{*0}$ Decay and Improvement

There are three similar previous studies [2, 3, 56]. Reference [56] is an analysis of the $B \to D^0\overline{D}^0\pi^0 K$ decay at Belle, and Refs [2, 3] are analyses of the $B \to D^0\overline{D}^{*0}K$ decays at *BABAR* and Belle, respectively. A comparison of analysis methods with previous studies is summarized in Table 2.4. The latter two analyses apply a D^{*0} selection and a mass-constrained fit to the D^{*0} candidates. While this has the advantage of improving the signal-to-background ratio, it has the disadvantage of disallowing entries below the $D^0\overline{D}^{*0}$ threshold, which is important for the study of the structure. We also adopt this technique, given the limited size of our data sample. The disadvantage of requiring the D^{*0} is partially compensated for by analyzing the Flatté model, in which we can obtain a lineshape reflecting poles of the scattering amplitude.

The major updates from the previous measurements (Refs [2, 3]) are the following two items.

- Statistics; They are improved by not only using the higher statistical data at Belle but also adding more D^0 decay modes and loosening the requirement of the previous studies about the D^0 decay modes, in which at least one $D^0(\overline{D}^0)$ decays to $K^-\pi^+(K^+\pi^-)$.
- A broken-signal PDF; We additionally consider a lineshape for broken-signals depending on the assumed lineshape (Sec. 5.2). Especially in the previous study at Belle [2], they considered the broken-signal as an indistinguishable shape from the signal. Its systematic uncertainty was relatively large for measuring the parameters of the relativistic Breit-Wigner lineshape.

2.6 Objective of This Thesis

In this study, we measure the $X(3872)$ lineshape using a sample of $X(3872) \to D^0\overline{D}^{*0}$ candidates, produced in the exclusive decay $B \to D^0\overline{D}^{*0}K$ in the full Belle dataset. The following items are measured using the analysis procedure improved from the previous studies at Belle and *BABAR* [2, 3, 56].

Table 2.4: Comparison of analysis methods with previous studies.

	Belle 2006 [56]	*BABAR* 2008 [3]	Belle 2010 [2]	This work
Amount of data [fb^{-1}]	414	347	605	711
$X(3872)$ decay mode	$D^0\overline{D}^0\pi^0$	$D^0\overline{D}^{*0}$	$D^0\overline{D}^{*0}$	$D^0\overline{D}^{*0}$
The number of D^0 decay modes	4	3	5	6
	$\left(\begin{array}{l} K^-\pi^+ \\ K^-\pi^+\pi^-\pi^- \\ K^0_S\pi^+\pi^- \\ K^-K^+ \end{array}\right)$	$\left(\begin{array}{l} K^-\pi^+ \\ K^-\pi^+\pi^0 \\ K^-\pi^+\pi^-\pi^- \end{array}\right)$	$\left(\begin{array}{l} K^-\pi^+ \\ K^-\pi^+\pi^0 \\ K^-\pi^+\pi^-\pi^- \\ K^0_S\pi^+\pi^- \\ K^-K^+ \end{array}\right)$	$\left(\begin{array}{l} K^-\pi^+ \\ K^-\pi^+\pi^0 \\ K^-\pi^+\pi^-\pi^- \\ K^0_S\pi^+\pi^- \\ K^0_S\pi^0\pi^+\pi^- \\ K^-K^+ \end{array}\right)$
Requirements of D^0 mode	No	Yes	Yes	No
		$\left(\begin{array}{c}\text{At least one } D^0 \\ \text{decays to } K^-\pi^+\end{array}\right)$	$\left(\begin{array}{c}\text{At least one } D^0 \\ \text{decays to } K^-\pi^+\end{array}\right)$	
D^{*0} mass requirement	No	Yes	Yes	Yes
Mass constraint fit on D^{*0}	No	Yes	Yes	Yes
				(Mass difference)

- The relativistic Breit-Wigner lineshape: the mass and width parameters, the branching fraction of $\mathcal{B}(B \to KX(3872)) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0})$ and the relative branching fraction between $B^0 \to X(3872)K^0$ and $B^+ \to X(3872)K^+$.
- The Flatté lineshape: the parameter g, as the undetermined parameter at LHCb [4].

Throughout this thesis, charge conjugation is always included. We do not distinguish $D^0\overline{D}^{*0}$ from $\overline{D}^0D^{*0}$ unless otherwise indicated.

Chapter 3

Belle Experiment and Data Set

Belle is the B-factory experiment designed and optimized to test the Kobayashi–Maskawa mechanism for CP-violation in B-meson decays. It consists of an accelerator with low background via e^+e^- collisions, KEKB, and a multi-purpose 4π detector with high performance, Belle. These components enable us to address various physics projects, for example, not only B physics but also hadron physics, tau physics, two-photon physics and so on. The data containing a lot of B meson decays is the best experimental data for this study because it is capable of reconstructing photons and π^0 mesons, which are essential for the D^{*0} reconstruction in this study. This chapter describes overviews of the KEKB accelerator and the Belle detector.

3.1 KEKB Accelerator

KEKB is an asymmetric energy electron-positron collider at Tsukuba aiming at producing B meson pairs mainly [57, 58]. Its schematic view is shown in Fig. 3.1. It consists of a LINear ACcelerator (LINAC) and two main rings for electrons and positrons installed in a tunnel with a circumference of 3 km. Electrons and positrons are accelerated to 8.0 GeV and 3.5 GeV, respectively in LINAC, and then injected into the High Energy Ring (HER) for electrons and the Low Energy Ring (LER) for positrons. The electrons and positrons collide at the interaction point (IP) where the two rings cross. The crossing angle is set to 22 mrad, which plays an important role in terms of simplifying the collision area with no bending magnets for beam separation.

The center-of-mass (CM) energy is mainly tuned to 10.58 GeV. This value is just above the $B\overline{B}$ threshold and corresponds to the mass of the $\Upsilon(4S)$ state composed of a $b\bar{b}$ quark pair. The $\Upsilon(4S)$ state is produced by the collisions of electrons and positrons via a virtual photon. Then it decays to a $B\overline{B}$ pair with a branching fraction of over 96%. The production cross-section of $e^+e^- \to \Upsilon(4S) \to B\overline{B}$ is about 1.2 nb. In addition to the process, various physics processes are produced, whose cross sections are summarized in Table 3.1. Among them, hadrons originating from the $e^+e^- \to q\overline{q}$ process are one of the background sources in analyses of hadronic B decays.

The accelerator was operated from 1999 to 2010. The peak instantaneous luminosity reached 2.11×10^{34} cm^{-2}s^{-1}, which is about twice as large as the design value. The integral luminosity corresponding to the data acquired by the Belle detector is 1040 fb^{-1}. The amount of the data collected at a CM energy of 10.58 GeV is 711 fb^{-1} of the data, which corresponds to 772×10^6 $B\overline{B}$ pairs. The other data was collected at various CM energies from 9.4 GeV to 11.2 GeV, depending on the physics motivations.

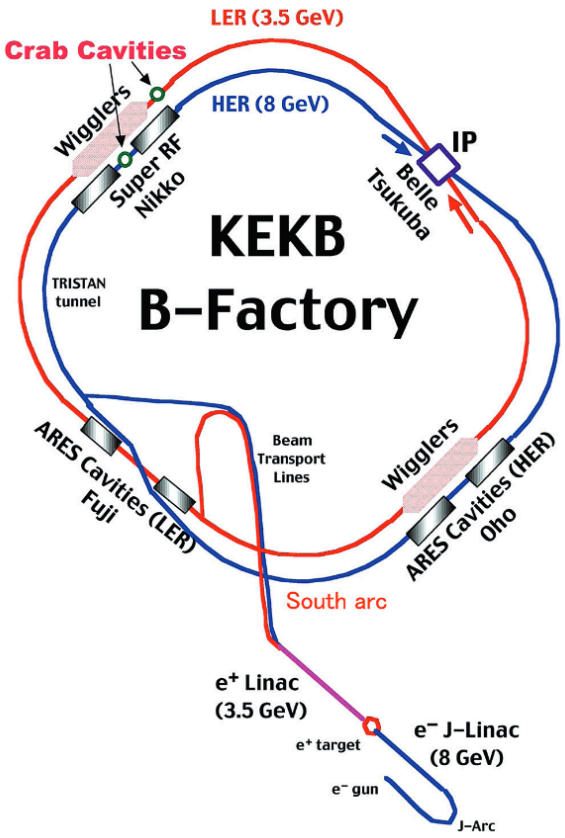


Figure 3.1: Schematic view of KEKB [58].

3.2 Belle Detector

The Belle detector [59] is a large-solid-angle magnetic spectrometer installed around the IP. It is designed and optimized to measure time-dependent CP violation in B-meson decay. The performance, e.g., vertex resolution, particle identification, and low-energy photon detection, is superior.

The configuration of this detector is shown in Fig. 3.2. The beam pipe is surrounded by a silicon vertex detector (SVD) to measure B mesons decay vertices. A central drift chamber is located outside of SVD for tracking of charged particles. Charged particles are identified not only by energy loss in CDC but also by information from aerogel Cherenkov counters (ACC) and time-of-flight counters (TOF). Photons and electrons are detected as an electromagnetic shower by an electromagnetic calorimeter (ECL) installed inside a super-conducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return located outside of the coil is installed to detect K_L^0 mesons and to identify muons (KLM). These sub-detectors cover a wide region of the polar angle from 17° to 150°. In the uncovered angles, an extreme forward calorimeter (EFC) is installed, consisting of an array of bismuth germanate oxide (BGO) crystals. Its main purpose is to improve the sensitivity of some physical processes, such as $B \to \tau\nu$. It is also used as an active shield against the beam background and a monitor of online luminosity.

The subsequent subsection gives a more detailed description of each sub-detector used in this analysis. Since the information from EFC is not used in this analysis, a detailed description is omitted in this thesis; Details are written in Sec. 3 of Ref. [59]. Hereafter,

Table 3.1: Physics processes and total cross-sections at the CM energy of 10.58 GeV [59]
.

Physics process	Cross section [nb]
$e^-e^+ \to \Upsilon(4S) \to B\overline{B}$	1.2
$e^-e^+ \to q\overline{q}$ $(q = u, d, s, c)$	2.8
$e^-e^+ \to l^-l^+$ $(l = \mu, \tau)$	1.6
Bhabha scattering	44
$\gamma\gamma$	2.4
two photons process	~ 15
Total	~ 67

the coordinate system of the Belle detector is used. The definition is a left-handed system with the origin at IP and the z-axis corresponding to the opposite direction of the positron beam. The radius, polar and cylindrical angles are denoted by r, θ and ϕ, respectively.

3.2.1 Silicon Vertex Detector

The vertex detector is designed to fulfill a position resolution of about 100 μm to measure the time-dependent CP violation of B mesons. In addition, this detector can contribute to tracking and reducing background tracks with the information of the impact parameter.

SVD consists of several approximately cylindrical layers of silicon sensors. The initial SVD installed at the beginning of the experiment (SVD1) is a three-layer structure. Because it was degraded due to accumulated radiation damage, it was replaced with an upgraded SVD with a four-layer structure (SVD2) after five years of operation. The data acquired using SVD1 are about 20% of the full data. The respective configurations are shown in Fig. 3.3. The detector covers a polar angle of $17° < \theta < 150°$ ($23° < \theta < 139°$), and the radii of the layers are 20 mm, 43.5 mm, 70 mm and 88 mm (20 mm, 45.5 mm and 60.5 mm), where the value without (with) the parentheses is the specification of the SVD2 (SVD1). Each layer consists of independent ladders, and each ladder consists of double-sided silicon detectors. On one side, p-type strips are stretched along the z-direction, and on the opposite side, n-type strips are stretched along the ϕ-direction. Since these two types of strips are orthogonal, the position of a passing charged particle can be measured in the z–ϕ plane.

3.2.2 Central Drift Chamber

The main role of CDC is the precise determination of three-dimensional trajectories and momenta of charged particles. In addition, it provides information for the trigger system (See Sec. 3.3) and energy loss (dE/dx) used for charged particle identification.

Its structure is asymmetric in the z-direction, covering a region of polar angle of $17° < \theta < 150°$ and radius of 83 mm $< r <$ 874 mm, as shown in Fig. 3.4. The chamber consists of six axial superlayers and five small-angle-stereo superlayers with 8400 drift cells. The drift cells are approximately rectangular, and each cell consists of one sense wire of gold-plated tungsten and eight field wires of unplated aluminum. High voltage, typically 2.35 kV, is applied to the sense wires, and the field wires are connected to the ground. The maximum drift distance is 8–10 mm. In the central part, the axial superlayers and the small-angle stereo superlayers are installed alternately. The stereo layers are used to

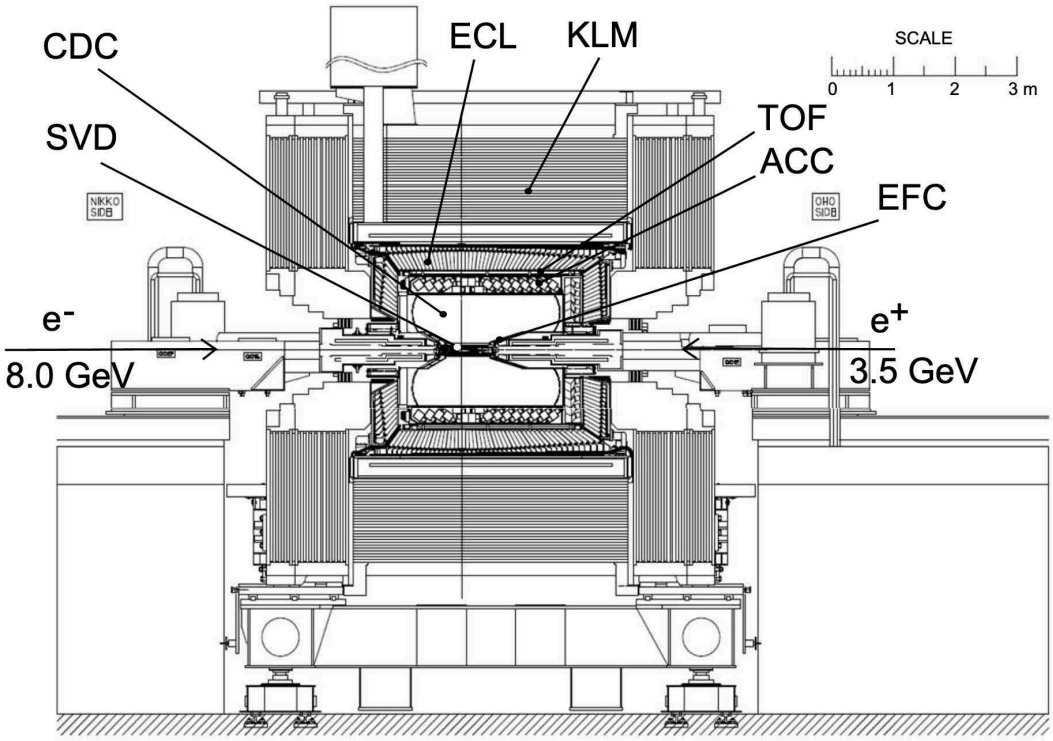


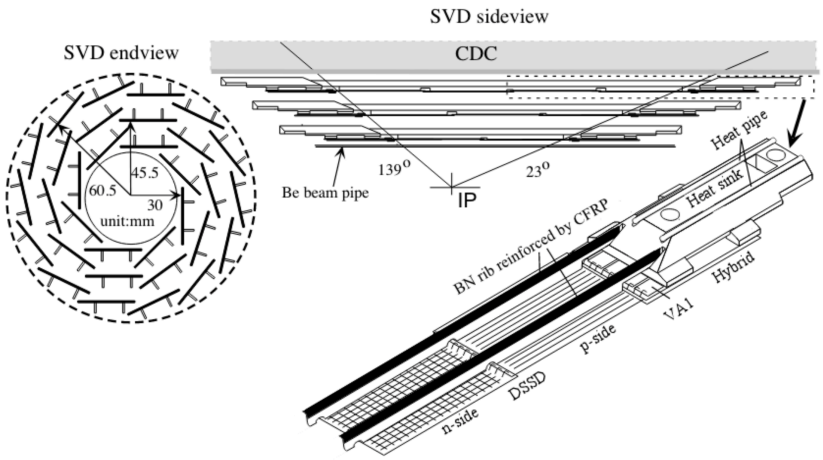
Figure 3.2: Side view of the Belle detector [60].

measure the z-direction in combination with the axial layers. The inner part consists of one axial superlayer with smaller drift cells and three cathode strip layers. The cathode strip layers are attached to the inner cylinder surface of the chamber and have the role of measuring the z-direction. The cathode strips are divided into eight segments in the ϕ direction and have a pitch of 8.2 mm in the z-direction. When the SVD was replaced in 2003, The design of the inner part was changed to make enough space for SVD2. It is a two-layer structure with 128 drift cells of 5 mm in both radial and azimuthal directions.

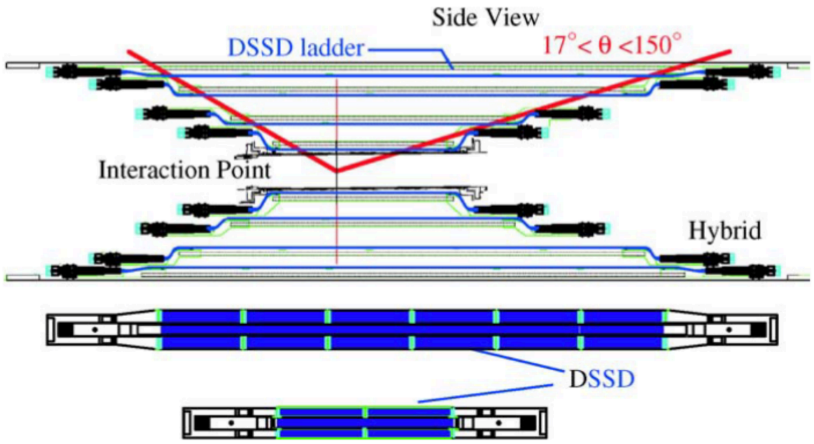
The detector inside is filled with a gas mixture consisting of 50% helium and 50% ethane. Helium gas is one of the low-Z gases that can minimize multiple Coulomb scattering. In addition, it has the advantage of suppressing background from synchrotron radiation due to a smaller photo-electric cross-section than argon. On the other hand, ethane has a long enough radiation length (640 m) and a property in which the drift velocity can saturate at a relatively low electric field in the ethane gas. The latter is important to simplify calibration by reducing the voltage dependence of the distance-time function. In addition, the large ethane component provides good dE/dx resolution.

3.2.3 Time of Flight Counter

The TOF detector using plastic scintillation counters is installed outside CDC. This detector is used to measure the time of flight. The velocity of a charged particle is determined from the time of flight and the distance of flight measured by CDC, and then the charged particle is identified from the velocity and the momentum measured by CDC. It is designed to have a time resolution of 100 ps so that particles with a momentum below 1.2 GeV/c can be well identified. In addition, it provides fast-timing signals for the trigger system.



(a) SVD1 [59]



(b) SVD2 [61]

Figure 3.3: Detector configitaion of SVD .

This TOF system consists of 64 modules installed in a cylindrical section of $34° < \theta < 120°$. Each module consists of two trapezoidal TOF counters and a thin trigger scintillation counter (TSC) with a 1.5-cm gap in the radial direction (Fig. 3.6). Sizes of TOF and TSF are $4.0 \times 6.0 \times 255.0$ cm^3 and $0.5 \times 12.0 \times 263.0$ cm^3, respectively. Scintillator materials are Bicron BC408 for the TOF counters and Bicron BC412 for the TSC counters. At each edge of the TOF counter, a fine-mesh photomultiplier tube of a 2-inch diameter and 24 stages, which is operated in a magnetic field of 1.5 T, is attached with a 0.1 mm air layer. For each TSC counter, the photomultiplier tube is glued to the light guides at the backward edge.

3.2.4 Aerogel Cherenkov Counter

ACC is an array of aerogel threshold Cherenkov counters. Charged particles are identified using momentum measured in CDC and whether Cherenkov light is detected in ACC according to the condition for emitting Cherenkov light when passing through a radiator, $\beta > 1/n$. Here, n and β denote a refractive index and a particle velocity. By selecting a silica aerogel with a refractive index of 1.01 to 1.03 as a radiator, it is possible to identify pion and kaon particles in the high momentum region of 1.0–2.5 GeV that cannot be

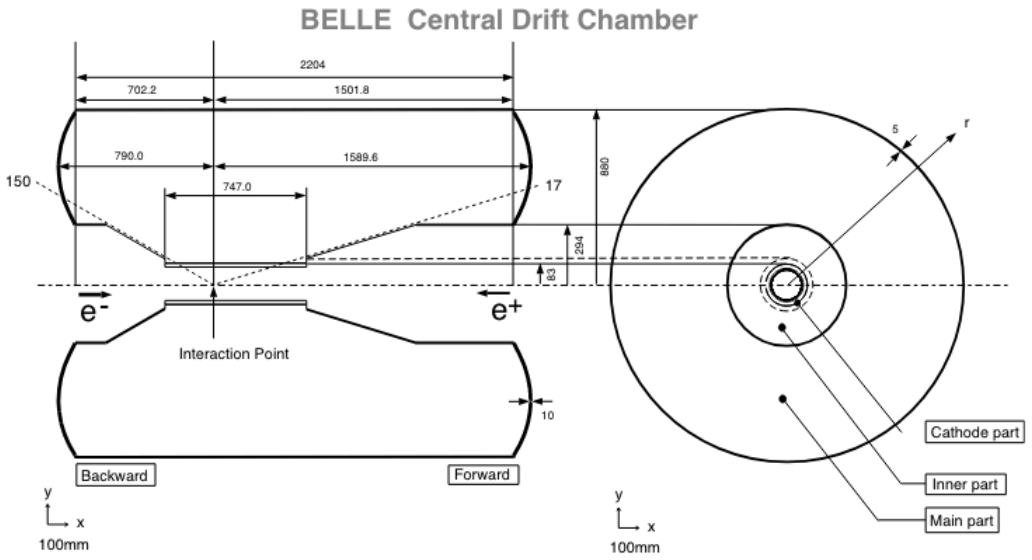
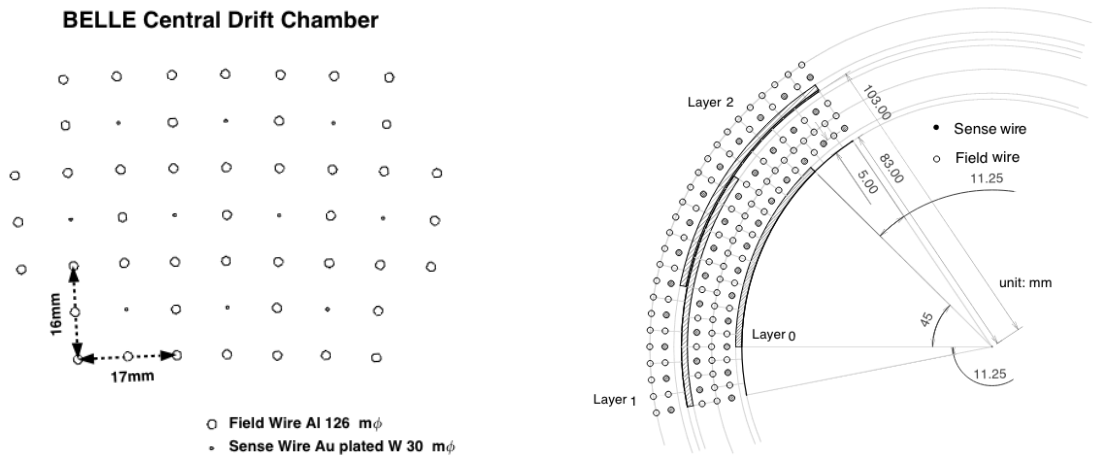


Figure 3.4: The overview of CDC structure [59]



(a) Main part (b) Inner part

Figure 3.5: Cell structure of CDC [59]

covered by dE/dx or time-of-flight.

The detector consists of 960 counter modules in the barrel section ($34° < \theta < 126°$) and 228 modules in the forward endcap region ($17° < \theta < 34°$). All the counters are arranged to point to IP. The typical sizes of the modules are $120 \times 120 \times 120$ mm^3 and $100 \times 100 \times 100$ mm^3 in the barrel and endcap region, respectively. Inside each module, five silica aerogel tiles are stacked in a box, made of 0.2 mm-thick aluminum in the barrel region and 0.5 mm-thick CFEP in the endcap region. To detect Cherenkov light, a fine mesh-type photomultiplier tube, which can be operated in a 1.5 T magnetic field, is used. Two photomultiplier tubes are directly attached to both sides of the box in the barrel region, and one tube is attached to one side through an air light guide in the endcap region.

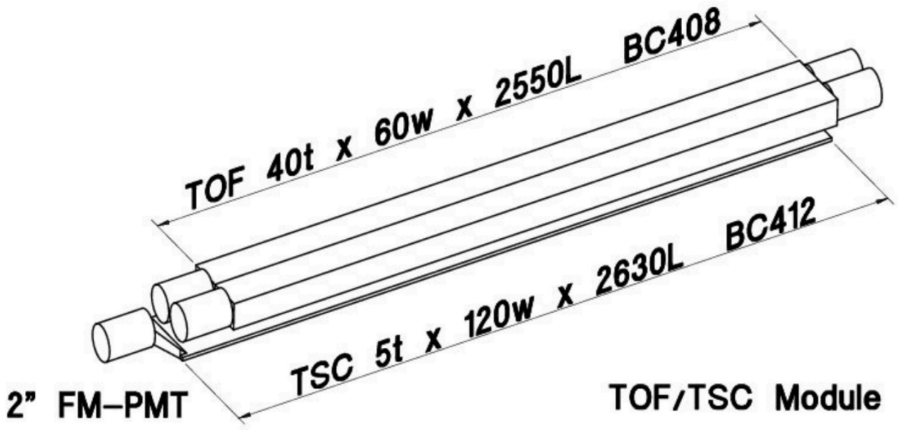
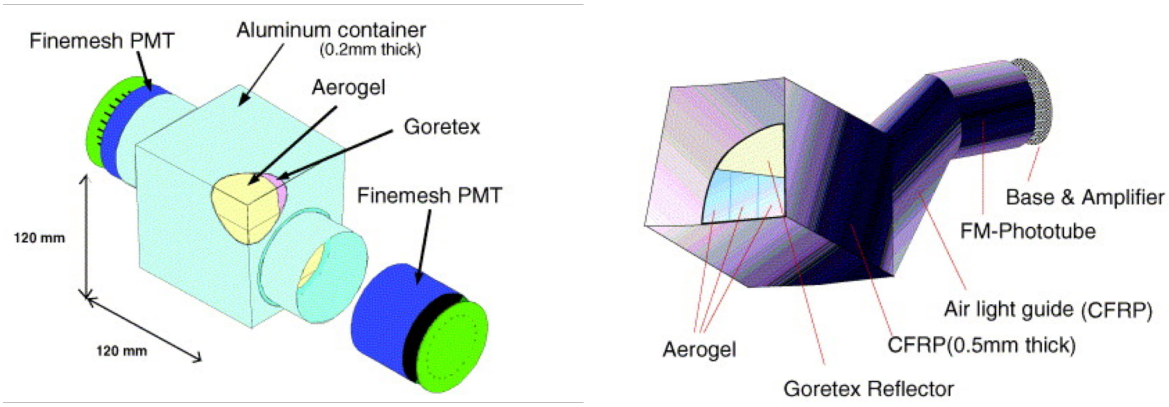


Figure 3.6: Dimensions of a TOF/TSC module [60].



(a) Barrel ACC module (b) Endcap ACC module

Figure 3.7: Schematic drawing of a typical ACC counter module [59].

3.2.5 Electromagnetic Calorimeter

ECL measures the energy deposited by the electromagnetic shower produced when a photon or an electron incidents on a heavy material. It is designed to be detectable over a wide energy range, e.g., from the typical energy of photons produced from B meson, $O(10-100)$ MeV, to high energy photon from two-body B decay, $O(1)$ GeV. The measured energy is also utilize to identify electrons.

The system consists of 8736 CsI(Tl) counters installed in the forward endcap, barrel, and backward endcap regions. It covers the polar angle region of $17.0° < \theta < 150°$ at a radius of 1.25 m. The material of CsI(Tl) was chosen because of its many advantages, such as large photon yield, weak hygroscopicity, and mechanical stability.

Each counter has a tower-like shape and is arranged to point toward the IP. The design of each counter is shown in Fig. 3.8. The crystal size is typically 55×55 mm^2 on the front face and 65×65 mm^2 on the rear face. The length is set to 30 cm, corresponding to a 16 radiation length, to avoid degradation of the resolution of high-energy photons due to leakage of the electromagnetic shower from the rear. On the rear face, two silicon PIN photodiodes with a sensitive area of 10×20 mm^2 and preamplifiers are attached for the readout.

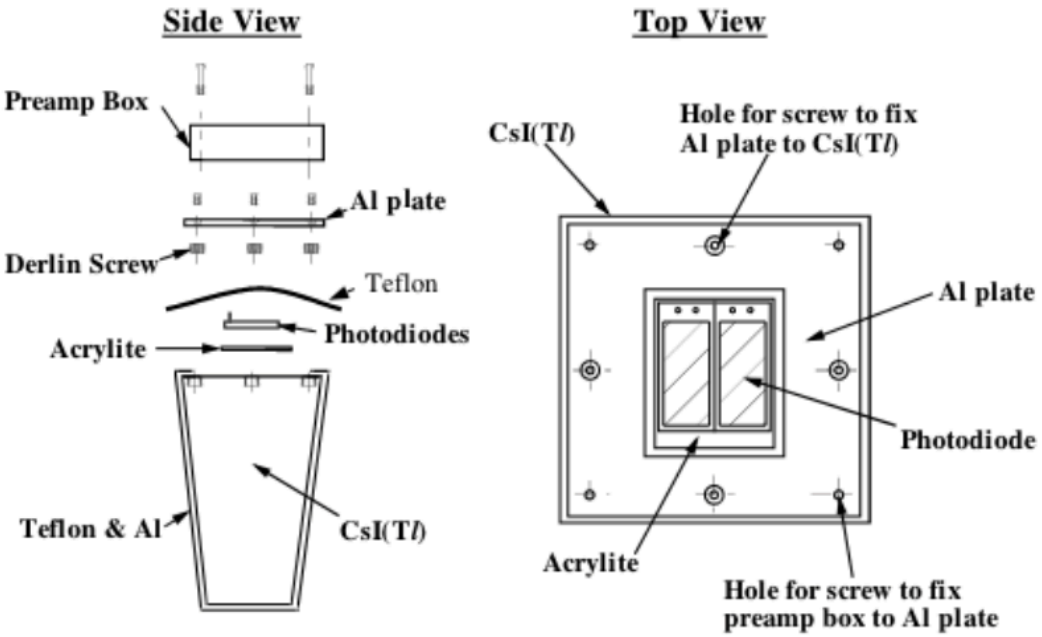


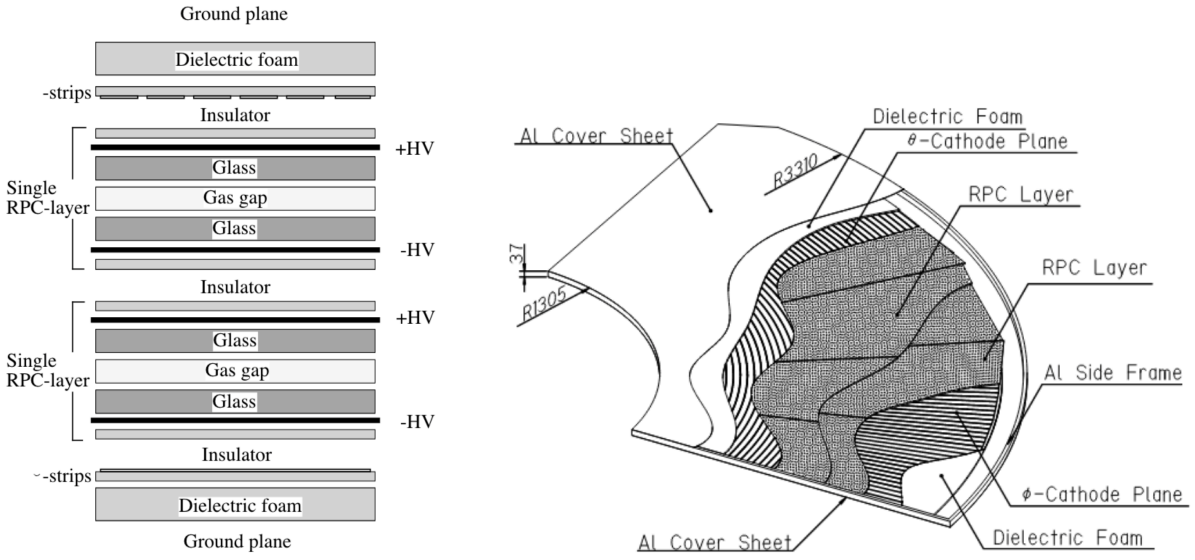
Figure 3.8: Mechanical assembly of the ECL counter [59].

3.2.6 K_L^0 and Muon Detector

KLM is designed to detect muons and K_L^0's with high efficiency in the wide momentum region above 600 MeV/c. It covers both the barrel region of $45° < \theta < 125°$ and the endcap region of $20° < \theta < 45°$ and $125° < \theta < 155°$. The detector consists of 15 (14) layers of charged particle detectors and 14 layers of steel plates of 4.7 cm thickness arranged alternately where the value without (with) the parentheses is the specification of the barrel (endcap) region.

This amount of iron plates corresponds to 3.9 hadronic interaction lengths of material for a particle passing vertically. When a charged particle passes through the steel plates, a charged hadron makes multiple scattering while a muon does not. Therefore, it is identified as a muon with this detector. In the case of K_L^0, it interacts with the steel plates or ECL. The direction of K_L^0 can be determined by the position of this shower, although its energy cannot be measured due to varying size of the shower.

The detector layer is made of glass-electrode resistive plate counters (RPCs). The RPC consists of two parallel plate electrodes of 2.4 mm thick float glass, which have a high bulk resistance ($\geq 10^{10}$ Ωcm) at room temperature. The electrodes are separated by a 1.9 mm gap, and the gap is filled with a non-combustible mixture of 62% HFC-134a, 30% argon, and 8% butane-silver. In order to supply high voltage on the electrodes, the outer surface is coated with carbon ink with high resistivity. The electrodes are supplied with high voltage. The cathodes are set at -3.5 kV. The anodes are set at $+4.7$ kV for the barrel RPCs and $+4.5$ kV for the endcap RPCs. When a charged particle passes through, the charge stored in the high-resistance glass electrode plates is discharged. This discharge gives a signal to the external pickup, and the location and time of ionization can be measured. To read out many pickup strips, VME-based discriminator and time multiplexing boards are used.



(a) Barrel (b) Endcap

Figure 3.9: View of a superlayer module of KLM [59].

3.3 Trigger, Data Acquisition and Preselection

The data acquisition (DAQ) system collects the signals from each sub-detector. In this system, the readout for all the subsystems except SVD is processed with a charge-to-time conversion chip combined with a common FASTBUS multi-hit TDC. These data are collected by event builders based on VME processors, and then processed by the online computer farm consisting of a cluster of VME processor modules. Finally, the data is sent via about 2 km of optical fibers to the KEK Computing Center, and recorded on the offline computing system. To cope with the high background rate arising from higher luminosity, some of the FASTBUS TDCs were replaced by the COPPER TDCs in the middle of the operation. In addition, the VME processors for the event builder were replaced by a set of Linux PC servers (EFARM), and the real-time reconstruction farm (RFARM) was introduced. The overview of the DAQ system is shown in Fig. 3.10.

Uninteresting events are eliminated using three levels of triggers: a Level-1 hardware trigger, a Level-3 software online trigger, and a Level-4 software offline trigger. The Level-1 trigger system consists of sub-detector trigger systems and a central trigger system (global decision logic, GDL), as shown in Fig. 3.11. As the sub-detector trigger systems, each trigger is provided with parallel signals from CDC, ECL, TOF, KLM, and EFC. The GDL receives the information from each sub-triggers and combines them to characterize an event type for the hadronic events, Bhabha events, $\mu^+\mu^-$ pair events, and so on. When an event is characterized, a trigger signal is formed in GDL and issued 2.2 μs after the event occurs. The trigger rate is typically 200–400 Hz. Concerning the hadronic events, the events are characterized by the following conditions: the presence of three or more charged tracks, high energy deposition in ECL, or four isolated neutral clusters in ECL. The trigger efficiency is greater than 99%.

On the online computer farm, event data is formatted for offline processing and the Level-3 trigger is performed to suppress beam background events. In the trigger, only those events are retained in which at least one of the charged-particle track candidates originates near the IP ($|dr| < 5.0$ cm).

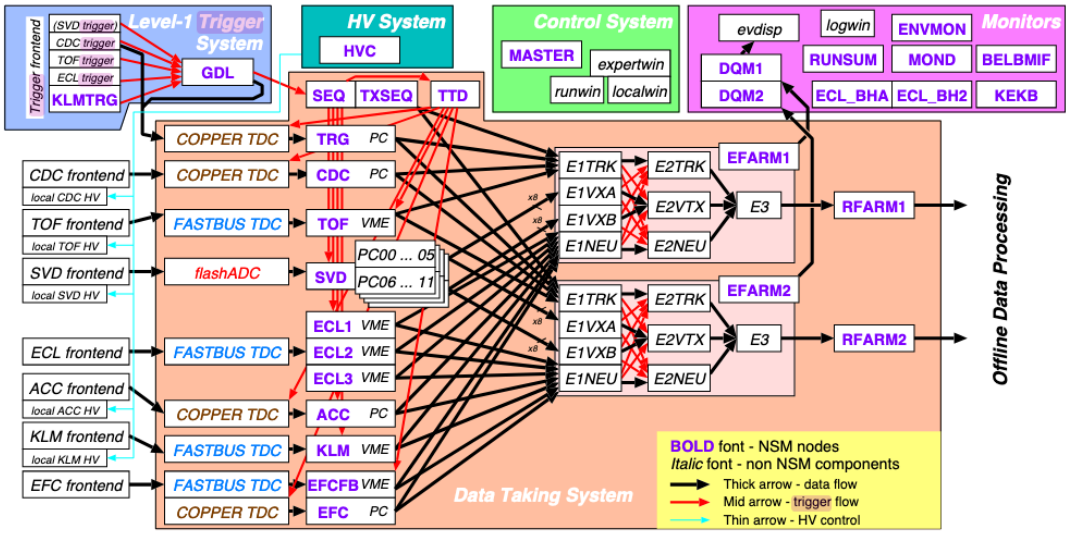


Figure 3.10: The configuration of the DAQ system at the end of data taking [60].

After the Level-3 trigger, the remaining events are recorded on the offline computing system. Just before full event reconstruction, the Level-4 trigger performs a fast track reconstruction and retains only events with at least one good track. A concrete criterion of the good track is that a transverse momentum is greater than 300 MeV/c, $|dr|$ is less than 1.0 cm, and $|dz|$ is less than 4.0 cm. This trigger reduces the background events to about 20%, whereas it retains about 98% of physical events.

After performing the Level-4 trigger, the remaining events are fully reconstructed. At the most downstream of the reconstruction flow, the event classification and skimming are performed to be efficient for the physics analyses. The data skimmed for hadronic events are used in this analysis. The primary requirements are track multiplicity and visible energy. The requirement for track multiplicity is that there are at least three tracks with transverse momentum of more than 0.1 GeV/c, $|dz|$ of at least 4.0 cm, and $|dr|$ of at least 2.0 cm. The requirement for visible energy is that the sum of the energy of charged tracks and reconstructed photons E^*_{vis} is greater than 20% of the $\sqrt{s}$. To reduce beam background and non-hadronic events, events satisfied with the following criterion are retained; (1) the vertex position is in the vicinity of IP if the vertex is well reconstructed, (2) there are two or more clusters detected at the large polar angular in the center-of-mass system, (3) the average ECL cluster energy is less than 1 GeV, (4) the total energy of ECL clusters is less 80% of $\sqrt{s}$, (5) an invariant mass of particles detected in hemispheres perpendicular to the event thrust axis, HJM, is greater than 1.8 GeV/c^2 or $HJM/E^*_{vis} > 0.18\sqrt{s}$. In addition, the inclusive J/ψ and $\psi(2S)$ samples are added.

3.4 Data Set

This analysis uses a real data sample to obtain physics results and a Monte-Calro (MC) sample to determine the event selection criteria and investigate the detector response. A description of each data set is given in this section.

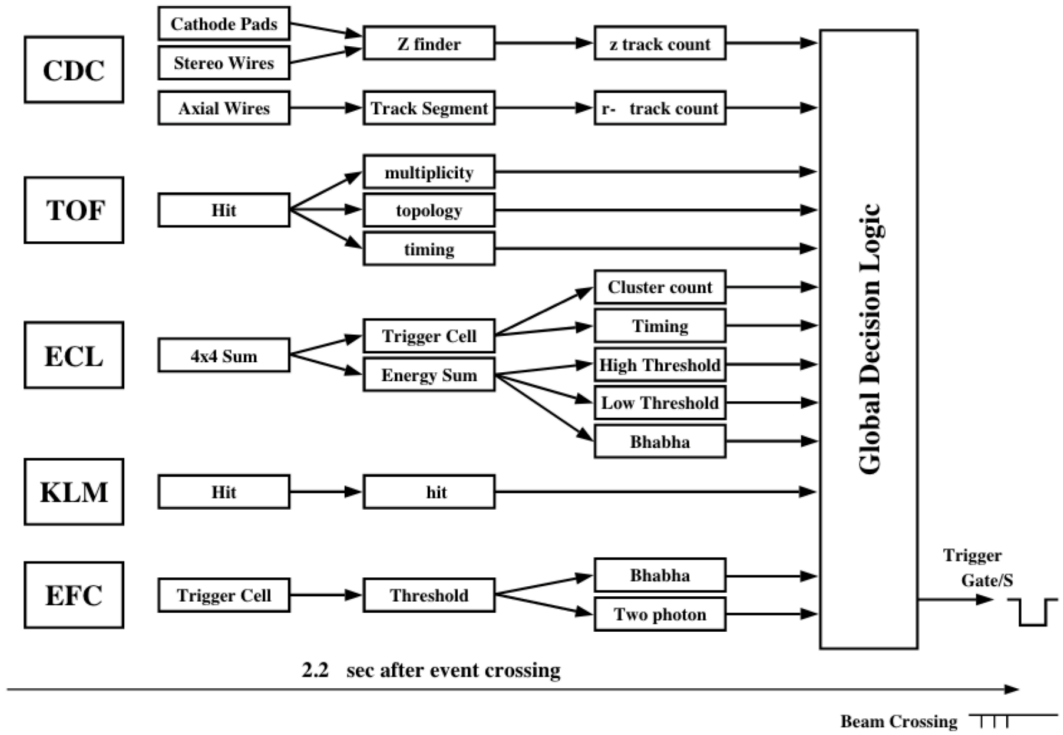


Figure 3.11: The overview of the Level-1 trigger system [59].

3.4.1 Real Data Sample

For this analysis, 711 fb^{-1} of data accumulated with the Belle detector at a center-of-mass energy of $\sqrt{s} = 10.58$ GeV. This data contains 772×10^6 $B\overline{B}$ pairs.

3.4.2 Monte-Calro Simulation Samples

In this analysis, we use a signal MC sample and a background MC sample according to the purpose. In both samples, amplitudes of B decay and other sequential decays are simulated using `EvtGen` [62]. For generic B decay, the branching fractions and the decay models are mainly specified to fit experimental measurements to date, theory, phenomenological models, and symmetry arguments. However, approximately 40% of B decays are not specified, and thus `PYTHIA` and `JETSET` [63] compensate them. For quark fragmentations in the continuum events, `PYTHIA` and `JETSET` are used, too. Subsequently, the detector response is simulated using `GEANT3` [64].

Signal MC sample

Signal MC samples are used to determine detector response in detail. The samples are generated separately for the following $\Upsilon(4S)$ decays at a branching fraction of 100%,

$$\Upsilon(4S) \to \begin{cases} B^+B^-, \ B^+ \to X(3872)K^+, \\ B^0\overline{B}^0, \ B^0 \to X(3872)K^0. \end{cases}$$

All of $X(3872)$ decays into the following modes

$$
\begin{aligned}
X(3872) &\to D^0\overline{D}^{*0}, \\
\overline{D}^{*0} &\to \begin{cases} \overline{D}^0\pi^0 & (64.7\%), \\ \overline{D}^0\gamma & (35.3\%), \end{cases} \\
D^0 &\to \begin{cases} K^-\pi^+ & (10.1\%), \\ K^-\pi^+\pi^0 & (36.7\%), \\ K^-\pi^+\pi^-\pi^+ & (20.9\%), \\ K^0_S\pi^+\pi^- & (7.1\%), \\ K^+K^- & (1.1\%), \\ K^-\pi^+\pi^-\pi^+\pi^0 & (10.9\%), \\ K^0_S\pi^+\pi^-\pi^0 & (13.2\%), \end{cases}
\end{aligned}
$$

where the fractions among the D^{*0} and D^0 modes follow the known ratios [18]. Although the $D^0 \to K^-\pi^+\pi^-\pi^+\pi^0$ mode is generated in this sample, it is not used for the final analysis due to a bad signal-to-background ratio.

When examining the detector response, we generally use MC samples in which the event generator generates the lineshape of interest. However, all lineshape models, e.g., the relativistic Breit-Wigner model and the Flatté model, are not defined in the event generator. We establish a model-independent method of determining detector response using two groups of signal MC samples: zero-width signal MC samples and finite-width signal MC samples. The former are necessary to parameterize the $D^0\overline{D}^{*0}$ invariant mass dependence of signal efficiency and mass resolution, which are essential to make the probability density functions of the signal contribution for any lineshapes. The latter are used to check the validity of the fitting function, the total efficiency and so on. Among the latter, the signal MC sample in which lineshapes are generated with 3.9 MeV width and 3.8729 GeV/c^2 mass [2] is called the *default signal MC sample*. Detailed input values are summarized in Table 3.2.

Background MC sample

To investigate background events except for the signal events, we use the MC samples of $e^+e^- \to B\overline{B}$ with generic B decay and the continuum MC samples. The amount of the MC samples corresponds to 711 fb^{-1} of integrated luminosity.

Table 3.2: Summary of the signal MC samples used in this study. The number of generated events is denoted by N_{sig}.

Type	Mass [GeV/c^2]	Width [MeV]	N_{gen}
Zero-width signal MC samples	3.8719	0.0	8×10^5
	3.8720		
	3.8721		
	3.8722		
	3.8723		
	3.8724		
	3.8725		
	3.8730		
	3.8735		
	3.8740		
	3.8745		
	3.8750		
	3.8755		
	3.8760		
	3.8765		
	3.8790		
	3.8815		
	3.8840		
	3.8865		
	3.8890		
	3.8915		
	3.8940		
	3.8965		
	3.8990		
	3.9100		
	3.9300		
	3.9500		
	3.9700		
	3.9900		
Finite-width signal MC samples Second one is "default signal MC samples"	3.8725	1.0	8×10^5
	3.8729	3.9	
	3.8729	10	

Chapter 4

Signal Reconstruction

We reconstruct $B^+ \to D^0\overline{D}^{*0}K^+$ or $B^0 \to D^0\overline{D}^{*0}K^0_S$ followed by $\overline{D}^{*0} \to \overline{D}^0\gamma$ or $\overline{D}^{*0} \to \overline{D}^0\pi^0$ (Fig. 4.1) using five types of final state particles: π^+, K^+, γ, π^0, and K^0_S. This chapter describes the reconstruction of the final state particles and the signal event selection. The last section describes the observable and the expected distribution of the $D^0\overline{D}^{*0}$ invariant mass.

4.1 Reconstruction of Final State Particles

In this section, the reconstructions of the final state particles, π^+, K^+, γ, π^0, and K^0_S, are described.

Charged Track Reconstruction

The stable charged particles such as π^+ and K^+ are reconstructed by finding helical tracks from a cluster of hits of CDC and SVD. The track is parameterized as a helix by fitting. The helix parameters are the following five:

dr: the signed distance of the Point Of Closest Approach (POCA) to the IP for the r-direction;

ϕ_0: the azimuthal angle to specify the IP with respect to the helix center;

κ: the inverse of track transverse momentum times a charge of the track;

dz: the signed distance of POCA to the IP for the z-direction; and

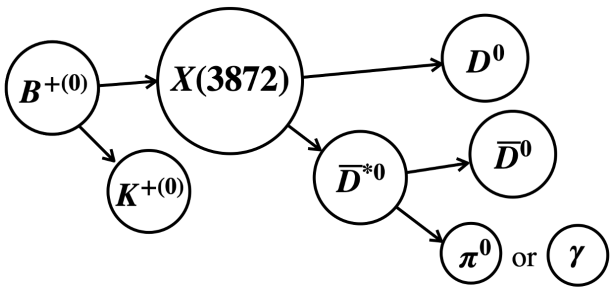


Figure 4.1: The decay chain of signal events.

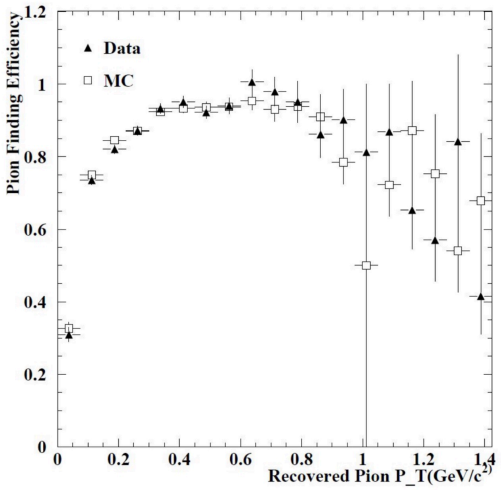


Figure 4.2: Reconstruction efficiency for charged tracks as a function of the transverse momentum of the particle [60].

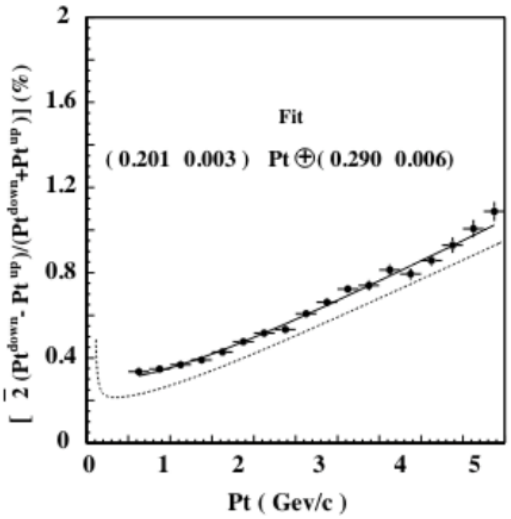


Figure 4.3: The p_t resolution. The solid curve shows the fitted result. The dotted curve shows the ideal expectation for $\beta = 1$ [59].

$\tan\lambda$: the tangent of the angle defined by the momentum at the POCA and the transverse plane.

The parameters are converted to the position and momentum of the charged particle.

Important performances for this analysis are the detection efficiency and the momentum resolution. The tracking efficiency was evaluated using a $D^{*+} \to \pi^+(D^0 \to \pi^+\pi^- K^0_S)$ sample. Figure 4.2 shows the efficiency as a function of transverse momentum. The efficiency is above 90% in the transverse momentum range from 400 MeV/c to 1000 MeV/c. For momentum above 200 MeV/c, the efficiency in the MC samples reproduces that in the data within 0.35%. For momentum below 200 MeV/c, the performance was precisely evaluated by the sample of soft-charged pions from D^{*-} in the $B^0 \to D^{*-}\pi^+$ decay. The data-MC difference is up to 15%. A correction factor is applied to the signal efficiency based on the ratio of tracking efficiency obtained for MC and data.

The momentum resolution is investigated using data of cosmic rays passing through the interaction region. It is good, as shown in Fig. 4.3, and the relative p_t resolution is obtained by

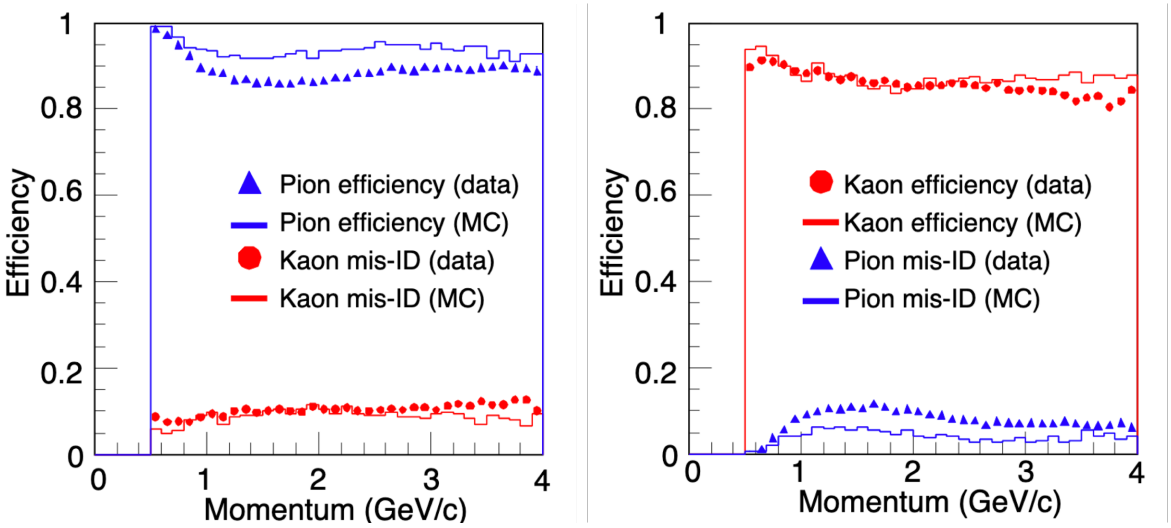
$$\sigma_{p_t}/p_t = (0.201 \pm 0.003)p_t \otimes (0.290 \pm 0.003)/\beta\%.$$

In the later analysis, the helix parameters are smeared in the MC samples so as to get better agreement between the mass resolution of the hadrons reconstructed with only charged particles for the MC samples and that for the data.

Charged Hadron Identification

The charged hadron is identified using a likelihood ratio derived from the CDC, TOF, and ACC information,

$$\frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi} \text{ (for } K^+ \text{ candidates) and } \frac{\mathcal{L}_\pi}{\mathcal{L}_K + \mathcal{L}_\pi} \text{ (for } \pi^+ \text{ candidates)},$$



(a) Pion ID (b) Kaon ID

Figure 4.4: Performance of the charged hadron identification [60].

where $\mathcal{L}_i$ is the total likelihood for the particle hypothesis i. It is derived from a product of the sub-detector likelihoods,

$$\mathcal{L}_i = \mathcal{L}_i^{\mathrm{CDC}} \times \mathcal{L}_i^{\mathrm{TOF}} \times \mathcal{L}_i^{\mathrm{ACC}}. \tag{4.1}$$

A likelihood of each detector is calculated as follows.

$\mathcal{L}_i^{\mathrm{CDC}}$; Based on dE/dx measured in CDC, a likelihood is calculated by the probability density function, which is a normalized Gaussian distribution with a mean of the expected dE/dx and a standard deviation corresponding to resolution for the particle i.

$\mathcal{L}_i^{\mathrm{TOF}}$; Based on time-of-flights measured at both ends of a TOF counter, a likelihood is calculated by the probability density function, which is a two-dimensional normalized Gaussian distribution with means of the expected time-of-flights and standard deviations of the resolution taking into account their correlation for the particle i.

$\mathcal{L}_i^{\mathrm{ACC}}$; Based on the number of photo-electron N_{pe} detected in ACC, a likelihood is calculated by the probability density function of N_{pe} expected from the MC simulation.

The identification performance is evaluated using the sample of the decay $D^{*+} \to D^0\pi^+$ followed by $D^0 \to K^-\pi^+$. The identification efficiency and the miss identification rate as a function of momentum are shown in Fig 4.4. The efficiency is more than 85% in the whole momentum range, while the miss identification rate is suppressed to about 10%. In the later analysis, a correction factor is applied to the signal efficiency based on the ratio of the hadron identification efficiencies obtained for MC and data.

Photon Reconstruction

Photon candidates are reconstructed by the energy deposition of an electromagnetic shower in ECL. A group of crystals in which energy is detected is called a cluster. The

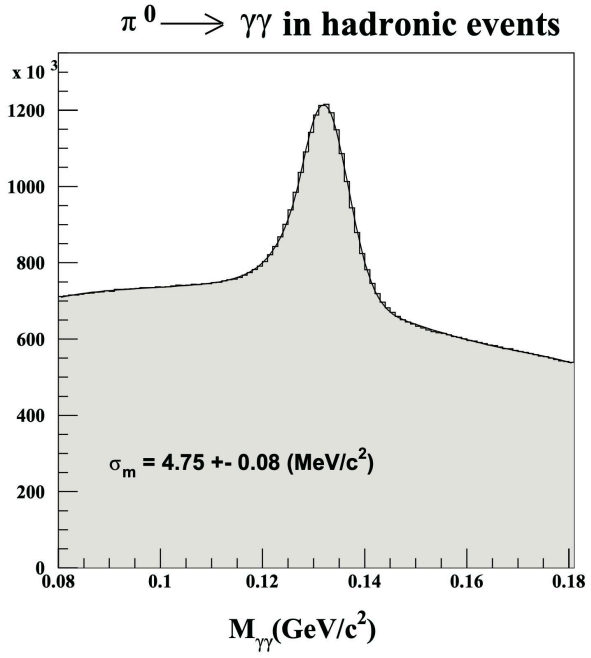


Figure 4.5: The $\gamma\gamma$ invariant mass distribution for hadronic events [59].

reconstructed clusters with no matching charge tracks are used as photon candidates. The bias of the energy and position measurements, caused by the energy deposition in ACC and TOF and the leakage of electromagnetic showers for high-energy photons, are always corrected based on the MC study.

The resolutions of energy and position and the detection efficiency are evaluated. The energy resolution is from about 4% at 100 MeV to 1.6% at 8 GeV. The angular resolution is from 13 mrad at low energy to 3 mrad at high energy. The detection efficiency is evaluated using the $\chi_{c1} \to J/\psi\gamma$ decay in the low energy region of $O(0.1)$ GeV, which is the energy range of signal photons. The difference in efficiency obtained for MC and data is 3.0%.

π^0 Reconstruction

Candidates of π^0 are reconstructed from the four-momenta of two-photon candidates since π^0 decays to two photons with a branching fraction of 98.8% [18]. Only candidates in which the $\gamma\gamma$ invariant mass is consistent with the nominal mass of 135 MeV/c^2 [18] are used in the analysis. At the initial stage of the analysis, the mass window is set to 80 MeV/c^2 $< M(\gamma\gamma) <$ 180 MeV/c^2, which is optimized later. Finally, a mass-constrained fit is performed on these π^0 candidates to update the three-momenta.

The mass resolution is evaluated using the hadronic events of the data, as shown in Fig. 4.5. The typical resolution is determined to be 5.0 MeV/c^2. The reconstruction efficiency is evaluated using the $\tau^- \to \pi^-\pi^0\nu_\tau$ sample for high momentum and the sample of $B^+ \to (D^{*0} \to \overline{D}^0\pi^0)\pi^+$ for low momentum. The efficiency for the data is about 4% lower than that for the MC simulation with a precision of about 1.5%. Based on the ratio of the π^0 efficiencies obtained for MC and data, a correction factor is applied to the signal efficiency.

K^0_S Reconstruction

The neutral kaon decays to two charged pions with opposite charges with a branching fraction of 69.2% [18]. Since K^0_S has a relatively long lifetime, the vertex position is separated from IP. Then, the vertex position is reconstructed by applying a vertex-constrained fit. In the fit, the momenta and positions of the daughter particles are updated to those derived from the helix parameter, assuming that the pivot equals to the vertex position.

4.2 Selection of Signal Event

Selection criteria are optimized so that Figure-Of-Merit (FOM) is improved while the signal efficiency keeps as high as possible. In this analysis, FOM is defined as $S/\sqrt{S+B}$, where S and B are the number of signal events and that of background events, respectively. The number of signal events is estimated by scaling the signal MC sample so that the branching ratio is 0.8×10^{-4} [2]. The background consists of broken-signal and generic background. The former is the background coming from the wrong combination in the signal event, and the same scale factor as the signals is applied. The latter is the background coming from continuum events and $B\overline{B}$ events except for the signals. To focus on the $B \to X(3872)K$ signal region, both S and B are counted only in the region of $M(D^0\overline{D}^{*0}) < 3.877$ GeV/c^2, 5.271 GeV/$c^2 < M_{\rm bc} <$ 5.287 GeV/c^2, and $|\Delta E| <$ 16.1 MeV (33.6 MeV) for $B \to KD^0\overline{D}^0\pi^0(\gamma)$. In this section, the resulting selection is described; See Appendix A for the optimization details.

4.2.1 Final State Particles

Tracks coming not from K^0_S are selected using vertex information measured by the tracking system. The requirement is that $|dr|$ is less than 1.0 cm and $|dz|$ is less than 4.0 cm. In addition, pion and kaon candidates are selected using likelihood ratios. In this analysis, tracks with a likelihood ratio $\mathcal{L}_\pi/(\mathcal{L}_\pi+\mathcal{L}_K) > 0.1$ are used as charged pion candidates, and tracks with a likelihood ratio $\mathcal{L}_K/(\mathcal{L}_\pi + \mathcal{L}_K) > 0.1$ are used as charged kaon candidates. The hadron identification efficiency is approximately 97% for both pions and kaons. Tracks satisfying $\mathcal{L}_e/(\mathcal{L}_e + \mathcal{L}_{\tilde{e}}) > 0.95$ are identified as electrons and are eliminated. Here, $\mathcal{L}_e$ and $\mathcal{L}_{\tilde{e}}$ are distinct likelihoods for the electron and non-electron hypotheses, based on ECL, tracking, and other information. The particle identification is described in detail elsewhere [65].

K^0_S candidates are selected by the $\pi^+\pi^-$ invariant mass. The $\pi^+\pi^-$ invariant mass distribution is shown in Fig. 4.6. The $\pi^+\pi^-$ invariant mass is required to agree with the known K^0_S mass (498 MeV/c^2 [18]) within 7.0 MeV/c^2 ($\approx 3.6\sigma$ of the resolution). The candidates are selected using a neural network classifier [66] with various kinematic variables as input. To improve the four-momenta of the K^0_S candidates, a mass- and vertex-constrained fit is applied.

Photon candidates are selected by the ratio of the energy deposited in the 3×3 array of the crystals centered on the crystal with the highest energy deposition to that in the 5×5 array of crystals, E_9/E_{25}. The requirement is that E_9/E_{25} is greater than 0.8.

For the selection of neutral pions candidates, an energy of the daughter photon is required to be greater than 30 MeV in the barrel region or 50 MeV in the endcap region. The $\gamma\gamma$ invariant mass is required to agree with the π^0 nominal mass (135 MeV/c^2 [18]) within

12 MeV/c^2. This mass window corresponds to 92% signal efficiency. The distribution of $\gamma\gamma$ invariant mass is shown in Fig. 4.7.

4.2.2 Candidates of D^0 meson

D^0 candidates are reconstructed in seven decay modes: $K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^-\pi^+$, $K^-\pi^+\pi^-\pi^+\pi^0$, $K^0_S\pi^+\pi^-$, $K^0_S\pi^+\pi^-\pi^0$ or K^+K^- mode. The $K^-\pi^+\pi^-\pi^+\pi^0$ mode is totally eliminated to improve sensitivity, because the mode has high background. Thus, the used decay modes cover 35.0% of the total D^0 decay branching fraction [18]. The π^0 candidates used in this reconstruction are required to have momentum in the center-of-mass system greater than 100 MeV/c, and energy in the laboratory system greater than 150 MeV. If a π^0 is included, the reconstructed D^0 invariant mass is required to be within 16 MeV/c^2 of the nominal mass [18] corresponding to 85% signal efficiency; otherwise, it is required to be within 8.5 MeV/c^2 corresponding to 91% efficiency. Each invariant mass distribution of reconstructed D^0 is shown in Fig. 4.8. To improve the momentum resolution, a mass- and vertex-constrained fit is applied. The candidates with poor quality fit whose χ^2 probability is less than 0.0001 are eliminated.

4.2.3 Candidates of D^{*0} meson

$\overline{D}^{*0}$ candidates are reconstructed in two decay modes: $\overline{D}^0\gamma$ and $\overline{D}^0\pi^0$. For the $\overline{D}^0\gamma$ mode, only γ candidates with an energy greater than 90 MeV in the laboratory system are used. For the $\overline{D}^0\pi^0$ mode, π^0 candidates with a momentum in the center-of-mass system of less than 100 MeV/c and an energy in the laboratory system of less than 200 MeV are used. The difference of the reconstructed mass between $\overline{D}^{*0}$ and $\overline{D}^0$ is required to agree with the nominal value (142 MeV/c^2 [18]) within 9.0 MeV/c^2 and 2.0 MeV/c^2 for $\overline{D}^0\gamma$ and $\overline{D}^0\pi^0$, respectively, corresponding to 90% signal efficiency. Figure 4.9 shows the distributions of the mass difference and the selection criteria.

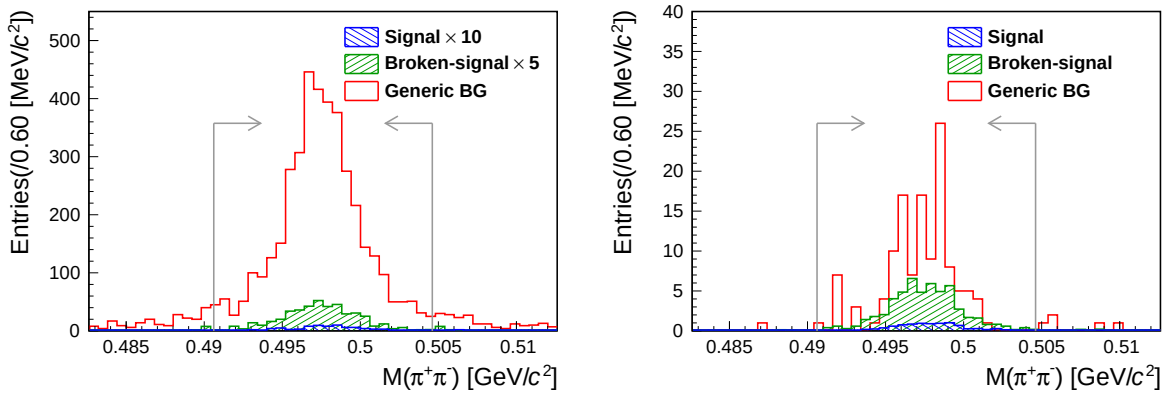
4.2.4 Candidates of B meson

B meson candidates are reconstructed in the decay modes of $D^0\overline{D}^{*0}K^+$ and $D^0\overline{D}^{*0}K^0_S$. To reduce wrong combinations, the daughter K^+ is required to have $\mathcal{L}_K/(\mathcal{L}_\pi+\mathcal{L}_K) > 0.6$, corresponding to the identification efficiency of 89%. The B candidates are selected based on the beam-energy constrained mass, $M_{\rm bc} \equiv \sqrt{(E^{\rm cms}_{\rm beam})^2 - (p^{\rm cms}_B)^2}$ and the difference of the energy in the center-of-mass system between the B candidate and the beam, $\Delta E \equiv E^{\rm cms}_B - E^{\rm cms}_{\rm beam}$, where $E^{\rm cms}_{\rm beam}$ is the beam energy in the center-of-mass system corresponding to half of $\sqrt{s}$, and $p^{\rm cms}_B$ and $E^{\rm cms}_B$ are the energy and momentum of B candidates in the center-of-mass system, respectively. Figures 4.10–4.11 shows their distributions. We retain events with $M_{\rm bc} > 5.2$ GeV/c^2 and $|\Delta E| < 50$ MeV for later analysis. The $M_{\rm bc}$ signal region is defined as $|M_{\rm bc} - m_B| < 4.5$ MeV/c^2 ($\approx 2\sigma$) for $\overline{D}^{*0} \to \overline{D}^0\gamma$ and $|M_{\rm bc} - m_B| < 6.0$ MeV/c^2 ($\approx 2.5\sigma$) for $\overline{D}^{*0} \to \overline{D}^0\pi^0$, where m_B denotes the nominal B mass [18]. The ΔE signal region is defined as $|\Delta E| < 12$ MeV ($\approx 2\sigma$).

4.2.5 Continuum Supression

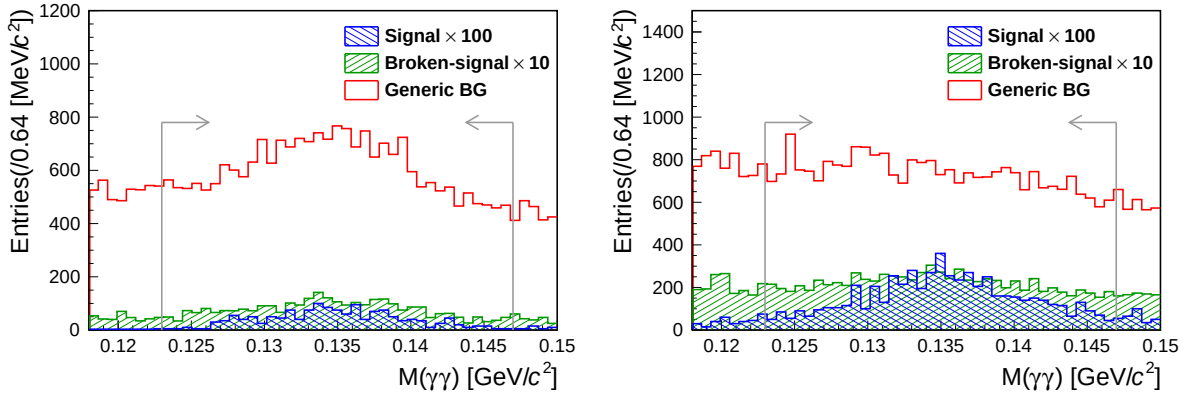
For suppression of continuum events, we use a FastBDT classifier [67] trained on the simulation sample with event-shape information as input; See Appendix B for input variables, training, and classifier performance. The distributions of the output are shown in

Fig. 4.12. Events for which the classifier output is less than 0.15 are eliminated. This requirement retains 96% of the signal candidates and rejects 49% of the candidates of continuum events.



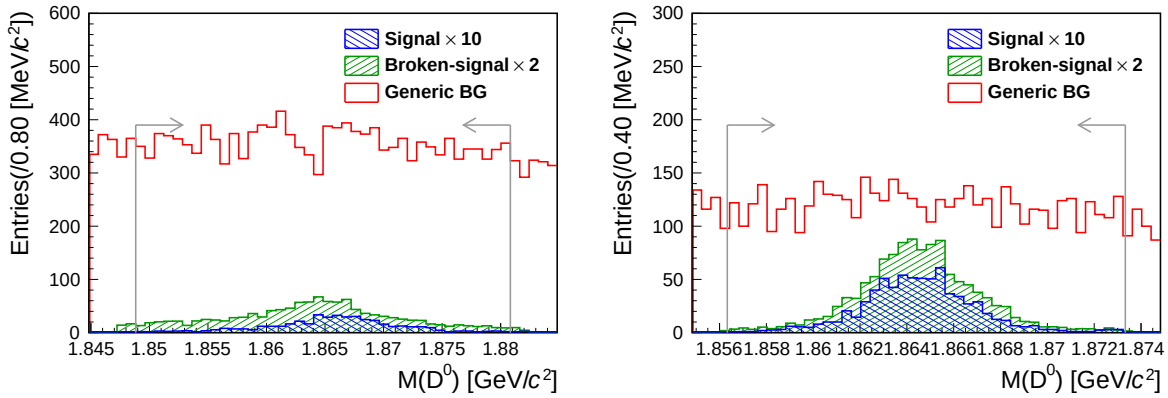
(a) K^0_S candidates from the D^0 decay (b) K^0_S candidates from the B^0 decay

Figure 4.6: The $\pi^+\pi^-$ invariant mass distribution in the $X(3872)$ and B signal regions for the MC samples. The blue, green, and red histograms show distributions for signal, broken-signal, and generic background events, respectively. The gray vertical line shows the criterion used for the later analysis.



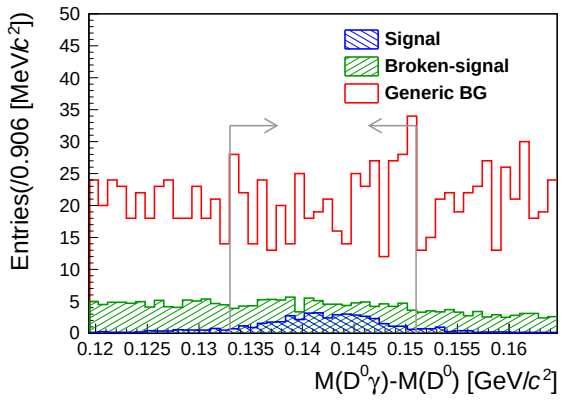
(a) π^0 candidates from the D^0 decay (b) π^0 candidates from the D^{*0} decay

Figure 4.7: The distributions of the $\gamma\gamma$ invariant mass for each source of π^0 candidates in the $X(3872)$ and B signal regions for the MC samples. The notations are the same as in Fig. 4.6.

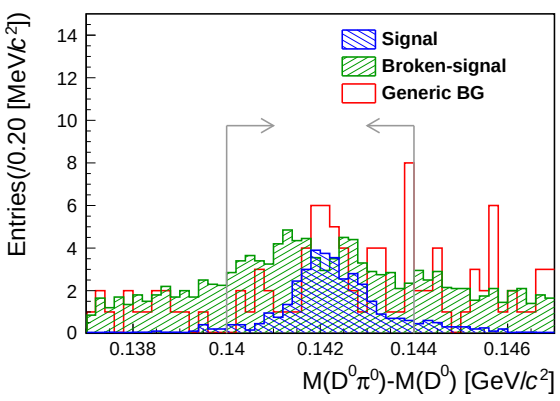


(a) D^0 candidate reconstructed with π^0 (b) D^0 candidate reconstructed without π^0

Figure 4.8: The distribution of the reconstructed D^0 mass in the $X(3872)$ and B signal regions for the MC samples. The notations are the same as in Fig. 4.6.

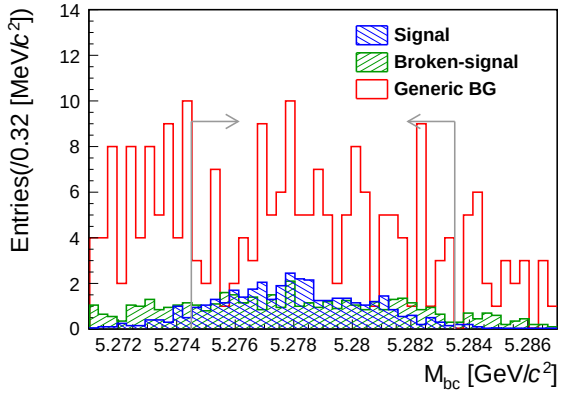


(a) Decay mode of $D^{*0} \to D^0\gamma$

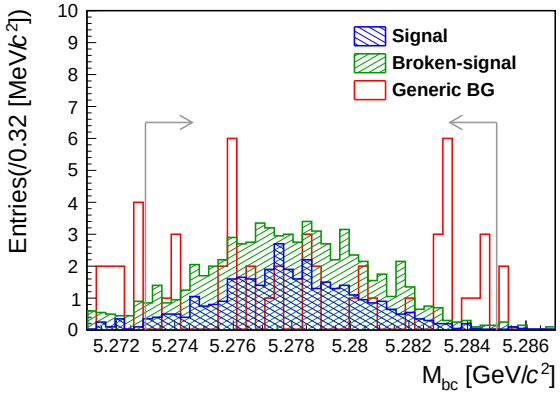


(b) Decay mode of $D^{*0} \to D^0\pi^0$

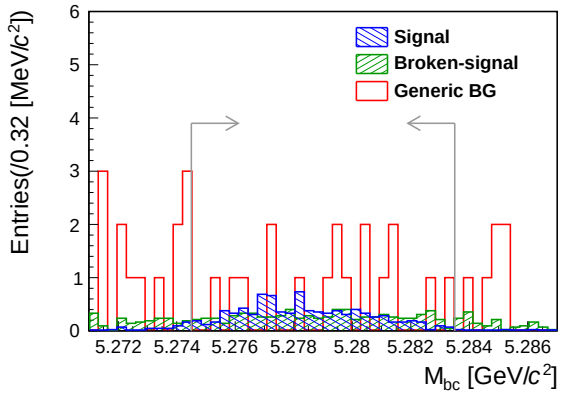
Figure 4.9: The distribution of the difference of the reconstructed mass between $\overline{D}^{*0}$ and $\overline{D}^0$ in the $X(3872)$ and B signal region for the MC samples. The notations are the same as in Fig. 4.6.



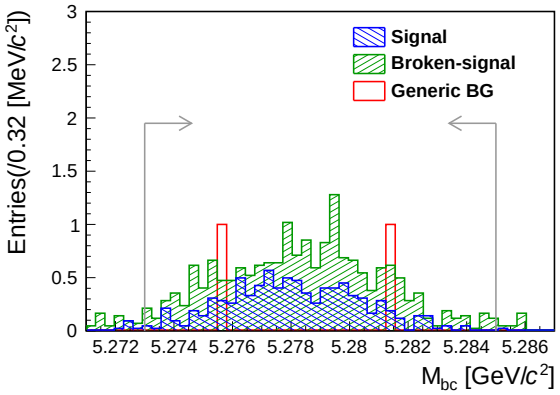
(a) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$



(b) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

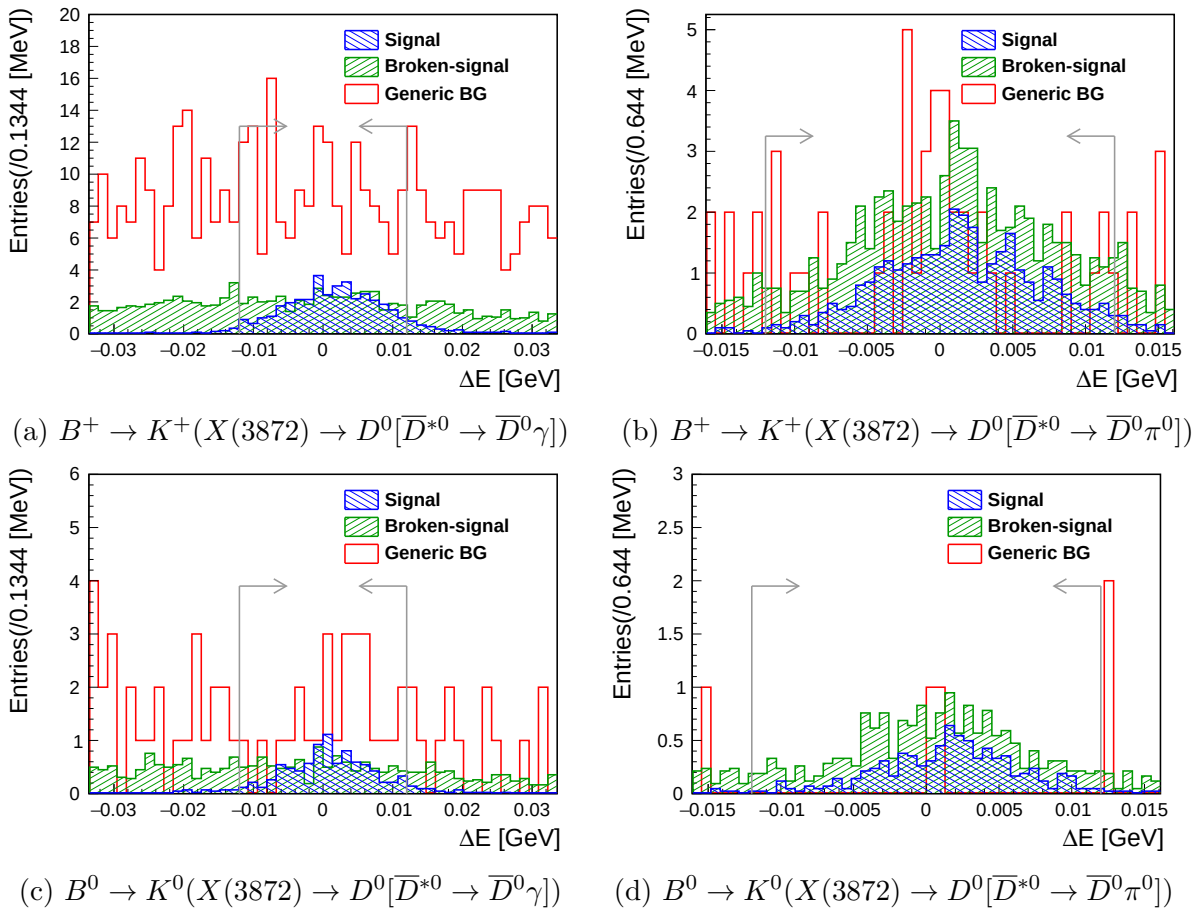


(c) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$



(d) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

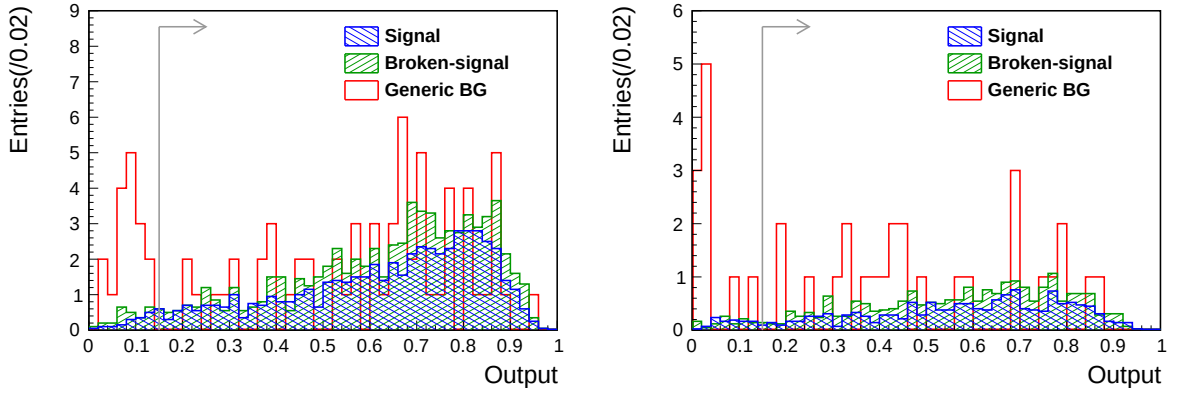
Figure 4.10: The $M_{\rm bc}$ distribution in the $X(3872)$ and B signal regions for the MC samples. The notations are the same as in Fig. 4.6.



(a) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (b) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

(c) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (d) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

Figure 4.11: The ΔE distribution in the $X(3872)$ and B signal regions for the MC samples. The notations are the same as in Fig. 4.6.



(a) The B^+ decay mode (b) The B^0 decay mode

Figure 4.12: The distribution of the FastBDT output of the continuum suppression in the $X(3872)$ and B signal regions for the MC samples. The notations are the same as in Fig. 4.6.

4.2.6 Best Candidate Selection

After this selection, the average number of signal B candidates per event is 1.8, because $D^0\overline{D}^{*0}$ and $D^{*0}\overline{D}^0$ are often indistinguishable and double-counted. To avoid multiple counting of signal events, we select the candidate that has the highest value of the product of the following likelihood $\mathcal{L}$ and prior probability $\mathcal{P}$. The prior probability is a measure to select a candidate based on expected branching fractions and reconstruction efficiency.

The likelihood $\mathcal{L}$ is the product of the likelihoods of the measured D^0 mass, $\overline{D}^0$ mass, $\overline{D}^{*0}$ mass, and ΔE; and, for the $\overline{D}^{*0} \to \overline{D}^0\pi^0$ mode, the likelihood of the measured π^0 mass,

$$\mathcal{L} = \mathcal{L}_{M(D^0)} \times \mathcal{L}_{M(\overline{D}^0)} \times \mathcal{L}_{M(\overline{D}^{*0})-M(\overline{D}^0)} \times \mathcal{L}_{\Delta E}\ [\times\mathcal{L}_{M(\pi^0)}]. \tag{4.2}$$

Each likelihood is obtained using PDFs of double Gaussian with a common mean, which are determined using the default signal MC sample. The PDF parameters are summarized in Table 4.1.

The probability $\mathcal{P}$ is obtained from the probability that a signal event can be reconstructed ε_{ijk}, the average number of the B candidates per single event ζ_{ijk}, and the branching fractions, when D^0, $\overline{D}^0$, and $\overline{D}^{*0}$ are reconstructed from i, j, and k modes, respectively,

$$\mathcal{P} = \frac{\varepsilon_{ijk}}{\zeta_{ijk}} \times \mathcal{B}(D^0 \to i) \times \mathcal{B}(\overline{D}^0 \to j) \times \mathcal{B}(\overline{D}^{*0} \to k). \tag{4.3}$$

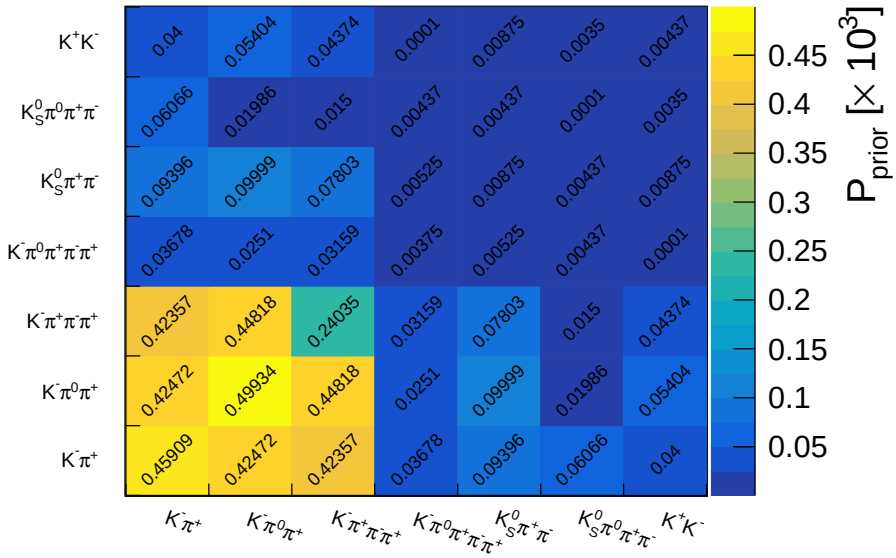
The values of ε_{ijk} and ζ_{ijk} are determined using the default signal MC sample. The $\mathcal{P}$ values are summarized in Fig. 4.13.

For the validation, the efficiency is evaluated and compared with the best candidate selection using the likelihood only. The efficiency is derived as

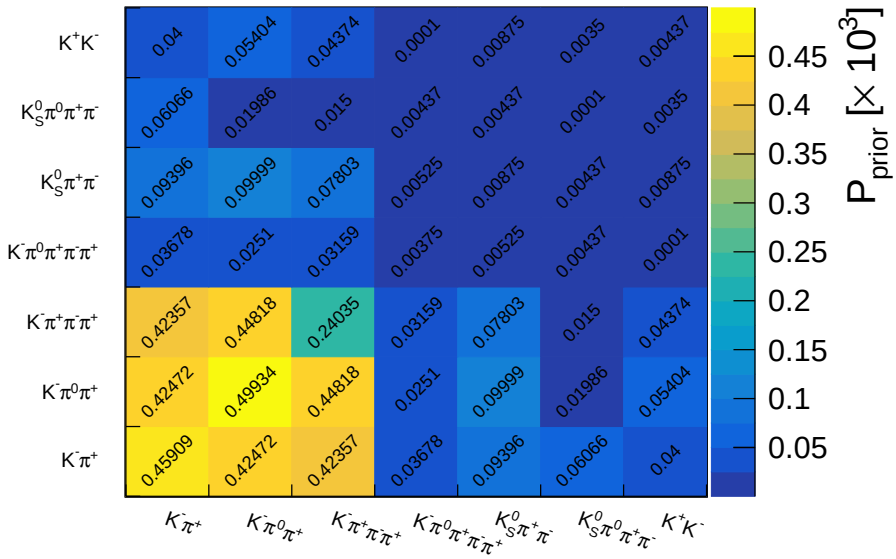
$$\frac{\text{(The number of signal events after BCS)}}{\text{(The number of events including a correctly reconstructed signal before BCS)}}.$$

Table 4.2 shows the efficiency for each B mode. The signal efficiency is almost the same for the B^+ mode. For the B^0 mode, it is improved by several percent. Therefore, we confirm that this method works well for both B modes.

The $(M_{\rm bc}, \Delta E)$ distribution of the selected B candidates is shown in Fig. 4.14. The red solid (blue dashed) rectangle shows the $(M_{\rm bc}, \Delta E)$ signal region for $\overline{D}^{*0} \to \overline{D}^0\gamma$ ($\overline{D}^0\pi^0$); B candidates used in the lineshape study are selected from this region.



(a) $\overline{D}^{*0} \to \overline{D}^0\gamma$



(b) $\overline{D}^{*0} \to \overline{D}^0\pi^0$

Figure 4.13: The $\mathcal{P}$ value of the best candidate selection for each D^0 (vertical axis) and $\overline{D}^0$ mode (horizontal axis).

Table 4.1: Parameters of the double Gaussian with a common mean for the resolution functions used in the best candidate selection. The smaller and wider standard deviation are denoted by $\sigma_{\rm core}$ and $\sigma_{\rm tail}$. The fraction represents that to the core component.

Variable	Mode	$\sigma_{\rm core}$ [MeV/c^2]	$\sigma_{\rm tail}$ [MeV/c^2]	Fraction
$M(D^0),\ M(\overline{D}^0)$	$D^0 \to K^-\pi^+$	4.32	9.91	0.799
	$D^0 \to K^-\pi^+\pi^0$	6.64	17.3	0.493
	$D^0 \to K^-\pi^+\pi^-\pi^+$	4.20	10.4	0.723
	$D^0 \to K^-\pi^+\pi^-\pi^+\pi^0$	6.61	15.9	0.675
	$D^0 \to K_S\pi^+\pi^-$	4.00	11.6	0.769
	$D^0 \to K_S\pi^+\pi^-\pi^0$	8.35	28.1	0.834
	$D^0 \to K^+K^-$	4.11	18.9	0.893
$M(\overline{D}^{*0}) - M(\overline{D}^0)$	$\overline{D}^{*0} \to \overline{D}^0\gamma$	3.96	5.69	0.273
	$\overline{D}^{*0} \to \overline{D}^0\pi^0$	0.733	2.02	0.670
$M(\pi^0)$	$\overline{D}^{*0} \to \overline{D}^0\pi^0$	4.63	10.2	0.408
ΔE	$\overline{D}^{*0} \to \overline{D}^0\gamma$	6.70	17.0	0.903
	$\overline{D}^{*0} \to \overline{D}^0\pi^0$	5.66	19.3	0.926

Table 4.2: The efficiency of the best candidate selection (BCS) for each B mode.

Signal mode	BCS with $\mathcal{L}$	BCS with $\mathcal{L} \times \mathcal{P}$
$B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$	75.7%	75.4%
$B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma))$	80.1%	80.1%
Total for B^+ mode	77.9%	77.9%
$B^0 \to K^0_S(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$	73.3%	75.1%
$B^0 \to K^0_S(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$	74.4%	80.9%
Total for B^0 mode	73.9%	78.2%

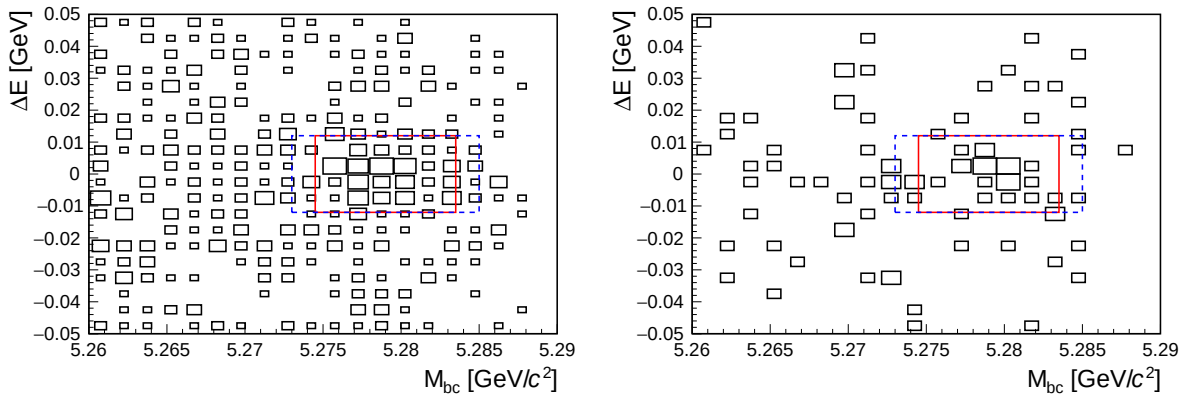


Figure 4.14: Distributions of $(M_{\rm bc}, \Delta E)$ for B^+ (left) and B^0 (right) candidates in the $M(D^0\overline{D}^{*0}) < 3.88$ GeV/c^2 region, where the signal-to-background ratio for $X(3872)$ in data is relatively high. The red solid and blue dashed rectangles show the $(M_{\rm bc}, \Delta E)$ signal regions for $\overline{D}^{*0} \to \overline{D}^0\gamma$ and $\overline{D}^0\pi^0$, respectively.

Table 4.3: Summary of signal, broken-signal, and generic background yields for each decay mode when assuming the mass and width of the relativistic Breit-Wigner lineshape, and branching fraction measured in Ref. [2]. Values in parentheses are products of the efficiency and the branching fraction of the intermediate states.

Mode	Signal	Broken-signal	generic BG
$B^+ \to K^+ D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$	22.5 (3.54×10^{-4})	23.0 (3.62×10^{-4})	731
$B^+ \to K^+ D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$	21.3 (3.35×10^{-4})	24.0 (3.77×10^{-4})	2545
$B^0 \to K^0_S D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$	5.40 (9.00×10^{-5})	5.36 (8.93×10^{-5})	120
$B^0 \to K^0_S D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$	5.23 (8.71×10^{-5})	5.55 (9.24×10^{-5})	343
Total	54.4 (8.66×10^{-4})	57.8 (9.21×10^{-4})	3739

4.3 Observable and Expected Distribution

In measuring the $X(3872)$ lineshape, we use the $D^0\overline{D}^{*0}$ invariant mass distributions in the region below 4.0 GeV/c^2 for all events surviving in the selection. The $D^0\overline{D}^{*0}$ invariant mass is calculated subtracting the reconstructed $\overline{D}^{*0}$ mass instead of applying a mass-constrained fit to improve the mass resolution:

$$M(D^0\overline{D}^{*0}) = \begin{cases} M(D^0\overline{D}^0\gamma) - M(\overline{D}^0\gamma) + m_{\overline{D}^{*0}} & \text{for } \overline{D}^{*0} \to \overline{D}^0\gamma, \\ M(D^0\overline{D}^0\pi^0) - M(\overline{D}^0\pi^0) + m_{\overline{D}^{*0}} & \text{for } \overline{D}^{*0} \to \overline{D}^0\pi^0. \end{cases} \tag{4.4}$$

Here $M(D^0\overline{D}^0\gamma[\pi^0])$ and $M(\overline{D}^0\gamma[\pi^0])$ are the $D^0\overline{D}^0\gamma[\pi^0]$ and $\overline{D}^0\gamma[\pi^0]$ invariant mass respectively, where the particles without [with] the square brackets are the specification of the $\overline{D}^{*0} \to \overline{D}^0\gamma$ [$\overline{D}^{*0} \to \overline{D}^0\pi^0$] mode. Because the D^{*0} width is predicted to be nonzero (65.5 ± 6.2 keV [41]), the techniques of the resolution improvement cause small bias. To make the bias more tractable, the mass difference technique is adopted. The bias is considered as a systematic uncertainty and is discussed in Chapter 6.

The expected distributions of the $D^0\overline{D}^{*0}$ invariant mass are shown in Fig. 4.15, in which the mass, the width, and the branching fractions are assumed to be those measured in Ref. [2]. The expected peak yield under the assumption is summarized in Table 4.3. Here, methods for making the shape function of the peak and background and efficiency determinations are explained in the next chapter. In the previous study, the peak yield was $50.1^{+14.8}_{-11.1}$, while it is expected to be 112.2 in this study, where the signal and broken-signal contributions are included. The improvement is a factor of 2.2. The main contribution is adding the D^0 decay modes. For example, whereas at least on $D^0(\overline{D}^0)$ is required to decay to $K^-\pi^+(K^+\pi^-)$ in the previous studies to reduce the background, we remove the requirement, add further $D^0 \to K^0_S\pi^+\pi^-\pi^0$ mode, and re-optimize the event selection to improve FOM. The signal-to-background ratio is about two times worse than that in the previous study. Since the background contribution near the peak is small, sensitivity is expected to improve. The concrete improvement of statistical uncertainty is evaluated using pseudo-experiments with the relativistic Bright-Wigner lineshape (Chapter 5.4).

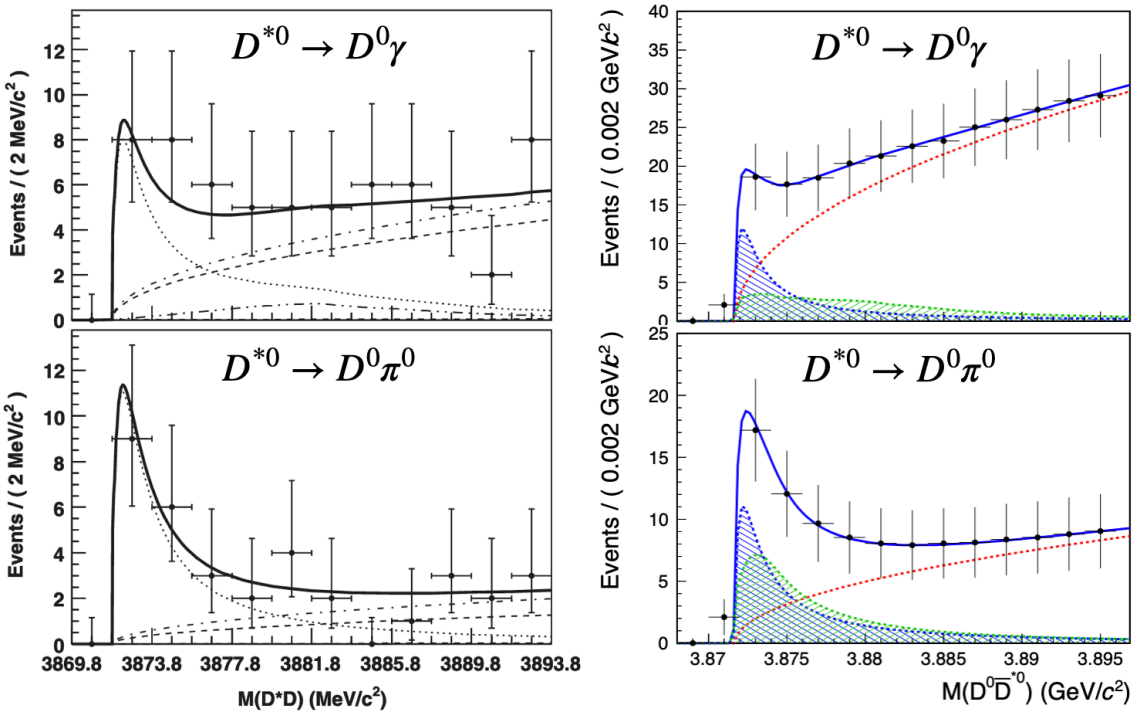


Figure 4.15: (Left) The distributions of the $D^0\overline{D}^{*0}$ invariant mass with the data of 605 fb^{-1} in Ref. [2]. (Right) The expected distributions of the $D^0\overline{D}^{*0}$ invariant mass in this study with the full data (711 fb^{-1}). To suppress the fluctuation of the distributions, a distribution with a thousand times yields is generated and scaled. The top and bottom rows are the $D^{*0} \to D^0\gamma$ and $D^0\pi^0$ modes, respectively. For each panel, distributions of B^0 and B^+ are always combined.

Chapter 5

Method of Signal Extraction

When signal events are reconstructed correctly, the invariant mass distribution has a peak consisting of the natural lineshape convolved with the mass-dependent detector response. To extract the natural lineshape, we perform fits in this analysis. Since the signal-to-background ratio depends on the D^{*0} decay mode, fits are performed separately for $D^{*0} \to D^0\gamma$ and $D^{*0} \to D^0\pi^0$. In addition, fits are performed separately for B^0 and B^+ candidates to determine the relative branching fraction between $B^0 \to X(3872)K^0$ and $B^+ \to X(3872)K^+$.

The distribution of the $D^0\overline{D}^{*0}$ invariant mass comprises three types of components. The first component is the correct combination of signal daughter particles, which we refer to as the signal. It is the most sensitive to the natural lineshape. The second component is the incorrectly reconstructed signal; hereafter called the broken-signal. It is also slightly sensitive to the natural lineshape because the final state particles have similar momentum to the signal after applying the selection. The third component is the background coming from continuum events and $B\overline{B}$ events except for the signals, hereafter called the generic background. It is insensitive to the lineshape. Respective components are discussed in Sec. 5.1–5.3. At the end of this chapter, the fit procedure and its validation are given.

5.1 Signal Reconstructed Correctly

The signal PDF is created by convolving the assumed lineshape with the detector response. For analyses of near-threshold states using a decay mode corresponding to the threshold, the detector response strongly depends on the mass because the detector efficiency is reduced for low-momentum final-state particles and the resolution function generally depends on the Q value. This response, i.e., the mass dependence of the signal efficiency and the mass resolution, is studied and parameterized using the zero-width signal MC samples; see Sec. 5.1.1. The shape determination is described in Sec. 5.1.2. In the fit, the ratio of signal yields between $\overline{D}^{*0} \to \overline{D}^0\gamma$ and $\overline{D}^{*0} \to \overline{D}^0\pi^0$ is fixed using the total signal efficiency to combine the D^{*0} and B sub-samples. The method to evaluate total signal efficiency is discussed in Sec. 5.1.3.

5.1.1 Detector Response

True Mass dependence of the Signal Efficiency

The true mass dependence of the signal efficiency is determined using the zero-width signal MC samples. In what follows, the efficiency always includes branching fractions

of intermediate particles, for example, D^0, D^{*0}, π^0, and K^0_S. Figure 5.1 (a, b) shows the efficiency as a function of the true mass for B^+ mode. It shows that the signal efficiency varies depending on the mass by a few tens percent relatively, especially around the threshold. The following exponential threshold function is used to parameterize the efficiency:

$$\varepsilon_{\text{reco}}(M) = \begin{cases} p_0(1 - p_1 e^{p_2\{M-(m_{D^0}+m_{D^{*0}})\}}) \\ \qquad\qquad + p_3\{M - (m_{D^0} + m_{D^{*0}})\} & \text{for } M \leq 3.920 \text{ GeV}/c^2, \\ \varepsilon_{\text{reco}}(3.920 \text{ GeV}/c^2) & \text{for } M > 3.920 \text{ GeV}/c^2. \end{cases} \tag{5.1}$$

where the case is divided to represent the convergence for large true mass. The threshold of the cases is determined by fitting with $\varepsilon_{\text{reco}}(M)$ floated it. Subsequently, we determine the parameters p_0–p_3 by fitting the relation with the function for the B^+ mode. The resulting values are summarized in Table 5.1.

For the B^0 mode, a similar structure is obtained in the relation between the efficiency and the true mass (Fig. 5.1 (c, d)). We fit it with $\varepsilon_{\text{reco}}(M)$ fixed parameters related to the structure, p_1–p_3. The p_0 value is determined as shown in Table 5.1.

True Mass Dependence of Resolution

The mass resolution depends on the invariant mass, especially in the vicinity of the threshold. We parameterize it using the zero-width signal MC samples. For each input $X(3872)$ mass sample, the resolution PDF is determined to reproduce a distribution for the difference between the $D^0\overline{D}^{*0}$ invariant mass and true $X(3872)$ mass, M_{diff}.

We use the sum of a Gaussian and a Crystal Ball function [68] with a common mean multiplied by a turn-on curve on the threshold as the resolution function f_{res}:

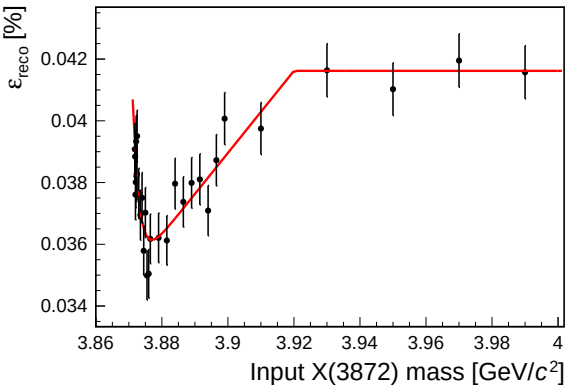
$$\begin{aligned} f_{\text{res}}(M_{\text{diff}}) =& [c \cdot f_{\text{gauss}}(M_{\text{diff}}, \mu, \sigma_{\text{gauss}}) + (1-c) \cdot f_{\text{CB}}(M_{\text{diff}}, \mu, \sigma_{\text{CB}}, n, \alpha)] \\ & \times f_{\text{turn-on}}(M_{\text{diff}}), \end{aligned} \tag{5.2}$$

where f_{gauss} and f_{CB} represent a Gaussian PDF and a Crystal Ball PDF, respectively. μ is a common mean, σ is a standard deviation, n is a value of the exponent of the power law, α is a transition point between the Gaussian and the power-law tail, and c is the fraction. For the turn-on curve $f_{\text{turn-on}}$, the following soft threshold function that rises from zero to one at the threshold using an error function (f_{er}) is used,

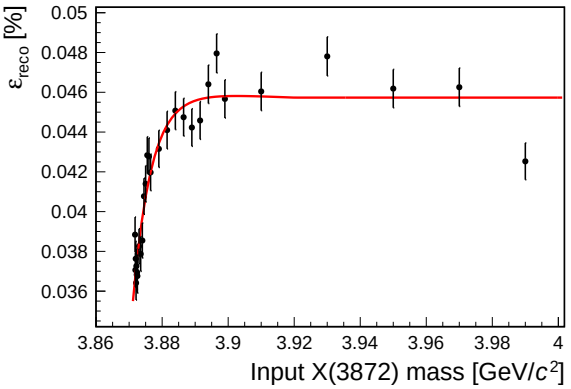
$$f_{\text{turn-on}} = 0.5 \times (f_{\text{er}}(p_0 \cdot M_{\text{diff}}) + 1), \tag{5.3}$$

where p_0 is determined to the best result which reproduces the distribution in the zero-width signal MC samples; the value is fixed to 36065.6. For the standard deviations, they degrade with the square root of the difference between the mass and the threshold; otherwise, since there are no requirements for the other parameters, we need to parameterize them by checking the relation among them.

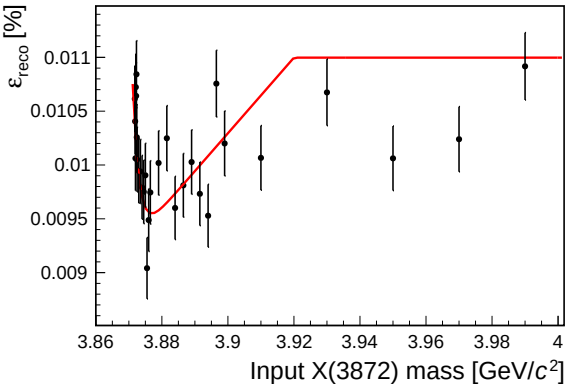
The PDF is parameterized using B^+ mode only because the resolution is determined approximately by the momentum resolution of the $X(3872)$ daughter particles regardless of the B decay modes. The details of the parameterization are described in Appendix C.2. The resulting resolution functions are shown in Fig. 5.2. For lower input masses, e.g., 3.8724GeV/c^2, there is a structure of $\mathcal{O}(0.1)$ MeV/c^2 in the pull distribution. It can be ignored because it is very small compared to the width of the Breit-Wigner lineshape measured by Ref. [2], $\mathcal{O}(10)$ MeV/c^2.



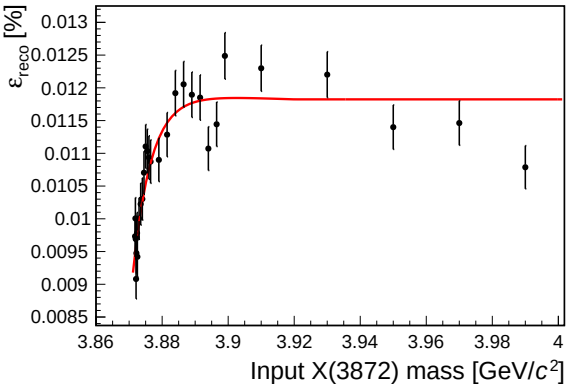
(a) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$



(b) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$



(c) $B^0 \to K^0_S(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$

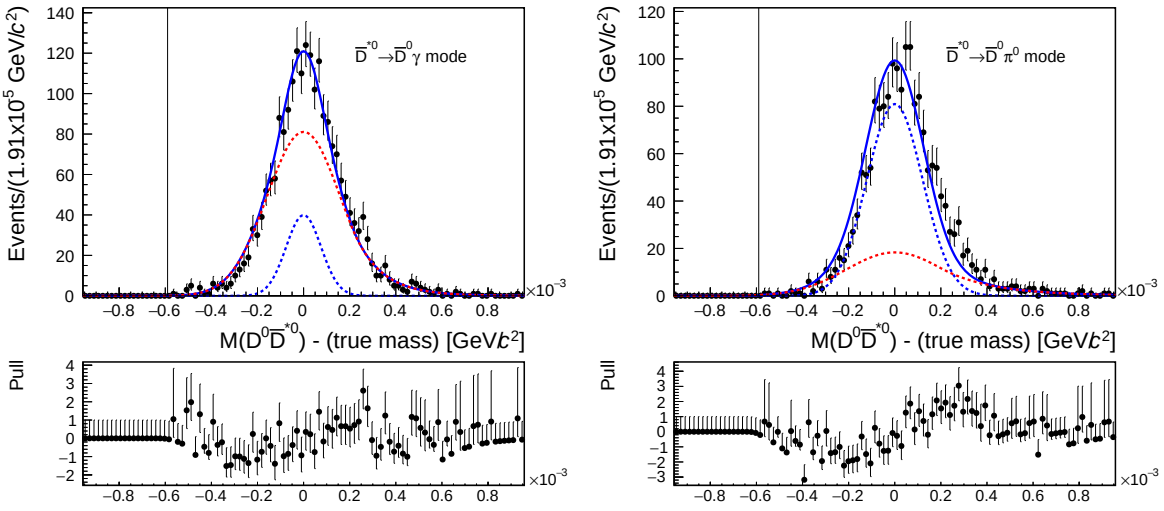


(d) $B^0 \to K^0_S(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

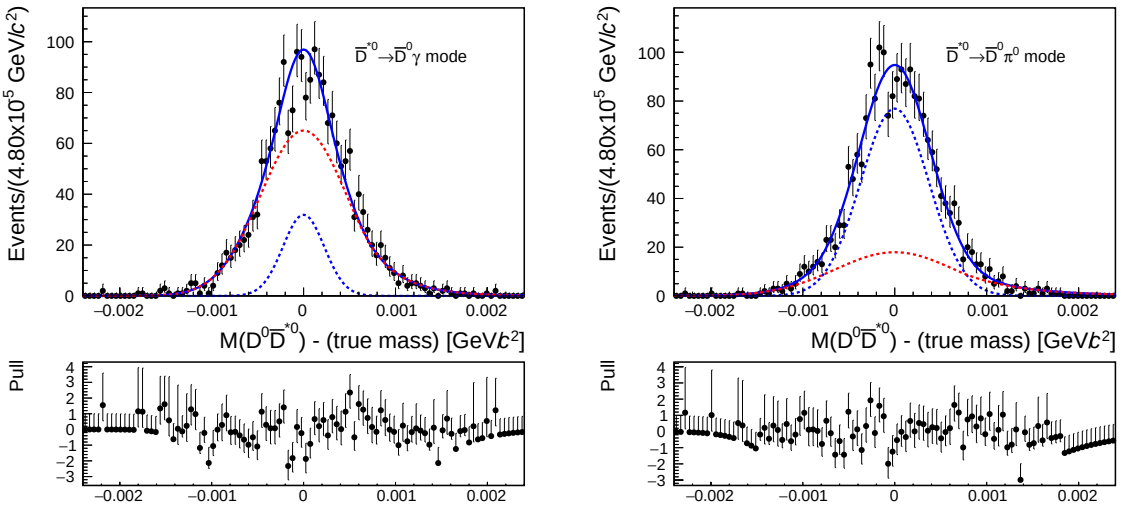
Figure 5.1: The signal efficiency as a function of the true mass with the zero-width signal MC samples. The red line shows the fit result with $\varepsilon_{\text{reco}}(M)$.

Table 5.1: $\varepsilon_{\text{reco}}(M)$ parameters determined by the fits in Fig. 5.1.

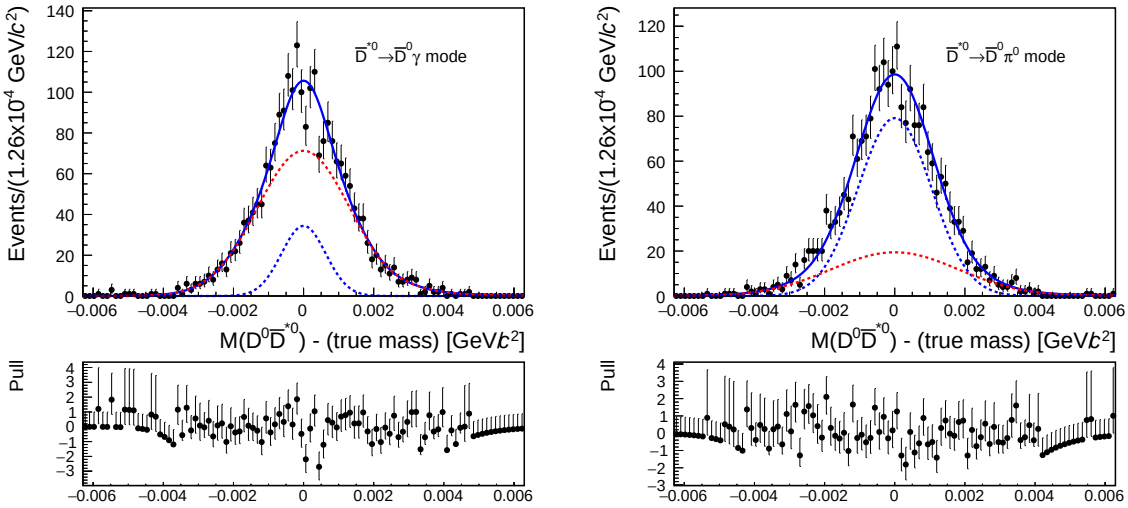
Mode	Parameter	Value
$B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$	p_0	$(3.521 \pm 0.041) \times 10^{-2}$ %
	p_1	$(-1.11 \pm 0.16) \times 10^{-1}$
	p_2	$(54 \pm 16) \times 10$
	p_3	3.78 ± 0.43
$B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$	p_0	$(4.61 \pm 0.10) \times 10^{-2}$ %
	p_1	$(2.04 \pm 0.18) \times 10^{-1}$
	p_2	180 ± 40
	p_3	-0.15 ± 0.60
$B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$	p_0	$(9.304 \pm 0.051) \times 10^{-3}$ %
	p_1	Fixed to -1.11×10^{-1}
	p_2	Fixed to 54×10
	p_3	Fixed to 3.78
$B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$	p_0	$(1.1908 \pm 0.0066) \times 10^{-2}$ %
	p_1	Fixed to 2.04×10^{-1}
	p_2	Fixed to 180
	p_3	Fixed to -0.15



(a) Generated $X(3872)$ mass is 3.8724 GeV/c^2



(b) Generated $X(3872)$ mass is 3.8790 GeV/c^2



(c) Generated $X(3872)$ mass is 3.9300 GeV/c^2

Figure 5.2: The $M_{\rm diff}$ distributions for $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (left side) and $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$ (right side) with the zero-width signal MC samples. In each panel, the solid blue line shows the fit result with the resolution PDF fixed the shape parameters. The dashed blue and red lines show the Crystal Ball and Gaussian components, respectively. The vertical black line represents the $D^0\overline{D}^{*0}$ threshold. The bottom figure of each panel shows the pull distribution.

5.1.2 Determination of Shape

To obtain the signal shape ($f_{\rm sig}$), the true lineshape multiplied by $\varepsilon_{\rm reco}(t)$ is convoluted with the mass-dependent resolution PDF

$$f_{\rm sig} = \int_0^\infty dt f_{\rm res}(M-t) \times (f_{\rm true}(t) \times \varepsilon_{\rm reco}(t)). \tag{5.4}$$

To reduce the computational time, convolution with the specific resolution function at a mass of 3871.9 MeV/c^2 near the peak is adopted as an approximation, because the natural lineshape is broad [2, 3] and the smearing effect due to the resolution is small at masses away from the peak. To reproduce the behavior near the threshold, the signal function is multiplied by $f_{\text{turn-on}}$ after the convolution.

The differences in lineshape between the convolution methods are shown in Figs. 5.3–5.4. The original convolution (dashed red lines) and the approximation (dashed black lines) are in good agreement for both lineshape models. Therefore, the effect of the approximation is negligible.

5.1.3 Total Reconstruction Efficiency

The total efficiency is necessary for not only the measurement of the branching fraction but also the fit constraint on a yield ratio between $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$ and $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$. Because the efficiency depends on the $D^0\overline{D}^{*0}$ invariant mass, the total efficiency is varied by the true lineshape. In addition, since the Flatté lineshape has a long tail and some signal leaks out of the fitting range, the total signal efficiency $\epsilon_{\rm tot}$ needs to include such effects. Thus, the total efficiency is calculated by

$$\epsilon_{\rm tot} = \epsilon_{\rm reco} \times \epsilon_{\rm retain}, \tag{5.5}$$

where $\epsilon_{\rm reco}$ is the signal efficiency for the reconstruction and the selection, and $\epsilon_{\rm retain}$ is the signal efficiency that remains in the fit range without leaking. The calculation method for each efficiency is shown as follows.

$\epsilon_{\rm reco}$; It is derived by an average of the signal efficiency as a function of true mass $\varepsilon_{\rm reco}(M)$ weighted with the assumed lineshape $f_{\rm true}$,

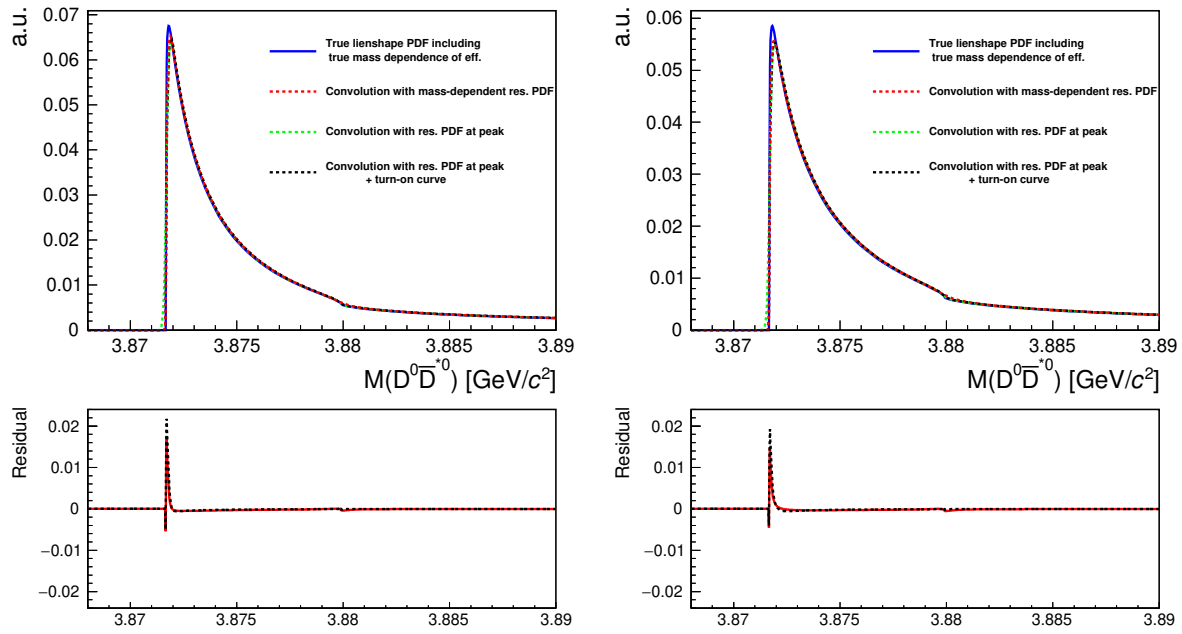
$$\epsilon_{\rm reco} = \frac{\sum_i \varepsilon_{\rm reco}(M_i) \times f_{\rm true}(M_i)}{\sum_i f_{\rm true}(M_i)}, \tag{5.6}$$

where M_i is one of 5000 points taken at equal intervals from the invariant-mass range from $D^0\overline{D}^{*0}$ threshold to $m_B - m_K$.

$\epsilon_{\rm retain}$; It is derived as a ratio of the values obtained by definite integration of lineshape for the fit range over that for the total range,

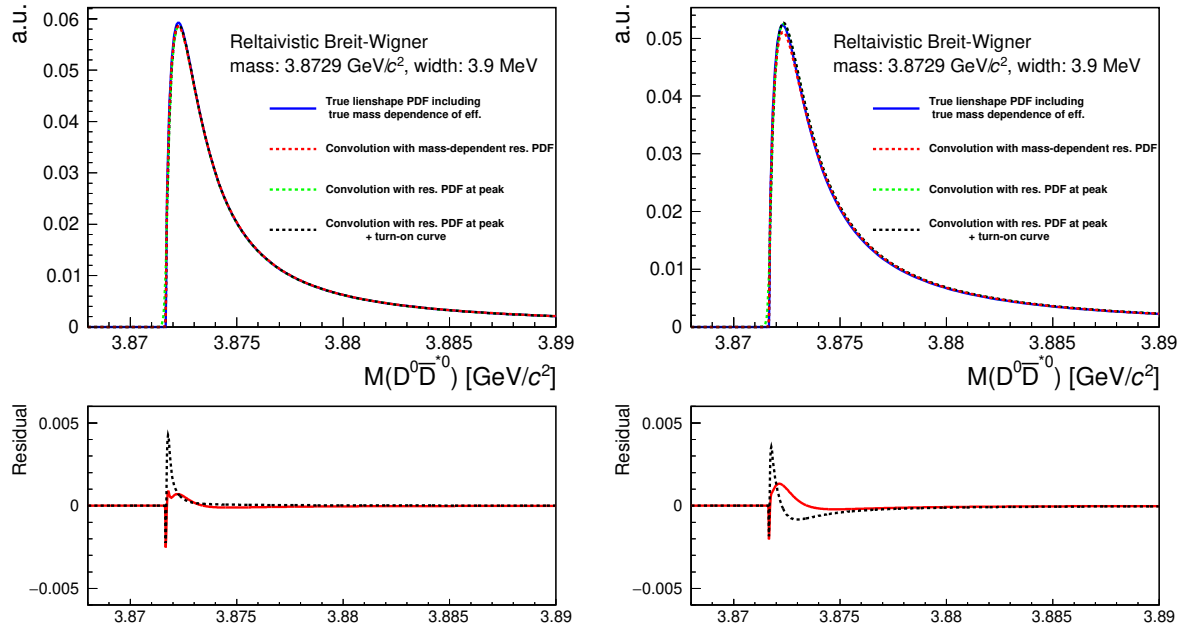
$$\epsilon_{\rm retain} = \frac{\int_{M_{\rm min}}^{M_{\rm max}} f(M)dM}{\int_0^{m_B-m_K} f(M)dM}, \tag{5.7}$$

where $M_{\rm max}$ and $M_{\rm min}$ denote the maximum and minimum of the fit range, respectively.



(a) $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$ (b) $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$

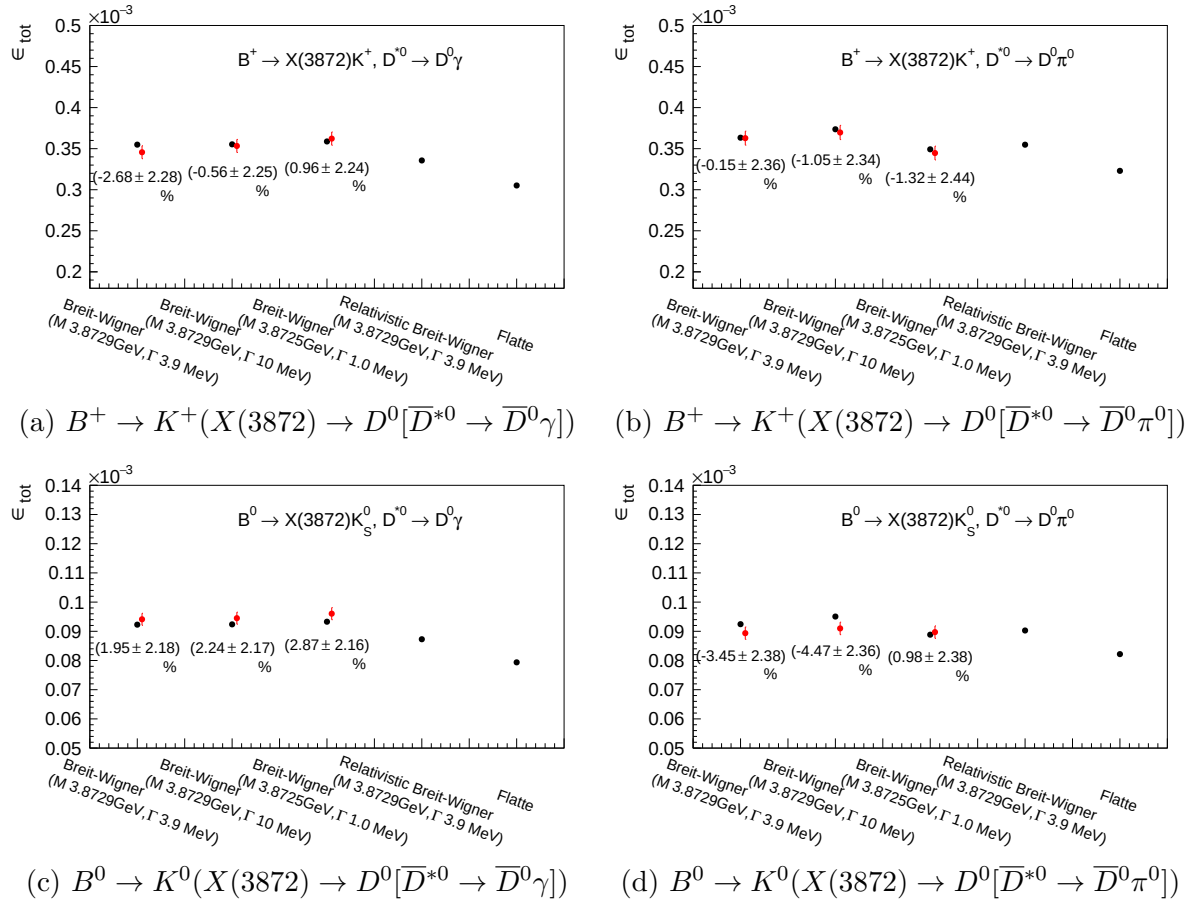
Figure 5.3: A Comparison of lineshapes with different convolution methods in the case of the Flatté lineshape model. The solid blue line shows the assuming lineshapes after considering the efficiency as a function of a true mass. The dashed red line shows the lineshapes convoluted with the mass-dependent PDF. The dashed green line shows the lineshapes convoluted with the mass-independent PDF, and the dashed black line shows the product of the dashed green line and $f_{\text{turn-on}}$. The bottom figure of each panel shows residuals from the assuming lineshape considering the efficiency to the lineshape convoluted with the mass-dependent PDF (solid red line), and that to the product of the lineshape convoluted with the mass-independent PDF and $f_{\text{turn-on}}$ (dashed black line).



(a) $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$ (b) $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$

Figure 5.4: A Comparison of lineshapes with different convolution methods in the case of the relativistic Breit-Wigner lineshape. The notations are the same as in Fig. 5.3.

To verify this calculation method, we compare the efficiency, derived from the ratio of the number of the correctly reconstructed signal events to that of the generated events. Figure 5.5 shows the total efficiency when the $M(D^0\overline{D}^{*0})$ range is to 4.0 GeV/c^2. In each panel, the first to third points from the left are the results using the finite-width signal MC samples. The efficiencies with the calculation method are consistent with those obtained by counting the number of events. It means that our calculation method works well. The fourth and fifth points from the left are examples of the efficiency calculated for the relativistic Breit-Wigner lineshape and the Flatté lineshape. The method is used for the later analysis.



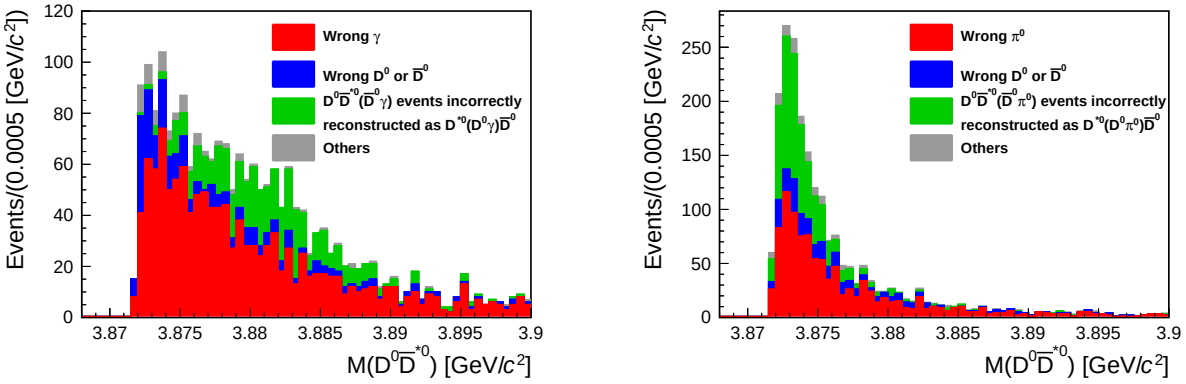
(a) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (b) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

(c) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (d) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

Figure 5.5: Total signal efficiency for five lineshapes when the $M(D^0\overline{D}^{*0})$ range is to 4.0 GeV/c^2: from left, the signal MC samples with a mass of 3.8729 GeV/c^2 and a width of 3.9 MeV, the signal MC samples with a mass of 3.8729 GeV/c^2 and a width of 10 MeV, the signal MC samples with a mass of 3.8725 GeV/c^2 and a width of 1.0 MeV, the relativistic Breit-Wigner with a mass of 3.8729 GeV/c^2 and a width of 3.9 MeV, and the Flatté lineshape with the parameters of Eq. (2.20). In each panel, the black points represent the calculation method results, and the red points represent the results obtained by counting the number of events. The values under the points show the relative difference between the methods.

5.2 Signal Reconstructed Incorrectly

The broken-signal contribution is mainly caused by three sources (Fig. 5.6), where a wrong π^0 or γ is combined in the $\overline{D}^{*0}$ reconstruction, the D^0 is reconstructed incorrectly, or a $D^{*0}\overline{D}^0$ signal event is misinterpreted as $D^0\overline{D}^{*0}$ by combining π^0 or γ from D^{*0} incorrectly with the $\overline{D}^0$ to make a fake $\overline{D}^{*0}$. This contribution is slightly sensitive to the natural lineshape. In this section, the PDF of the broken-signal is studied for each D^{*0} mode. Also, the yield of the broken-signal relative to the signal is determined for the fit constraint.



(a) $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$ (b) $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$

Figure 5.6: The histogram stacked for three background sources where a wrong π^0 or γ is combined in the $\overline{D}^{*0}$ reconstruction (red), the D^0 is reconstructed incorrectly (blue), a $D^{*0}\overline{D}^0$ signal event is misinterpreted as $D^0\overline{D}^{*0}$ by combining π^0 or γ from D^{*0} incorrectly with the $\overline{D}^0$ to make a fake $\overline{D}^{*0}$ (green) and the others (gray).

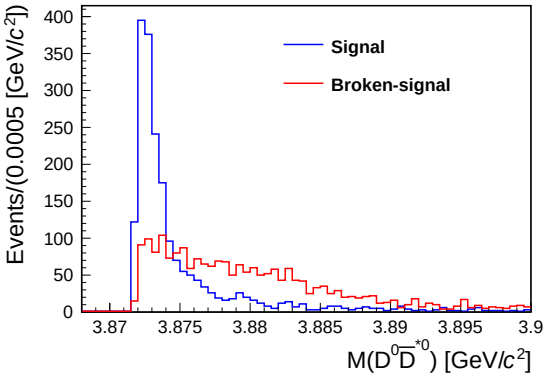


Figure 5.7: The $D^0\overline{D}^{*0}$ invariant mass distributions with the default signal MC samples for $\overline{D}^{*0} \to \overline{D}^0\gamma$. The blue and red histograms show the signal and broken-signal components, respectively.

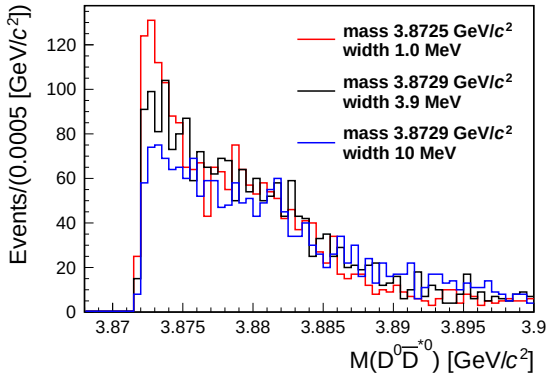


Figure 5.8: The comparison of the distributions of the broken-signal with the finite-width signal MC samples for $\overline{D}^{*0} \to \overline{D}^0\gamma$.

5.2.1 $\overline{D}^{*0} \to \overline{D}^0\gamma$

For $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$, the background peak is much wider than that for the signal, as shown in Fig. 5.7. The comparison of the distributions for the broken-signal with three types of the finite-width signal MC samples (Fig. 5.8) shows that the shape and the yield are similar among them; thus, sensitivity to the natural lineshape is very small. For the fit PDF, the histogram PDF depending on the true lineshape is used to reduce the systematic uncertainty due to the shape. In the fit, we fix the yield of the broken-signal relative to the signal determined from the signal MC samples.

Shape

Since the slight true lineshape dependence on the shape exists (Fig. 5.8), we determine the PDF depending on the assumed lineshape. Because we confirmed that the distribution never changes between B^+ and B^0, the histogram PDF is determined using only B^+ samples. The procedure is as follows.

1. Plot the $M(D^0\overline{D}^{*0})$ distributions of the broken-signal with the zero-width signal MC samples. Here, the bin widths are adjusted to increase as the mass increases so that the statistical fluctuation is suppressed.

2. Scale each histogram by the assumed lineshape. Here, the scale factor for the MC sample with mass of m_i is derived by the integral value from the midpoint between m_{i-1} and m_i to that between m_i and m_{i+1} because the intervals of input masses of the samples are not constant,

$$\int_{(m_{i-1}+m_i)/2}^{(m_i+m_{i+1})/2} f_{\rm true}(\mu)d\mu.$$

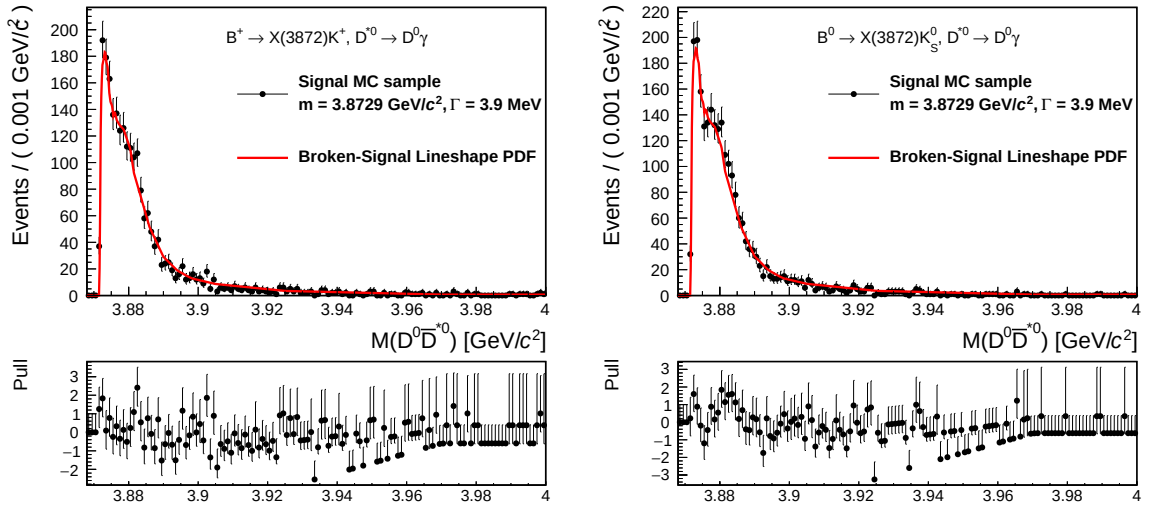
3. Sum up all of the scaled histograms.

4. Generate a histogram PDF from the total histogram with the first-order interpolation.

The validation of this method is done using the finite-width signal MC samples. Figure 5.9 shows the comparison between the distribution obtained from the sample and the histogram PDF assuming the generated lineshape in the sample. The histogram PDF reproduces the distribution well. The lineshape-dependent histogram PDF is used for fitting.

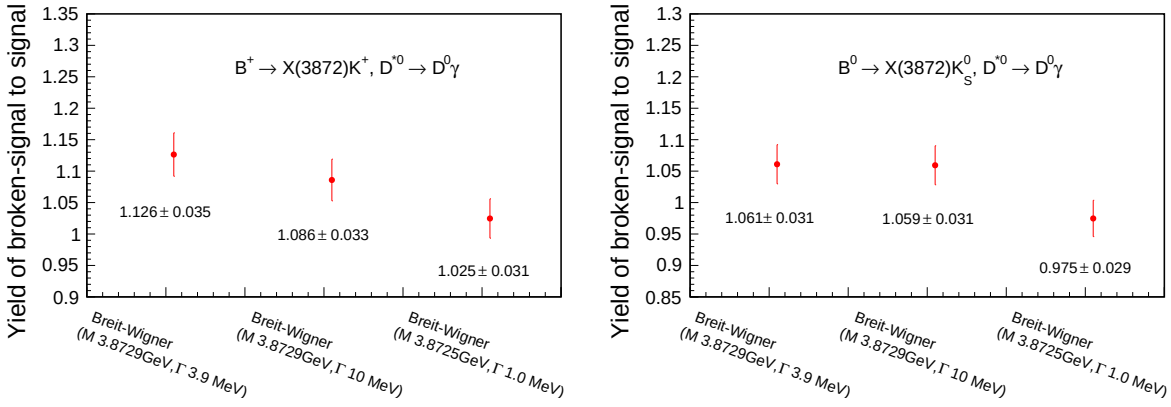
Yield of Broken-Signal Relative to Signal

The dependence of the yield of the broken-signal relative to the signal on true lineshape is studied. A comparison among the finite-width signal MC samples is shown in Fig. 5.10. Since the relative yields are consistent within 1σ, the impact of true lineshape dependence is small. In the fit, the relative yields for both B modes are fixed by the values for the default signal MC sample.



(a) $B^+ \to X(3872)K^+$ (b) $B^0 \to X(3872)K^0$

Figure 5.9: Comparison between the $D^0\overline{D}^{*0}$ invariant mass distribution of the broken-signal (black points) for $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$ and the histogram PDF assuming the generated lineshape in the default signal MC samples (red line). The bottom plot in each panel shows the pull distribution.



(a) $B^+ \to X(3872)K^+$ (b) $B^0 \to X(3872)K^0$

Figure 5.10: The yield of broken-signal relative to signal with three types of the finite-width signal MC samples. A value under the point shows the value for each ratio.

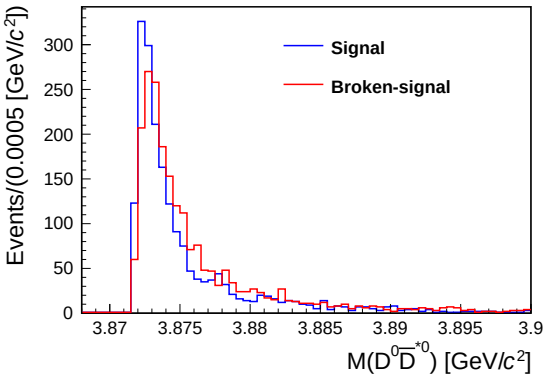


Figure 5.11: The $D^0\overline{D}^{*0}$ invariant mass distributions with the default signal MC samples for $\overline{D}^{*0} \to \overline{D}^0\pi^0$. The blue and red histograms show the signal and broken-signal components, respectively.

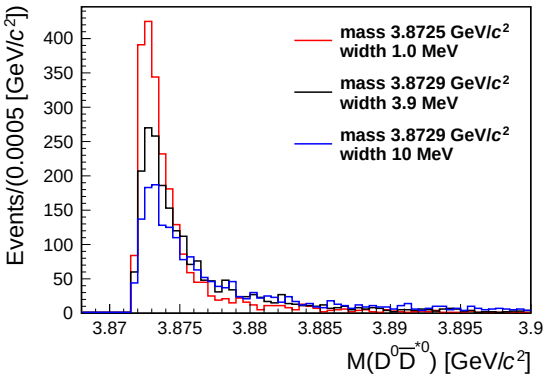


Figure 5.12: The comparison of the distributions of the broken-signal with the finite-width signal MC samples for $\overline{D}^{*0} \to \overline{D}^0\pi^0$.

5.2.2 $\overline{D}^{*0} \to \overline{D}^0\pi^0$

For $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$, the broken-signals produce a broad peak in the $M(\overline{D}^{*0}D^0)$ signal region and possibly distort the lineshape of the signal (Fig. 5.11). The distribution for the broken-signal depends on the true lineshape because the final state particles have similar momentum to the correctly reconstructed signal after applying the selection. Therefore, we prepare a shape and a broken-signal yield that depend on the true lineshape as with the signal.

True Mass Dependence of the Efficiency

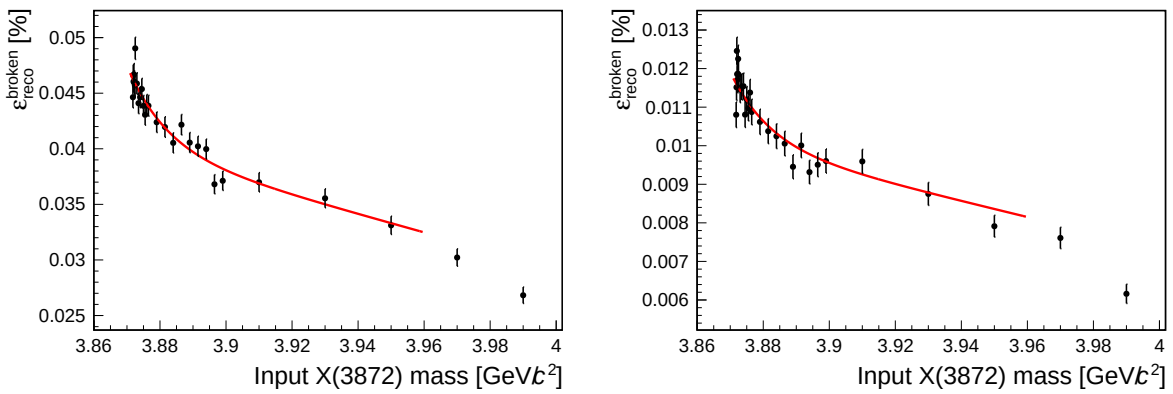
As with the signal, the mass dependence of the efficiency of the broken-signal events is studied. We call it *broken-signal efficiency*, and it is used to determine the PDF shape and the fraction of the background component in the fitting. The definition is

$$\frac{\text{(The number of broken events after reconstruction and selection)}}{\text{(The number of generated events)}}.$$

Figure 5.13 shows the efficiency as a function of the input mass using the zero-width signal MC samples. The dependence is parameterized with a similar threshold function to the signal. The function $\varepsilon_{\rm reco}^{\rm broken}$ is defined as

$$\varepsilon_{\rm reco}^{\rm broken}(M) = p_0(1 - p_1 e^{p_2\{M-(m_{D^0}+m_{D^{*0}})\}}) + p_3\{M - (m_{D^0} + m_{D^{*0}})\}. \qquad (5.8)$$

To determine p_0–p_3, the efficiency as a function of the true mass is firstly fitted with $\varepsilon_{\rm reco}^{\rm broken}(M)$ for the B^+ mode. The red line in Fig. 5.13 (a) shows the fit result, and $\varepsilon_{\rm reco}^{\rm broken}$ can reproduce the dependence. For the high mass region, the efficiency of the MC sample is expected to be lower than $\varepsilon_{\rm reco}^{\rm broken}$. The cause is leakage into the range above 4.00 GeV/c^2. In practice, this effect can be canceled because some events where the true mass is greater than 4.00 GeV/c^2 can flow into the region. Therefore, we ignore the high mass region from 3.96 GeV/c^2 in the fit. For the B^0 mode, the dependence is the same



(a) $B^+ \to X(3872)K^+$ (b) $B^0 \to X(3872)K^0$

Figure 5.13: The broken-signal efficiency as a function of the true mass with the zero-width signal MC samples for $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$. The red line shows the fit result with $\varepsilon^{\text{broken}}_{\text{reco}}(M)$. The fit range is from the $D^0\overline{D}^{*0}$ threshold to 3.96 GeV/c^2.

Table 5.2: $\varepsilon^{\text{broken}}_{\text{reco}}(M)$ parameters determined by the fits in Fig. 5.13.

Mode	Parameter	Value
$B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$	p_0	$(3.98 \pm 0.22) \times 10^{-2}$ %
	p_1	$(-1.63 \pm 0.60) \times 10^{-1}$
	p_2	82 ± 42
	p_3	-2.08 ± 0.73
$B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$	p_0	$(9.977 \pm 0.058) \times 10^{-3}$ %
	p_1	Fixed to -1.63×10^{-1}
	p_2	Fixed to 82
	p_3	Fixed to -2.08

as that for the B^+ mode, as shown in Fig. 5.13 (b). The parameter p_0 is determined from the fit fixing p_1–p_3 to the values obtained in the B^+ mode. The resulting parameters are summarized in Table 5.2.

Shape

The shape is determined with a similar method as the signal PDF. In what follows, the $M(D^0\overline{D}^{*0})$ spread at a certain true mass is called *broken-signal resolution*. The broken-signal resolution PDF is parameterized by the true mass so that the true lineshape is convoluted with the broken-signal resolution PDF.

The broken-signal resolution PDF $f^{\text{broken}}_{\text{res}}$ is determined by using the zero-width signal MC samples. We use a triple-Gaussian with a common mean as the PDF, which is multiplied by the turn-on curve to reproduce the behavior at the threshold. The resolution is parameterized using the MC samples of B^+ only, because no background sources originate from the B mode, i.e., the broken-signal resolution of the B^+ and B^0 modes are expected to be the same. Details of the parameterization are summarized in Appendix C.3. Figure 5.14 shows the parameterized broken-signal resolution PDFs. They reproduce the $M(D^0\overline{D}^{*0})$ distributions of the broken-signal for several choices of the input mass. Some difference is found for the low input mass, e.g., 3.8724 GeV/c^2. We ignore it since we

confirm it does not affect the lineshapes after the convolution.

Finally, the resolution convolution is performed to determine the shape. To reduce the computational time, a discrete convolution is adopted as an alternative method because the convolution performance is adequate for the poor resolution. The convolution can be represented by the sum of the broken-signal resolution PDFs weighted by the broken-signal efficiency and the true lineshape,

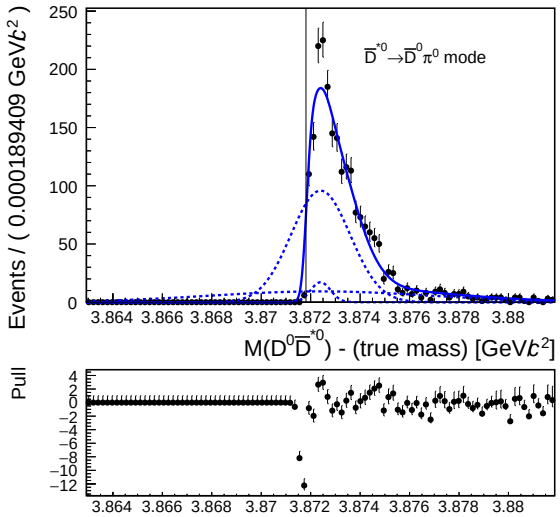
$$f_{\text{broken}} = \left(\sum_{\mu_i} f_{\text{res}}^{\text{broken}}(M - \mu_i, \sigma(\mu_i)) \times \varepsilon_{\text{reco}}^{\text{broken}}(M) \times f_{\text{true}}(\mu_i) \right) \times f_{\text{turn-on}}(M_{\text{diff}}), \quad (5.9)$$

where μ_i is a set of arbitrary masses. To obtain a smooth line, we select five hundred points at intervals at 0.8 MeV/c^2 from the $D^0\overline{D}^{*0}$ threshold as μ_i. The largest point is about 4.272 GeV/c^2.

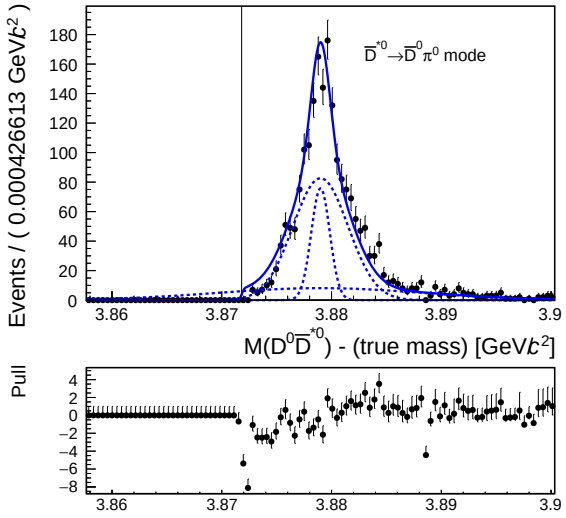
To verify the method, we compare the shape with the distribution obtained from the default signal MC sample in Fig. 5.15. The pull distributions show no difference for both B modes, and the lineshapes can reproduce the distributions well; therefore, the procedure is reasonable.

Total Broken-Signal Efficiency for Each Lineshape

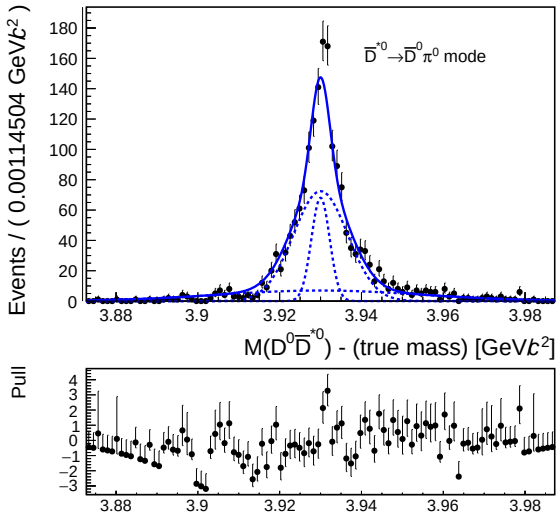
The method to calculate the total broken-signal efficiency is the same as the signal (Sec. 5.1.3). To verify it, we compare the results with those obtained from the ratio of the number of broken-signal events to that of generated events (Fig. 5.16). The dependence on the true lineshape is well reproduced, but there are differences in some cases. The largest difference is 1.7σ in the case using the signal MC sample with small width. The efficiency is used for the final fit, but the deviation is considered as a systematic uncertainty.



(a) Generated $X(3872)$ mass is 3.8724 GeV/c^2

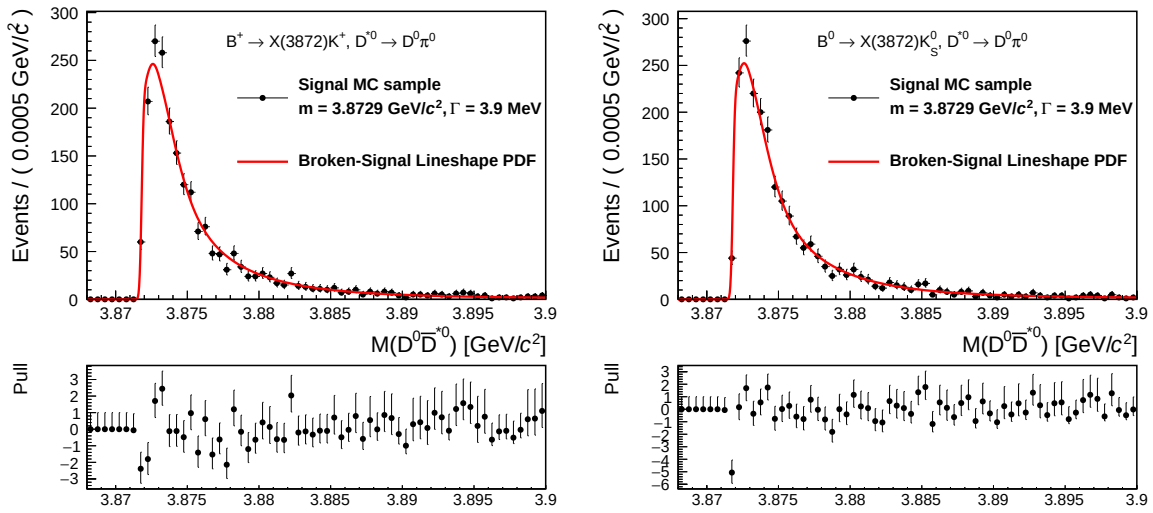


(b) Generated $X(3872)$ mass is 3.8790 GeV/c^2



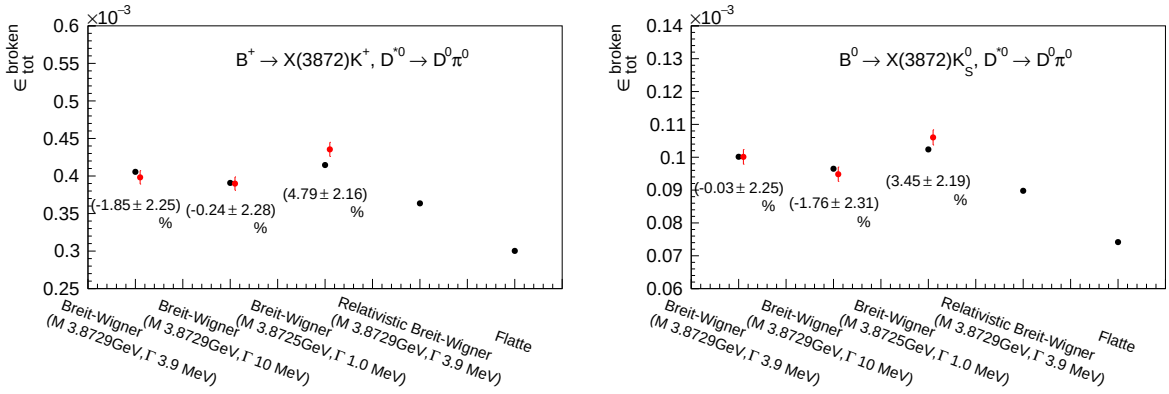
(c) Generated $X(3872)$ mass is 3.9300 GeV/c^2

Figure 5.14: The $M(D^0\overline{D}^{*0})$ distributions of the broken-signal for $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$ with the zero-width signal MC samples. The solid blue lines are the total fit results. The dashed blue lines show the components of Gaussians. The vertical black line represents the $D^0\overline{D}^{*0}$ threshold. In each panel, the bottom plot shows the pull distributions.



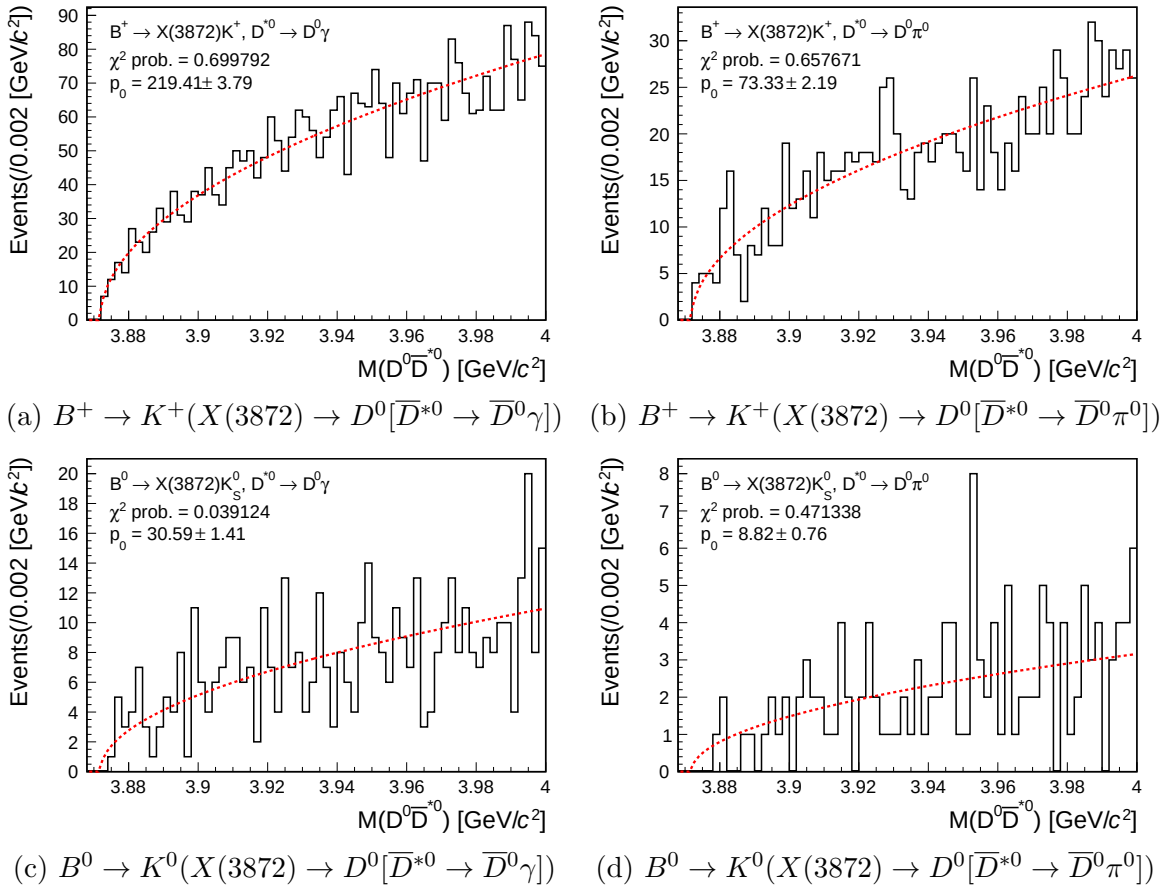
(a) $B^+ \to X(3872)K^+$ (b) $B^0 \to X(3872)K^0$

Figure 5.15: Comparison between the $D^0\overline{D}^{*0}$ invariant mass distribution of the broken-signal (black points) and the lineshape with the discrete convolution (red line) with the default signal MC samples. The bottom plot in each panel shows the pull distribution.



(a) $B^+ \to X(3872)K^+$ (b) $B^0 \to X(3872)K^0$

Figure 5.16: Total broken-signal efficiency for five lineshapes when the $M(D^0\overline{D}^{*0})$ range is to 4.0 GeV/c^2: from left, the signal MC samples with a mass of 3.8729 GeV/c^2 and a width of 3.9 MeV, the signal MC samples with a mass of 3.8729 GeV/c^2 and a width of 10 MeV, the signal MC samples with a mass of 3.8725 GeV/c^2 and a width of 1.0 MeV, the relativistic Breit-Wigner with a mass of 3.8729 GeV/c^2 and a width of 3.9 MeV and the Flatté lineshape with the parameters of Eq. (2.20). In each panel, black points represent the calculation method results, and red points represent the results obtained by counting the number of events. The values under the points show the relative difference between the methods.



(a) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (b) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

(c) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (d) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

Figure 5.17: The distributions of the $D^0\overline{D}^{*0}$ invariant mass for the generic background using the background MC samples. In each panel, the red dashed line shows the fit result with the threshold function.

5.3 Generic background

In this section, we check the generic background and determine the shape and yields expected for the data. The latter is necessary for fit validation with pseudo experiments.

5.3.1 Shape

The number of background events increases with the phase space, as shown in the background distribution of the $D^0\overline{D}^{*0}$ invariant mass in Fig. 5.17. The shape can be reproduced by the threshold function (red dashed line in Fig. 5.17)

$$f_{\mathrm{genBG}}(M(D^0\overline{D}^{*0})) = p_0\sqrt{M(D^0\overline{D}^{*0}) - p_1}, \tag{5.10}$$

where p_0 is a normalization factor, and p_1 is the threshold. p_1 can be deviated slightly from the $D^0\overline{D}^{*0}$ threshold due to the effect of the mass resolution. Since it is difficult to estimate precisely, it is fixed by the threshold.

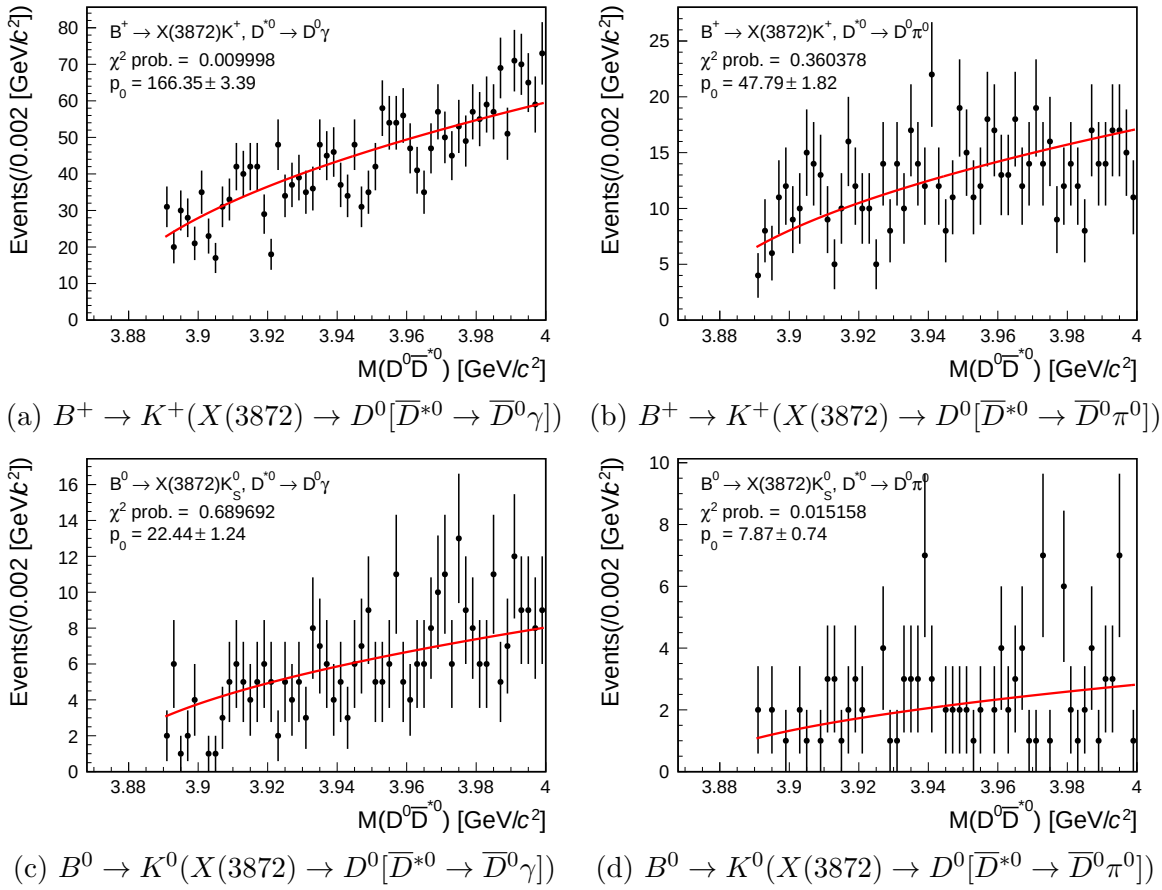
Table 5.3: Summary of the number of generic background events in the background MC sample, the scaling parameter, and the yield expected in the data.

Mode	Parameter	Value
$B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$	Yield in MC	1122
	p_0 (data)	47.8 ± 1.8
	p_0 (MC)	73.3 ± 2.2
	Scale factor	0.652 ± 0.031
	Yield expected for data	$1122 \times 0.652 = 731$
$B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$	Yield in MC	3357
	p_0 (data)	166.4 ± 3.4
	p_0 (MC)	219.4 ± 3.8
	Scale factor	0.758 ± 0.020
	Yield expected for data	$3357 \times 0.758 = 2545$
$B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$	Yield in MC	135
	p_0 (data)	7.87 ± 0.74
	p_0 (MC)	8.82 ± 0.76
	Scale factor	0.89 ± 0.11
	Yield expected for data	$135 \times 0.89 = 120$
$B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$	Yield in MC	468
	p_0 (data)	22.4 ± 1.2
	p_0 (MC)	30.6 ± 1.4
	Scale factor	0.734 ± 0.053
	Yield expected for data	$468 \times 0.734 = 343$

5.3.2 Yield Expected for the Data

The number of background events for the background MC samples may differ from that for the data because the branching fractions of B decay assumed for the background MC sample are not the same as the nature. Therefore, the yield expected for the data is calculated by a product of the yield in the background MC sample and the scaling factor of the data to the MC sample.

The yield in the fit region, 3.868 GeV/c^2 < $M(D^0\overline{D}^{*0})$ < 4.000 GeV/c^2, is determined with the background MC samples, summarized in Table 5.3. Next, we compare the yields between the data and MC samples to determine the scaling factors. To avoid artificial bias, the signal region of the data is hidden until the analysis procedure is completely validated; thus, the data sideband region, 3.890 GeV/c^2 < $M(D^0\overline{D}^{*0})$ < 4.000 GeV/c^2, is used in this study. Figure 5.18 shows the distribution of the generic background for the data sideband. Compared with Fig. 5.17, the number of background events in the data is less, although the shapes of the distributions are consistent. The scale factor is determined by fitting both distributions of the data and MC samples and taking the ratio of the variables corresponding to the area, p_0; the resulting p_0 values are summarized in Table 5.3. Finally, the scale factors for $B^+ \to K^+D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$ modes and $B^+ \to K^+D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$ are obtained as 0.652 and 0.758, respectively. For the B^0 mode, the scale factors for $B^0 \to K^0_S D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$ modes and $B^0 \to K^0_S D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$ are obtained as 0.89 and 0.734, respectively.



(a) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (b) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

(c) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (d) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

Figure 5.18: The distributions of the $D^0\overline{D}^{*0}$ invariant mass for the generic background using the data sideband. The solid red line shows the fit result with the threshold function in each panel.

5.4 Fit Procedure

A simultaneous Unbinned Extended Maximum-likelihood fit is performed on the $M(D^0\overline{D}^{*0})$ distributions. The distributions with different signal-to-background ratios due to the D^{*0} modes are separated. Furthermore, we treat the modes separately to measure the relative branching fraction between $B^+ \to X(3872)K^+$ and $B^0 \to X(3872)K^0$. In total, we simultaneously fit four distributions: $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$, $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$, $B^0 \to K^0_S(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$, and $B^0 \to K^0_S(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$.

We treat three components in the fit: signal, broken-signal, and generic background. The PDFs and constraints used for each component are summarized below.

- Signal component; Parameter conditions of the assumed lineshape are described in the subsequent subsections. The lineshape is multiplied by the mass-dependent signal efficiency and then convoluted the mass resolution at the peak to create a PDF for the signal. The ratio of signal yields between $\overline{D}^{*0} \to \overline{D}^0\gamma$ and $\overline{D}^{*0} \to \overline{D}^0\pi^0$ is fixed by the total efficiency ratio for the assumed lineshape. The typical yield ratio of $\overline{D}^{*0} \to \overline{D}^0\pi^0$ to $\overline{D}^{*0} \to \overline{D}^0\gamma$ is about 1.1 regardless of B modes. To measure the relative branching fraction between $B^+ \to X(3872)K^+$ and $B^0 \to X(3872)K^0$, the signal yield for each B mode is floated.

- Broken-signal component; For both decay modes, parameters of the assumed lineshape are always common with the signal PDF. For $\overline{D}^{*0} \to \overline{D}^0\gamma$, we use the lineshape-dependent histogram PDF estimated using the zero-width signal MC samples. The yield of the broken-signal relative to the signal is fixed to that obtained with the default signal MC sample. For $\overline{D}^{*0} \to \overline{D}^0\pi^0$, we use a PDF in which the assumed lineshape is multiplied by the efficiency and then discretely convoluted with the mass-dependent broken-signal resolution PDF. The yield of the broken-signal relative to the signal is fixed from the total efficiency ratio between the signal and the broken-signal.

- Generic background component; The shape is the threshold function using a square root. Note that it has no parameters for the shape. The parameter of the background yield for each decay mode is floated.

The subsequent subsection describes the detailed parameter conditions and a fit validation for each lineshape.

5.4.1 Breit-Wigner Lineshape

Parameter Conditions

The relativistic Breit-Wigner lineshape has two parameters: m and Γ_0. (See Sec. 2.4.) Both of them are floated in the fit. Figure 5.19 shows the comparison of the relativistic Breit-Wigner lineshapes with various m's and Γ_0's. For larger Γ_0, the peak is easy to move to the threshold. It implies that uncertainty is not always symmetric. We need to evaluate it with a likelihood profile, and `Minos` in RooFit [69] is used for this purpose.

Fit Validation

We check if a fit result reproduces the input values correctly with pseudo experiments. Pseudo experiments are performed using the following procedure.

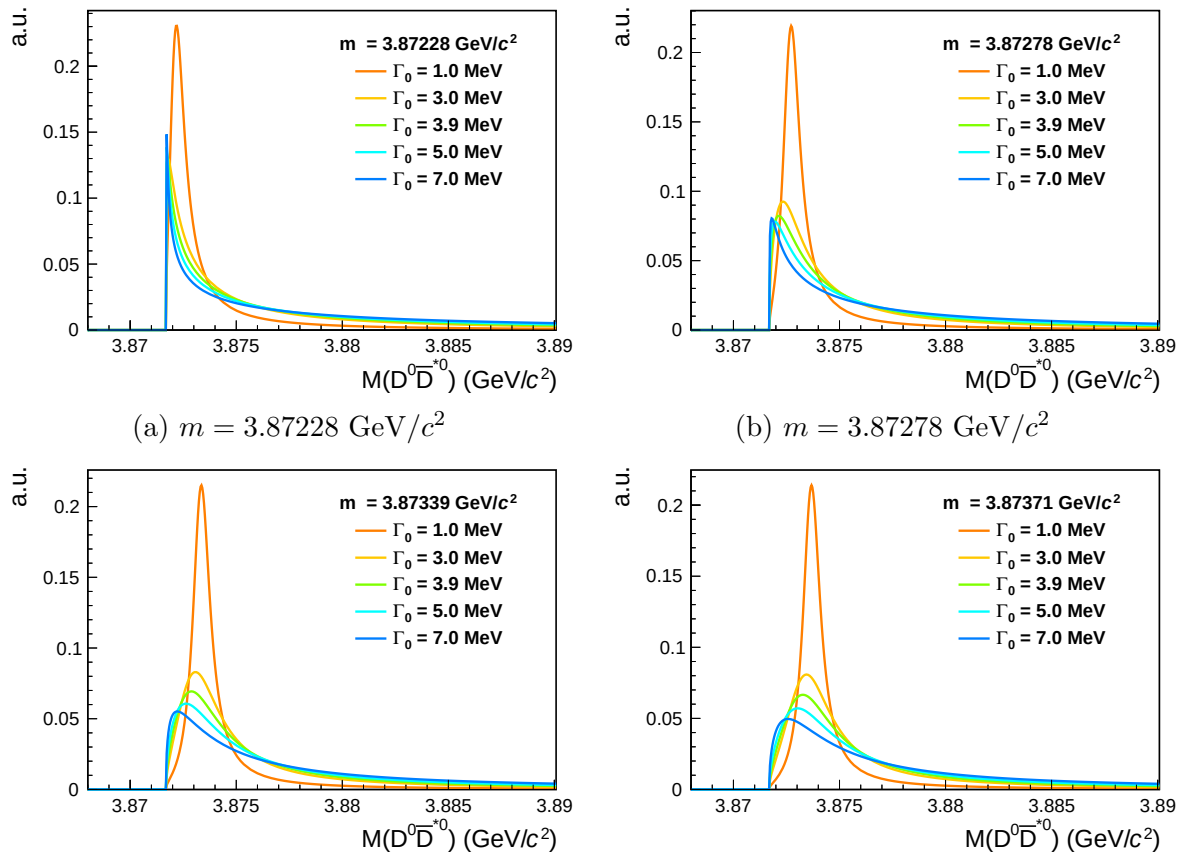
1. Generate the $D^0\overline{D}^{*0}$ distributions according to the PDFs with arbitrary m and Γ_0. The number of generated events is determined by a random number according to the Poisson distribution centered at the predicted value. For the signal and broken-signal components, the predicted values are derived by the total efficiency and an arbitrary branching fraction $\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0})$. For the generic background, we use the values obtained from the background MC sample and the data sideband.

2. Fit them with the PDFs floated m, Γ_0, the signal yield, and the generic background yield.

From eight hundred pseudo experiments, pull distributions of m, Γ_0 and signal yields are made to check the fit bias. The pull is defined as follows for asymmetric uncertainties [70]:

$$\begin{cases} \dfrac{(\text{true value}) - (\text{fit value})}{(\text{positive uncertainty})} & (\text{For (true value)} - \text{(fit value)} \geq 0), \\ \dfrac{(\text{fit value}) - (\text{true value})}{(\text{negative uncertainty})} & (\text{Otherwise}). \end{cases} \tag{5.11}$$

The pull distributions are evaluated by fitting with a Gaussian.



(a) $m = 3.87228$ GeV/c^2 (b) $m = 3.87278$ GeV/c^2

(c) $m = 3.87339$ GeV/c^2 (d) $m = 3.87371$ GeV/c^2

Figure 5.19: Comparison of the relativistic Breit-Wigner lineshapes with the various masses and widths.

The following five input values for m are to be verified under $\mathcal{B}(B \to KX(3872)) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}) = 0.80 \times 10^{-4}$, which is the center value of the previous measurement at Belle [2]:

$$m = \begin{cases} 3.87228 \text{ GeV}/c^2 & (-1.0\sigma \text{ from center value at Belle 2010}), \\ 3.87278 \text{ GeV}/c^2 & (\text{Center value at Belle 2010}), \\ 3.87339 \text{ GeV}/c^2 & (+1.0\sigma \text{ from center value at Belle 2010}), \\ 3.87371 \text{ GeV}/c^2 & (+1.5\sigma \text{ from center value at Belle 2010}), \\ 3.87510 \text{ GeV}/c^2 & (\text{Center value at } BABAR \text{ 2008}). \end{cases}$$

where we take $3.87278^{+0.0006}_{-0.0007}$ GeV/c^2 as the m measurements at Belle 2010, which is updated according to the updates of masses of D^0 and D^{*0}. For each input m, we verify five widths in the 1σ range of the previous measurements at Belle and *BABAR*:

$$\Gamma_0 = \{\ 1.0 \text{ MeV}, 3.0 \text{ MeV}, 3.9 \text{ MeV}, 5.0 \text{ MeV}, 7.0 \text{ MeV}\ \}.$$

In addition, the following seven branching ratios are assumed to verify the signal yields:

$$\mathcal{B}(B \to KX(3872)) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}) = \begin{cases} 0.58 \times 10^{-4} & (-1\sigma \text{ from center value at Belle 2010}), \\ 0.69 \times 10^{-4} & (-0.5\sigma \text{ from center value at Belle 2010}), \\ 0.80 \times 10^{-4} & (\text{Center value at Belle 2010}), \\ 0.91 \times 10^{-4} & (+0.5\sigma \text{ from center value at Belle 2010}), \\ 1.02 \times 10^{-4} & (+1\sigma \text{ from center value at Belle 2010}), \\ 1.67 \times 10^{-4} & (\text{Center value at } \textit{BABAR} \text{ 2008}), \\ 2.26 \times 10^{-4} & (+1\sigma \text{ from center value at } \textit{BABAR} \text{ 2008}). \end{cases}$$

Figures 5.20–5.21 summarize the means and standard deviations of pull for m and Γ_0. For the width, pull distributions are almost consistent with a Gaussian with a mean of 0.3 and a standard deviation of one regardless of the input mass. It indicates that the fit bias is less than statistical uncertainty, so it is taken in the systematic uncertainty. For the mass, the bias is smaller than the statistical uncertainty in most cases, and thus the bias is basically accounted for the systematic uncertainty rather than corrected the final results. However, the bias becomes large, especially when the input mass is around the threshold or the input width is large. In the former case, the peak is so sharp that the shape is indistinguishable due to the mass resolution. In the latter case, the PDF summing the signal and broken-signal components becomes almost the same even when m is changed, as shown in Figs. 5.22–5.23. A bias correction is applied only if the mass or width obtained for data is expected to have a large mass bias compared to its statistical uncertainty or the previous studies.

Means and standard deviations of pull for yield are summarized in Figs. 5.24–5.25. They are almost consistent with a Gaussian with a mean of 0.1 and a standard deviation of one. It means the bias is small enough compared to the statistical uncertainty, and thus it is taken in the systematic uncertainty.

Comparison of Statistical Uncertainty with Previous Studies

To check improvement, we compare the expected precision with the most precise previous measurement [2]. The expected statistical uncertainty of our measurement is evaluated by the medians of fit outputs, upper uncertainties, and lower uncertainties obtained from pseudo experiments. Here, the reason for using a median instead of a mean is to reduce the impact caused by outliers or asymmetry of the output distribution. The results are summarized in Table 5.4. Compared with the previous study, all statistical uncertainties are improved by about 5–19%. It is almost consistent with the sensitivity improvement expected when peak yield is improved by a factor of two (Sec. 4.3). Note that the statistical precision of this measurement depends on the lineshape parameters. The dependence on mass and width is shown in Fig. 5.26. The precision degrades for smaller mass or larger width.

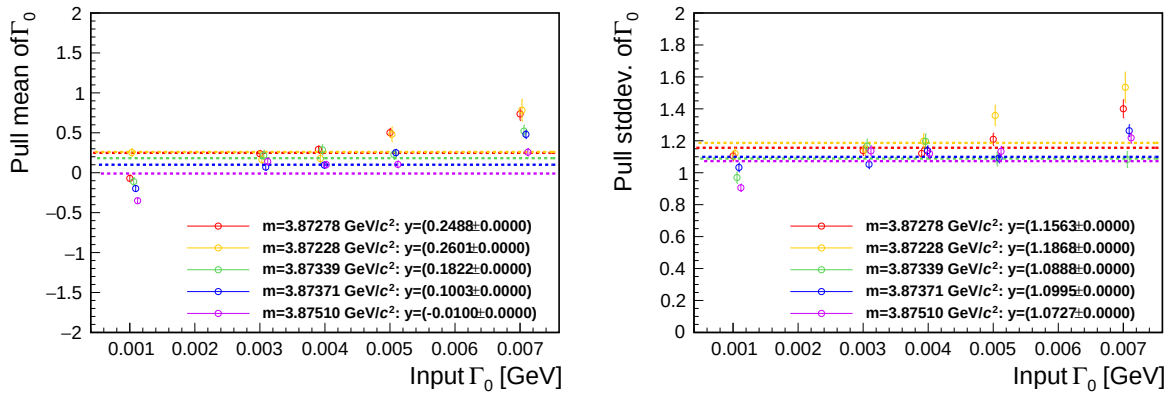


Figure 5.20: Means of pull (left) and standard deviations of pull (right) as a function of input Γ_0.

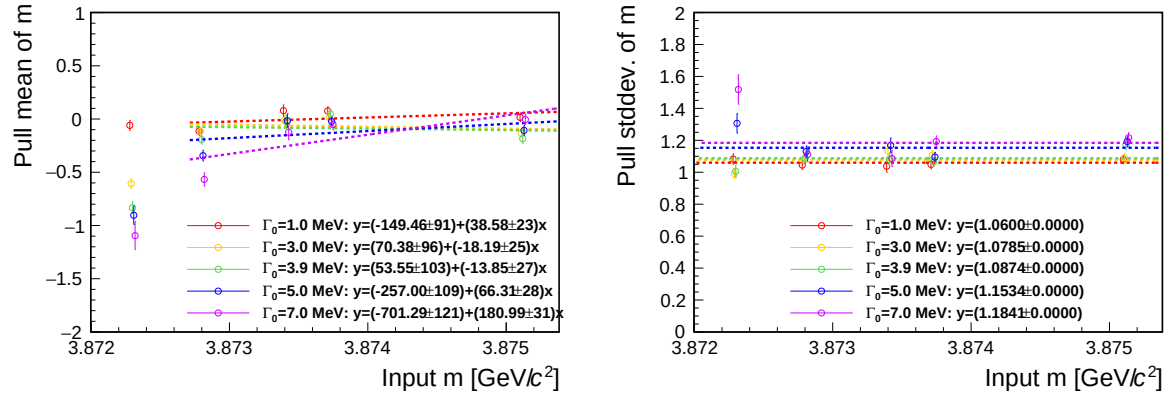
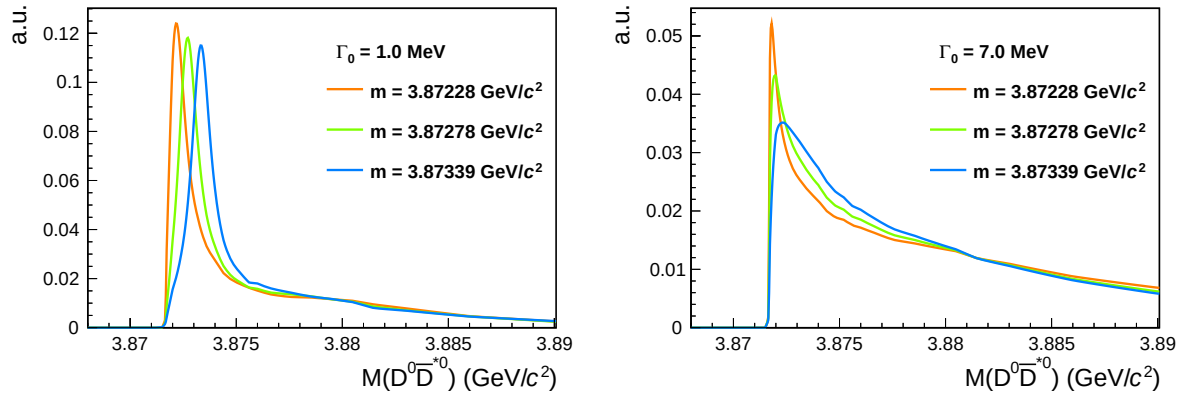


Figure 5.21: Means of pull (left) and standard deviations of pull (right) as a function of input m.

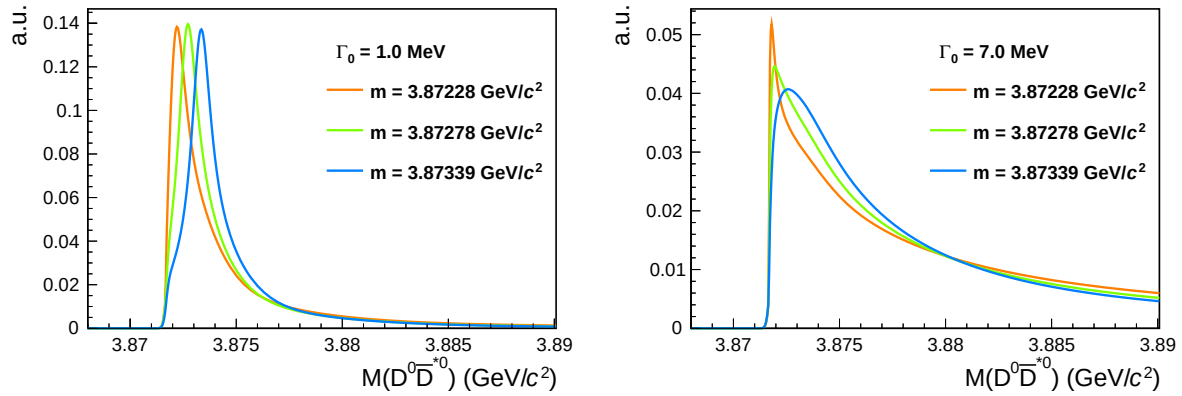
Table 5.4: Comparison of medians of the output m, Γ_0, and yield values with results in the previous study.

	Belle 2010	This study		Improvement
		Assumption	Output mean	
m (MeV/c^2)	3872.8 (+0.6/ − 0.4)	3872.78	3872.96 (+0.51/ − 0.44)	5%
Γ_0 (MeV)	3.9 (+2.8/ − 1.4)	3.9	3.7 (+2.3/ − 1.3)	10%
Yield	50.1 (+14.8/ − 11.1)	54.4	53.8 (+12.7/ − 9.9)	19%



(a) $\Gamma_0 = 1.0$ MeV (b) $\Gamma_0 = 7.0$ MeV

Figure 5.22: Comparison of the sum of the signal and broken-signal PDFs with various input masses for $D^{*0} \to D^0\gamma$.



(a) $\Gamma_0 = 1.0$ MeV (b) $\Gamma_0 = 7.0$ MeV

Figure 5.23: Comparison of the sum of the signal and broken-signal PDFs with various input masses for $D^{*0} \to D^0\pi^0$.

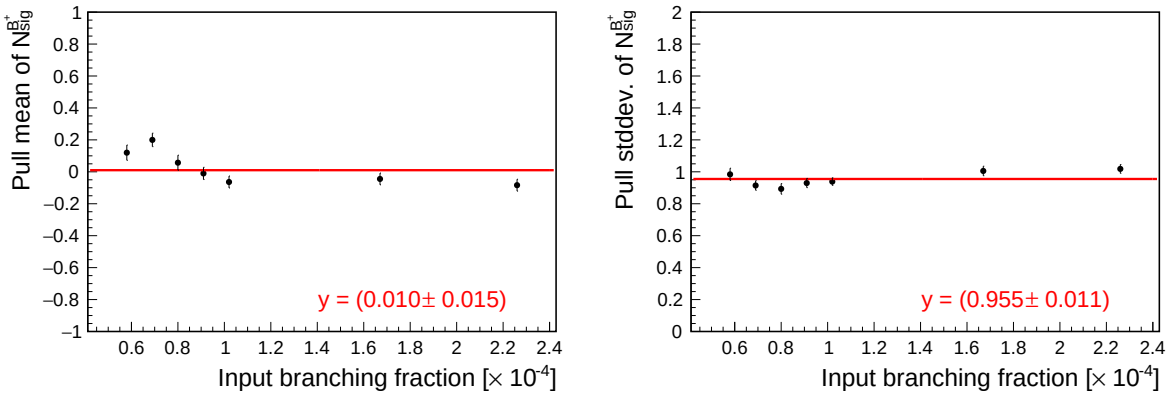


Figure 5.24: Means of pull (left) and standard deviations of pull (right) for input B^+ yield as a function of input branching fraction.

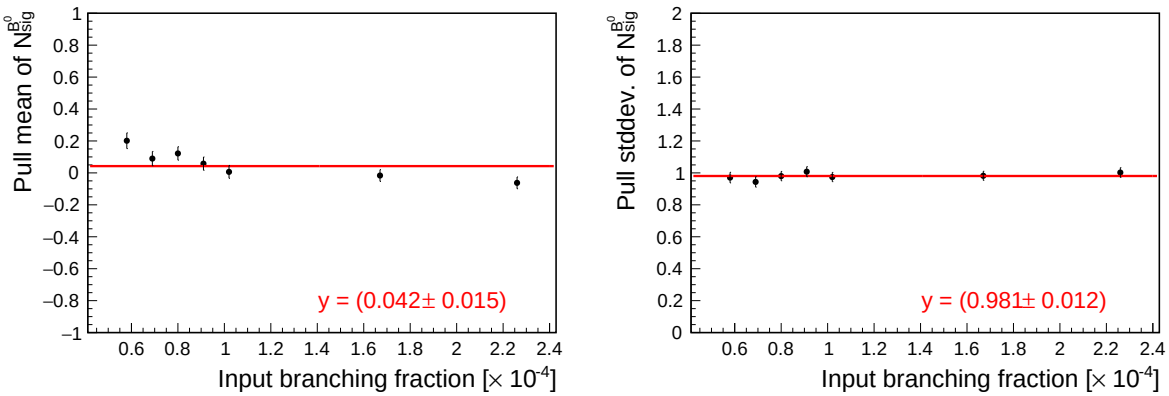
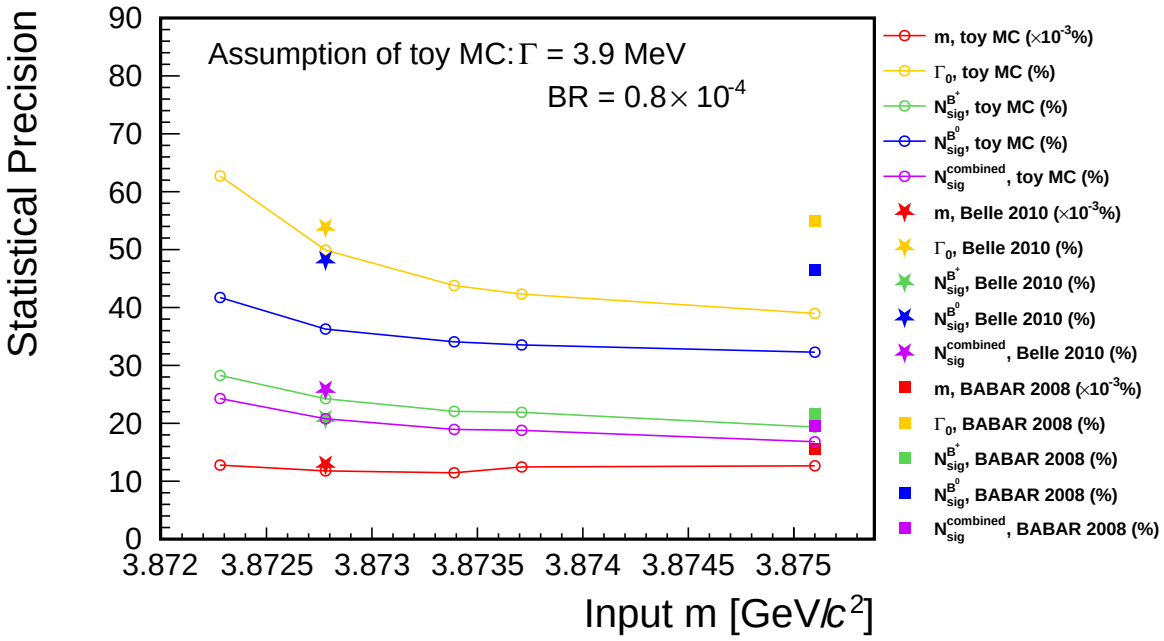
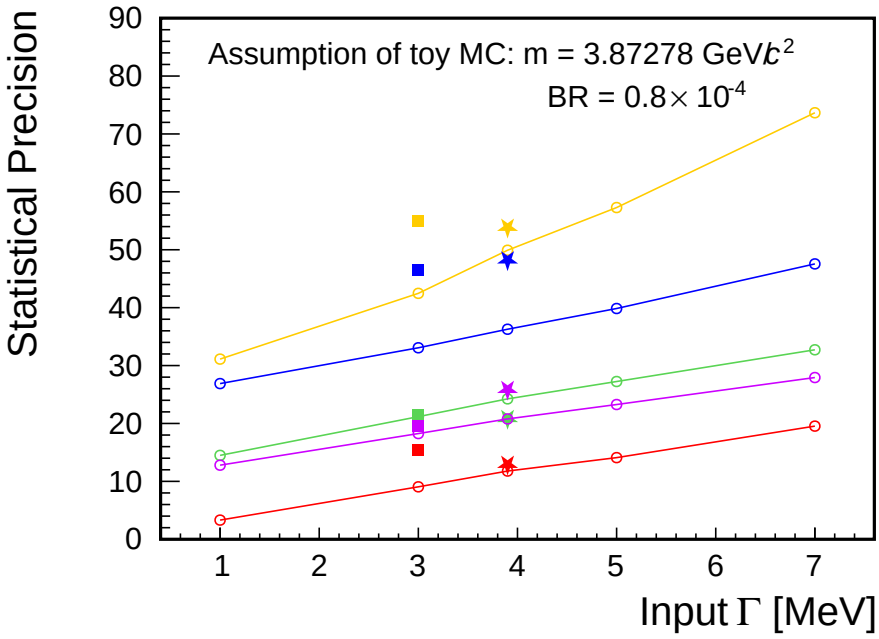


Figure 5.25: Means of pull (left) and standard deviations of pull (right) for input B^0 yield as a function of input branching fraction.



(a) Dependence of the input mass



(b) Dependence of the input width

Figure 5.26: Statistical precision, which is obtained from a center value over a statistical uncertainty, for m (red), Γ_0 (yellow), the signal yield from B^+ decays (green), the signal yield from B^0 decays (blue), and the combined signal yield (violet). The opened circles with lines represent the expectation of our measurement. The closed stars and squares represent the Belle and *BABAR* measurements [2, 3], respectively.

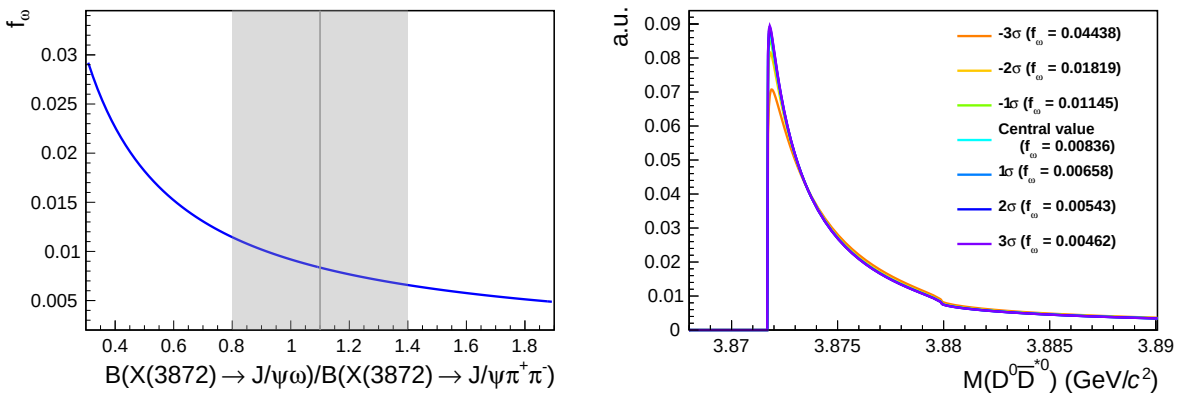


Figure 5.27: (Left) f_ω as a function of the branching ratio between $X(3872) \to J/\psi\omega$ and $J/\psi\pi^+\pi^-$. The gray line is the central value of the world-average branching ratio, and the gray region is $\pm 1\sigma$ of the uncertainty. (Right) Comparison among the lineshape PDFs with seven f_ω values: the value under the central value of the world-average branching ratio, and those under $\pm 1\sigma$, $\pm 2\sigma$, and $\pm 3\sigma$ from the center value. For simplicity, the other parameters are fixed to Eq. (2.20).

5.4.2 Flatté Lineshape

Parameter Conditions

The Flatté model has five free parameters: E_f, g, f_ρ, f_ω, and Γ_0 (Sec. 2.4). The fit floating all parameters is unstable because the $X(3872) \to D^0\overline{D}^{*0}$ sample does not have enough statistics in this analysis. Therefore, we apply some constraints to the fit.

According to the LHCb result [4], dg/dE_f is fixed as follows because it was measured with a precision of about 1%:

$$\frac{dg}{dE_f} = -15.11 \text{ GeV}^{-1}. \tag{5.12}$$

The uncertainty due to the dg/dE_f measurement is assigned in the systematic uncertainty.

The effective coupling constant to the $J/\psi\omega$ channel f_ω is fixed so that the branching fraction of the $J/\psi\pi^+\pi^-$ mode and that of the $J/\psi\omega$ mode are equal, consistent with experimental results to date [33–35]. It is predicted by the ratio of areas of the Flatté lineshapes as follows:

$$\frac{\mathcal{B}(X(3872) \to J/\psi\omega)}{\mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)} = \frac{\int_{m_{D^0}+m_{D^{*0}}}^{m_B-m_K} f_{\text{Flatte}}(M(J/\psi\omega))dM(J/\psi\omega)}{\int_{m_{D^0}+m_{D^{*0}}}^{m_B-m_K} f_{\text{Flatte}}(M(J/\psi\pi^+\pi^-))dM(J/\psi\pi^+\pi^-)}. \tag{5.13}$$

To save computational resources, the integration range is set from 3.77GeV/c^2 to 3.97 GeV/c^2, where the lineshape converges sufficiently. Figure 5.27 (left) is f_ω as a function of the ratio. When the branching ratio changes in the $\pm 1\sigma \equiv \pm 0.4$ range of the world-average uncertainty, it shows that f_ω changes by about $+37\%/-21\%$. However, a comparison of the lineshape PDFs under the f_ω changes, shown in Fig. 5.27 (right), indicates that the effect on the lineshape is negligible.

Figure 5.28 (left) shows a comparison of the lineshapes scaled by all parameters by a constant, which the parameter setting causes the scaling behavior for the $J/\psi\pi^+\pi^-$ mode. It shows that the width increases logarithmically rather than linearly when g increases

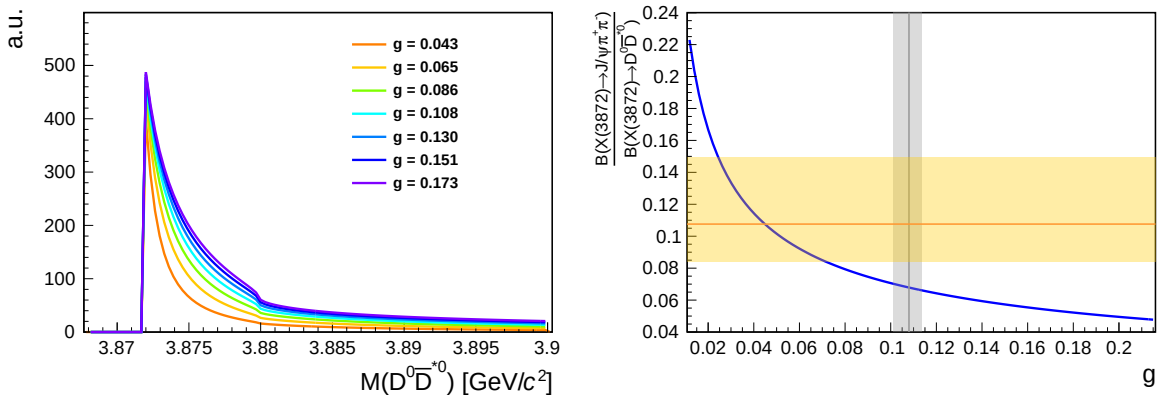


Figure 5.28: (Left) Comparison among the lineshapes (**NOT** PDFs) with seven g values: 0.043, 0.065, 0.086, 0.108, 0.130, 0.151 and 0.173. (Right) The branching ratio between $X(3872) \to D^0\overline{D}^{*0}$ and $J/\psi\pi^+\pi^-$ as a function of g. The ratios between the other parameters are fixed to Eq. (2.20). The gray line is the central value under the LHCb assumption (Eq. (2.20)), and the gray region is $\pm 1\sigma$ of the uncertainty. The orange line is the central value of the Belle previous measurement with the relativistic Breit-Wigner lineshape [2], and the orange region is $\pm 1\sigma$ of the uncertainty.

linearly. It is expected to cause asymmetric errors of g in the fit and makes it difficult to determine the upper statistical uncertainty. Therefore as a countermeasure for the former, we use `Minos` to obtain the asymmetric uncertainty from the likelihood profile. As a countermeasure for the latter, we introduce the constraint of the branching ratio between the $J/\psi\pi^+\pi^-$ channel and the $D^0\overline{D}^{*0}$ channel. The ratio is defined as

$$\frac{\mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)}{\mathcal{B}(X(3872) \to D^0\overline{D}^{*0})} = \frac{\int_{m_{D^0}+m_{D^{*0}}}^{m_B-m_K} f_{\text{Flatte}}(M(J/\psi\pi^+\pi^-))dM(J/\psi\pi^+\pi^-)}{\int_{m_{D^0}+m_{D^{*0}}}^{m_B-m_K} f_{\text{Flatte}}(M(D^0\overline{D}^{*0}))dM(D^0\overline{D}^{*0})}. \quad (5.14)$$

The branching fraction as a function of the coupling constant for f_ρ and Γ_0 scaled simultaneously is shown in Fig. 5.28 (right). The larger the coupling constant is, the smaller this ratio is. However, the $\mathcal{B}(X(3872) \to D^0\overline{D}^{*0})$ depends on its own measurement. For this analysis, we transform the formula and apply the constraint on $\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)$,

$$\begin{aligned}
&\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-) \\
&= \frac{\int_{m_{D^0}+m_{D^{*0}}}^{m_B-m_K} f_{\text{Flatte}}(M(J/\psi\pi^+\pi^-))dM(J/\psi\pi^+\pi^-)}{\int_{m_{D^0}+m_{D^{*0}}}^{m_B-m_K} f_{\text{Flatte}}(M(D^0\overline{D}^{*0}))dM(D^0\overline{D}^{*0})} \\
&\quad \times \mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}).
\end{aligned} \quad (5.15)$$

Here, the branching fraction of the $D^0\overline{D}^{*0}$ mode is calculated as follows:

$$\begin{aligned}
\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}) &= \frac{N_{\text{sig}}}{2N_{B\overline{B}}\mathcal{B}(\Upsilon(4S) \to B^+B^-)\epsilon_{\text{tot}}}, \\
\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}) &= \frac{N_{\text{sig}}}{2N_{B\overline{B}}\mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0)\epsilon_{\text{tot}}},
\end{aligned} \quad (5.16)$$

where N_{sig} is the signal yield on the $D^0\overline{D}^{*0}$ invariant mass distributions, which is determined by this fit. $N_{B\overline{B}}$ is the number of the $B\overline{B}$ pairs in the data, and $\mathcal{B}(\Upsilon(4S) \to f)$

Table 5.5: Input values for the calculation of $\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0})$.

Parameter	Value	Reference
$N_{B\overline{B}}$	$(772 \pm 11) \times 10^6$	
$\mathcal{B}(\Upsilon(4S) \to B^+B^-)$	$(51.4 \pm 0.6)\%$	[18]
$\mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0)$	$(48.6 \pm 0.6)\%$	[18]

is the branching fraction from $\Upsilon(4S)$ to a mode f. The factor of two in the denominator is used to count the number of B mesons in the $B\overline{B}$ pairs. Their input values are summarized in Table 5.5. $\epsilon_{\rm tot}$ is the total efficiency, and it is the sum of $\epsilon_{\rm tot}$ of the two D^{*0} modes defined in Sec. 5.1.3.

For the value of $\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)$, we evaluate the average of the independent measurements from *BABAR* [71] and Belle [31] and also LHCb [72] for that the B^+ mode (Table 5.6, Fig. 5.29). They are calculated by a weighted average of the measurements using their uncertainties. The measurements of Belle and *BABAR* contain systematic uncertainties of the same sources: the MC model and the secondary branching fractions. Therefore, we conservatively consider full correlation for the uncertainties of the sources and combine them. The values are obtained as

$$\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-) = (8.10 \times 0.30) \times 10^{-6}, \tag{5.17}$$

$$\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-) = (4.1 \times 1.1) \times 10^{-6}. \tag{5.18}$$

To include the uncertainty of $\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)$, we perform the fit with a soft constraint by multiplying the likelihood ($\mathcal{L}$) by constraint PDFs ($\mathcal{C}$) as a function of the branching fraction,

$$\begin{aligned}\mathcal{L}_c =& \mathcal{L} \times \mathcal{C}(\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)) \\ & \times \mathcal{C}(\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)).\end{aligned} \tag{5.19}$$

A normalized Gaussian is used as the constraint PDF,

$$\mathcal{C}(\mathcal{B}) = \frac{1}{\sqrt{2\pi}\sigma\mathcal{B}_{\rm Avg.}} \exp\left(- \frac{(\mathcal{B}_{\rm Avg.} - \mathcal{B}_{\rm Mes.})^2}{2(\sigma\mathcal{B}_{\rm Avg.})^2} \right), \tag{5.20}$$

where $\mathcal{B}_{\rm Mes.}$ is the branching fraction calculated by $N_{\rm sig}$ (Eq. (5.15)), $\mathcal{B}_{\rm Avg.}$ is the average branching fraction, and $\sigma\mathcal{B}_{\rm Avg.}$ is the $\mathcal{B}_{\rm Avg.}$ uncertainty. This method is also known as the Gaussian constraint [73]. Strictly speaking, the correlation between two branching ratios should be included. They are assumed to be uncorrelated since the branching fraction from B^+ depends exclusively on the LHCb measurement and that from B^0 depends on the other measurements.

There are insufficient events in our $X(3872) \to D^0\overline{D}^{*0}$ sample to simultaneously determine the three remaining parameters, i.e. g, f_ρ, and Γ_0. Therefore, we focus on the parameter regions where scaling behavior was observed at LHCb [4]. We search for the best lineshape fitted to the $M(D^0\overline{D}^{*0})$ distribution when the following ratios of parameters are fixed to the values measured at LHCb: f_ρ/E_f and Γ_0/E_f are fixed based on the measurements $f_\rho = 1.8 \times 10^{-3}$ and $\Gamma_0 = 1.4$ MeV under the assumption $E_f = -7.2$ MeV. Therefore, the free parameter for the signal PDF is only g.

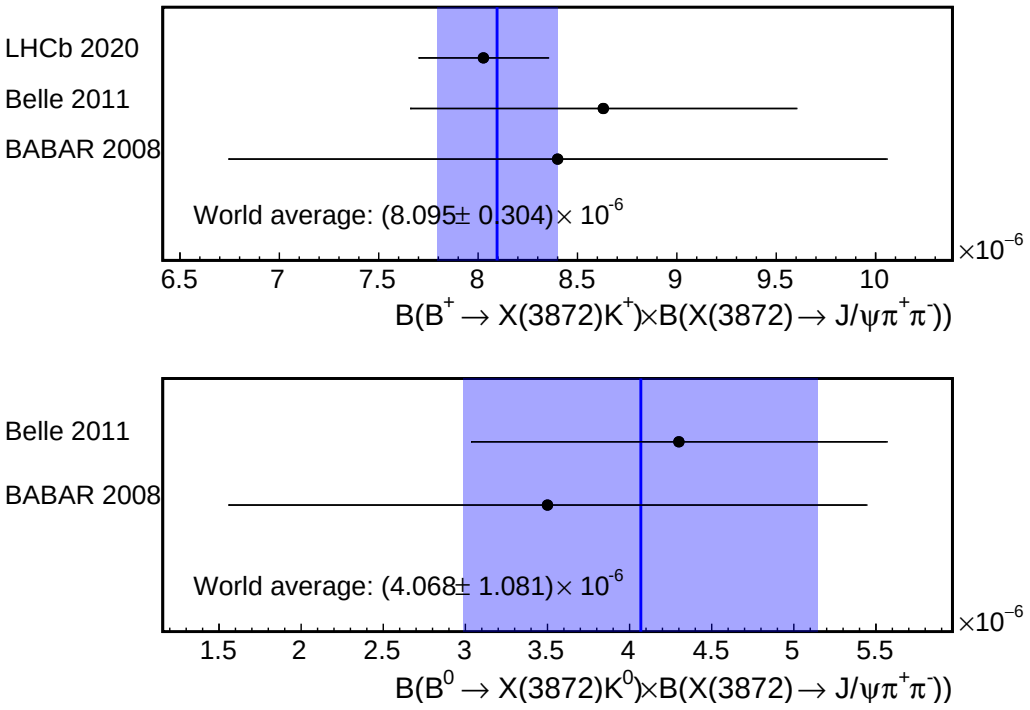


Figure 5.29: (Top) Comparison of $\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)$ at *BABAR* [71], Belle [31], and LHCb [72]. (Bottom) Comparison of $\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)$ between *BABAR* [71] and Belle [31]. In each panel, the blue line is the average value, and the blue region is the $\pm 1\sigma$ of the uncertainty.

Fit Response

The validity of the fit is confirmed using pseudo experiments. The procedure of the pseudo experiments is as follows.

1. Generate the $D^0\overline{D}^{*0}$ distributions according to the PDFs with an arbitrary g. Here, f_ρ and Γ_0 are also scaled so that the ratio to g is constant. The number of generated events is determined by a random number according to the Poisson distribution centered at the predicted value. For components of the signal and the broken-signal, the predicted values are derived by the total efficiency and an arbitrary branching fraction. For the generic background, we use the values obtained from the background MC sample and the data sideband.

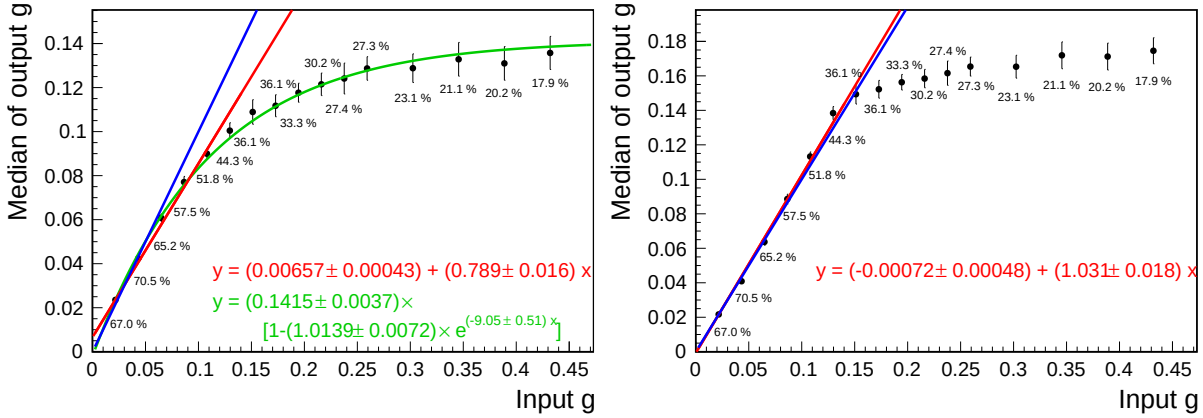
2. Fit them with the PDFs floated g of the Flatté parameters and the number of the signal and that of the generic background. Here the constraint value of $\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)$ is fluctuated by its uncertainty [73].

To verify the fit output, we check only the median of the output value distribution as a function of the input value because the pull distribution could not be evaluated well due to the large asymmetry.

The g response is shown in Fig. 5.30 (a). If g is smaller than 0.100, i.e., a narrow lineshape, g obtained by the fit can reproduce the input value, because it is well-separated from the generic background event even with the statistically limited data at Belle. In the case of large g, fits fail, especially in determining an upper statistical uncertainty.

Table 5.6: Summary of the $\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)$ values used in the calculation of the weighted average, Eqs. (5.17)-(5.18). For the LHCb result [72], $\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)$ is obtained by a product of $\frac{\mathcal{B}(B^+\to X(3872)K^+)\times\mathcal{B}(X(3872)\to J/\psi\pi^+\pi^-)}{\mathcal{B}(B^+\to\psi(2S)K^+)\times\mathcal{B}(\psi(2S)\to J/\psi\pi^+\pi^-)}$, $\mathcal{B}(B^+ \to \psi(2S)K^+)$, and $\mathcal{B}(\psi(2S) \to J/\psi\pi^+\pi^-)$.

Experiment	Branching fraction	Value [$\times 10^{-6}$]	Ref.
Belle	$\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)$	$(8.63 \pm 0.82 \pm 0.52) \times 10^{-6}$	[31]
	$\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)$	$(4.3 \pm 1.2 \pm 0.4) \times 10^{-6}$	
BABAR	$\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)$	$(8.4 \pm 1.5 \pm 0.7) \times 10^{-6}$	[71]
	$\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)$	$(3.5 \pm 1.9 \pm 0.4) \times 10^{-6}$	
LHCb	$\frac{\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)}{\mathcal{B}(B^+ \to \psi(2S)K^+) \times \mathcal{B}(\psi(2S) \to J/\psi\pi^+\pi^-)}$	$(3.69 \pm 0.07 \pm 0.06) \times 10^{-2}$	[72]
	$\mathcal{B}(B^+ \to \psi(2S)K^+)$	$(6.24 \pm 0.20) \times 10^{-4}$	[18]
	$\mathcal{B}(\psi(2S) \to J/\psi\pi^+\pi^-)$	$(34.68 \pm 0.30) \times 10^{-2}$	[18]



(a) Before the bias correction (b) After the bias correction

Figure 5.30: The median of the output g values with respect to the input g values based on 4000 pseudo experiments. The values near the points are the success rates of the fit in the pseudo experiment. The blue line is an ideal relation, $g_{\text{output}} = g_{\text{input}}$. The red line results from the fit with a first-order polynomial function in the input g range of 0.000–0.150. The green line shows the threshold function, Eq. (5.21).

The main reason is that the lineshape converges to a fixed form for large g (given the assumed ratios for the other parameters); See Eq. (2.18). It causes the underestimation of g. The bias is corrected in some cases by the relation between the median of the output parameters and the input value; It can be reproduced by the following threshold function (green lines in Fig. 5.30 (a)),

$$g_{\text{output}} = p_0(1 - p_1 \exp(p_2 \cdot g_{\text{input}})), \tag{5.21}$$

where g_{output} and g_{input} denote the output and input g values, respectively. The corrected result is shown in Fig. 5.30 (b). It confirms that better linearity is obtained for $g < 0.14$. If we observe a larger g than 0.14, we cannot determine the center value; therefore the lower limit is determined using the likelihood profile. If the fit fails for the data, the statistical uncertainty is re-evaluated with pseudo experiments.

Chapter 6

Fit to Data and Systematic Uncertainty

This chapter describes the fit result for the real data for each lineshape model.

6.1 Relativistic Breit-Wigner Lineshape

Figure 6.1 presents the $M(D^0\overline{D}^{*0})$ distributions obtained from the data. Here, unbinned maximum likelihood fits are performed simultaneously to the distributions for the $\overline{D}^{*0}$ decay modes, $\overline{D}^{*0} \to \overline{D}\pi^0$ and $\overline{D}\gamma$, and for the B^+ and B^0 samples, with the common fit parameters m and Γ_0. Table 6.1 summarizes the parameters obtained from the fit. The significance is determined from the log-likelihood ratio $-2\ln(\mathcal{L}_0/\mathcal{L})$ accounting for the difference in the number of degrees of freedom, where $\mathcal{L}_0$ and $\mathcal{L}$ are the fit likelihood without and with the peak component, i.e., the yield is dropped for the significance of each B mode, and the parameters m and Γ_0, and the yields of both B modes, are dropped for the combined significance. Here the likelihood is smeared to take account of the systematic uncertainties on the signal yields as described below. The significance is found to be 5.9σ for $B^+ \to X(3872)K^+$, and 5.2σ for $B^0 \to X(3872)K^0$. To investigate contributions from peaking background, we check the $M(D^0\overline{D}^{*0})$ distribution in the $(M_{\rm bc}, \Delta E)$ sideband region, as shown in Fig. 6.2; here the sideband region is defined as 12 MeV/c^2 $< |M_{\rm bc} - m_B| <$ 20 MeV/c^2 or 30 MeV $< |\Delta E| <$ 50 MeV. The absence of peaks confirms that any contribution from peaking background is small

The lineshape parameters are determined to be

$$m = 3873.71^{+0.56}_{-0.50}(\text{stat}) \pm 0.13(\text{syst})\ \text{MeV}/c^2,$$

$$\Gamma_0 = 5.2^{+2.2}_{-1.5}(\text{stat}) \pm 0.4(\text{syst})\ \text{MeV}.$$

Table 6.1: Results using the relativistic Breit-Wigner lineshape: the fitted mass, width and signal yield, the total signal efficiency, and the significance.

Mode	m (MeV/c^2)	Γ_0 (MeV)	$N_{\rm sig}$	$\epsilon_{\rm tot}$ ($\times 10^{-4}$)	significance
Combined	$3873.71^{+0.56}_{-0.50}$	$5.2^{+2.2}_{-1.5}$	$70.5^{+13.6}_{-11.5}$	8.70	7.5σ
$X(3872)K^+$	—	—	$53.2^{+11.6}_{-9.8}$	6.92	5.9σ
$X(3872)K^0$	—	—	$17.3^{+4.7}_{-4.1}$	1.78	5.2σ

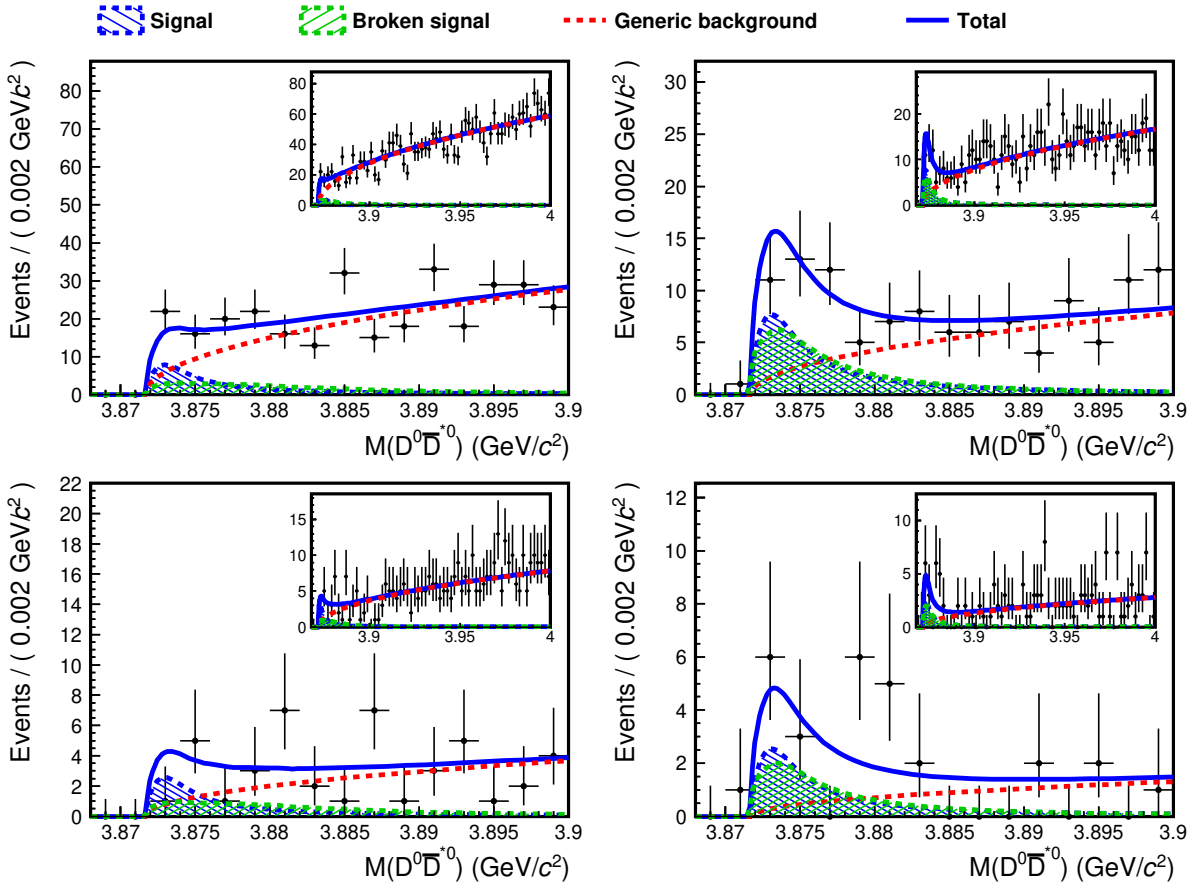


Figure 6.1: The $M(D^0\overline{D}^{*0})$ distributions with the fit result with the relativistic Breit-Wigner lineshape for $B^+ \to X(3872)K^+$ (top) and $B^0 \to X(3872)K^0$ (bottom). The left and right rows are for $\overline{D}^{*0} \to \overline{D}^0\gamma$ and $\overline{D}^{*0} \to \overline{D}^0\pi^0$, respectively. The points with error bars represent data. The solid blue line shows the total fit result. The dashed blue and green lines show the signal and broken-signal contributions, respectively. The dashed red line shows the generic background contribution.

We consider the following nine sources of systematic uncertainty on the mass, and the width, as listed in Table 6.2:

(i) The systematic uncertainty due to the assumed shape of the generic background is estimated by performing a fit after changing the PDF from the threshold function with a square root to an inverted ARGUS function [74]

$$f_{\rm ARGUS}(M) = p_0 \cdot M\sqrt{1-\left(\frac{M}{p_1}\right)^2} \cdot \exp\left[p_2\left(1-\left(\frac{M}{p_1}\right)^2\right)\right], \qquad (6.1)$$

where p_0 and p_1 are the normalization factor and the threshold, respectively. In the fit, p_1 is fixed to the $D^0\overline{D}^{*0}$ threshold, and p_2 is floated. The symmetric both-side uncertainty based on the change from the original result is taken as a conservative estimation of the systematic uncertainty.

(ii) The mass resolution is validated by comparing the data and MC ΔE resolution in the $B^+ \to \overline{D}^{*0}\pi^+\pi^-\pi^+$ control sample, which has a similar decay topology to $B \to (X(3872) \to D^0\overline{D}^{*0})K$. This resolution is consistent in data and MC, so no

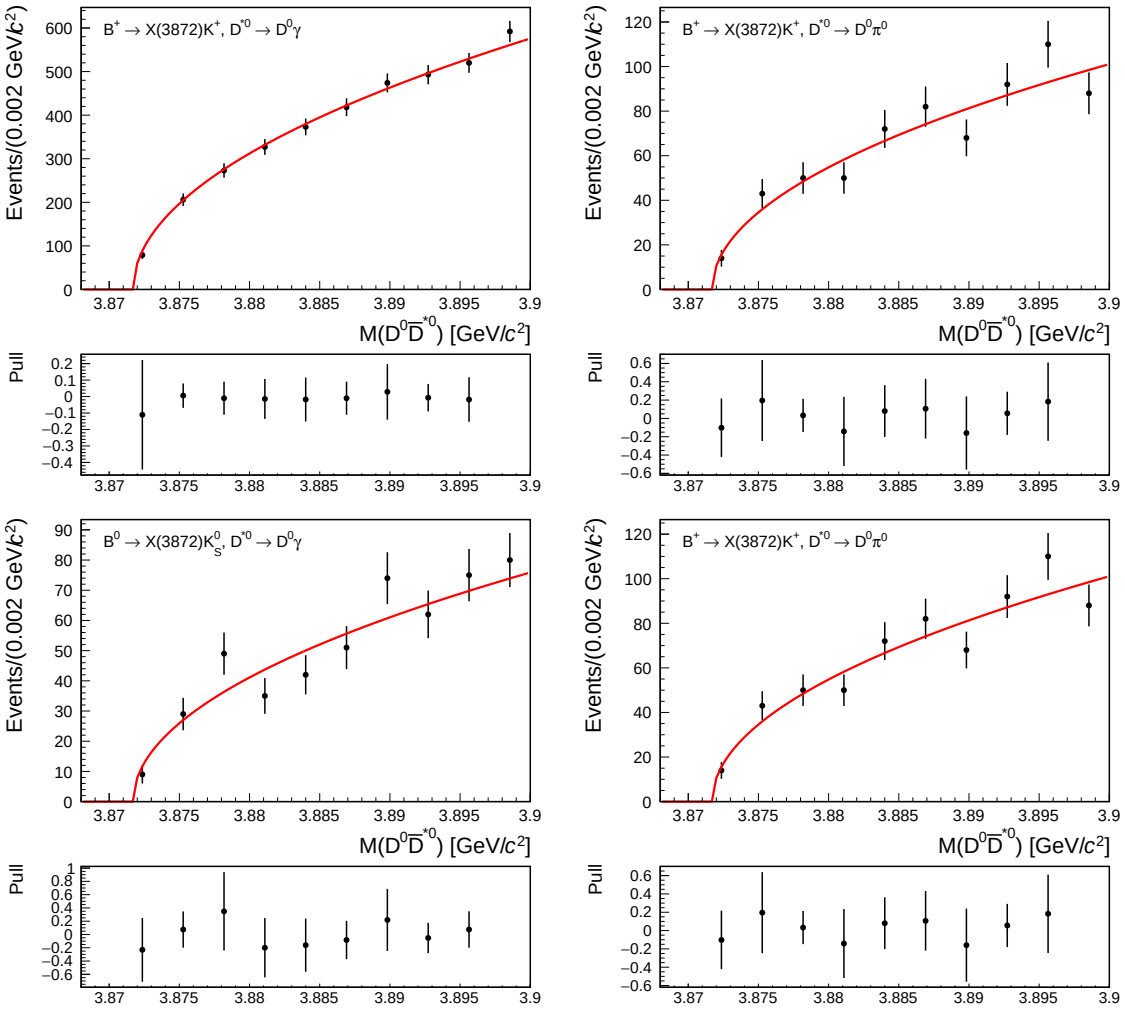


Figure 6.2: The $M(D^0\bar{D}^{*0})$ distributions in the $M_{\rm bc}$ or ΔE sideband region for $B^+ \to X(3872)K^+$ (top) and $B^0 \to X(3872)K^0$ (bottom). The left and right rows are for $\bar{D}^{*0} \to \bar{D}^0\gamma$ and $\bar{D}^{*0} \to \bar{D}^0\pi^0$, respectively. In each panel, the red solid line is the fit result with the generic background PDF. The bottom plot is the pull distribution.

correction is applied, and the associated uncertainty is assigned by performing fits with the resolution varied by the corresponding precision, $\pm 1\sigma \equiv \pm 13\%$. Here, because the broken-signal resolution of the $D^{*0} \to D^0\pi^0$ mode is also affected by the degradation of the resolution, it is varied simultaneously.

(iii) The systematic uncertainty arising from the mass dependence of the efficiency is evaluated by the quadratic sum of the changes induced by $\pm 1\sigma$ variations of the parameters p_1–p_3 in Eq. (5.1) and (5.8).

(iv) The systematic uncertainty arising from the ratio of the broken-signal to the signal is evaluated by the quadratic sum of the changes induced by $\pm 1\sigma$ variations of the corresponding parameters, e.g., p_0 in Eq. (5.1) and (5.8), and the yield of the broken-signal relative to the signal for $\overline{D}^{*0} \to \overline{D}^0\gamma$ obtained with the default signal MC sample. For the total broken-signal efficiency of $\overline{D}^{*0} \to \overline{D}^0\gamma$, a 1.7σ difference is found for a small width in the validation; See Sec. 5.2.2. Because the fitted width is large enough, we ignore the effect.

(v) The systematic uncertainty due to possible bias in the fit is evaluated by performing the pseudo experiments containing the signal and background components. The pseudo data is generated from the lineshape and the yield based on the data result. The input value of a parameter subtracted from the median of the parameter distribution is regarded as the corresponding uncertainty.

(vi) The uncertainty arising from the finite precision of the D^0 mass and the $\Delta(m_{D^{*0}} - m_{D^0})$ mass difference is taken as the $\pm 1\sigma$ uncertainty of $2m_{D^0} + \Delta(m_{D^{*0}} - m_{D^0}) = 3871.69 \pm 0.10$ MeV/c^2 following Ref. [18].

(vii) The nonzero D^{*0} width ($\Gamma_{D^{*0}}$) leads to three potential sources of bias: a bias arising from the mass difference technique, a bias arising from the consideration of the D^{*0} width in the lineshape model, and a bias due to the interference between $X(3872) \to D^0\overline{D}^{*0}$ and $\overline{D}^0 D^{*0}$. These three sources of bias are evaluated as follows. For the first bias, two $M(D^0\overline{D}^{*0})$ distributions are formed in MC with a broad lineshape: one where $m_{D^{*0}}$ in Eq. (4.4) is fixed to the nominal value (as in our analysis), and one where $m_{D^{*0}}$ is replaced by the true D^{*0} mass generated by `EvtGen`, where $\Gamma_{D^{*0}} = 65.5$ keV [42] is assumed. Each distribution is fitted with the PDF of the signal component, and the largest difference is regarded as the associated uncertainty. For the second bias, the distribution for data is fitted after smearing the assumed lineshape with a Breit-Wigner function of $\Gamma_{D^{*0}} = 65.5$ keV, and the change from the original result is regarded as the associated uncertainty. The third bias is ignored since the interference effect is negligible above the threshold [75]. The uncertainties associated with the first and second biases are added in quadrature.

(viii) Limited MC statistics lead to uncertainty on the shape of the broken-signal for $\overline{D}^{*0} \to \overline{D}^0\gamma$. This is evaluated by repeating the fit while varying each bin entry of the MC PDF histogram assuming Poisson distributions. The 68% interval of the distributions of the resulting fit values is used to assign the systematic uncertainty.

(ix) The uncertainty arising from the fixed ratio of the signal yields for $\overline{D}^{*0} \to \overline{D}^0\gamma$ to $\overline{D}^0\pi^0$ is evaluated by performing new fits, and varying the relative branching fraction between $D^{*0} \to D^0\gamma$ and $D^{*0} \to D^0\pi^0$ by $\pm 1\sigma$ [18]. The difference from the original result is treated as the uncertainty.

The product branching fraction is calculated as follows:

$$\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}) = \frac{N_{\rm sig}}{2N_{B\overline{B}}\mathcal{B}(\Upsilon(4S) \to B\overline{B})\epsilon_{\rm tot}}, \qquad (6.2)$$

where $\epsilon_{\rm tot}$ is the sum of products of the signal efficiency and the branching fraction of the intermediate states (Sec. 5.1.3), and $N_{B\overline{B}}$ is the number of $B\overline{B}$ meson pairs in the data. For $\mathcal{B}(\Upsilon(4S) \to B\overline{B})$, 0.514 and 0.486 are assigned for that of the B^+B^- mode and that of the $B^0\overline{B}^0$ mode, respectively [18]. The results are

$$\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}) = (0.97^{+0.21}_{-0.18}(\text{stat}) \pm 0.10(\text{syst})) \times 10^{-4},$$
$$\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}) = (1.30^{+0.36}_{-0.31}(\text{stat})^{+0.12}_{-0.07}(\text{syst})) \times 10^{-4},$$

where we consider eight sources of systematic uncertainty in addition to the sources (i)–(ix), as listed in Table 6.3:

(x) The uncertainty of the tracking efficiency is estimated using a $D^{*+} \to \pi^+(D^0 \to \pi^+\pi^- K^0_S)$ sample for tracks with high momentum. The efficiency is consistent in data and MC; the precision of the test, 0.35% per track, is taken as a systematic uncertainty. For tracks with low momentum, the sample of soft-charged pions from D^{*-} in the $B^0 \to D^{*-}\pi^+$ decay is used. The ratio of tracking efficiency obtained for MC and data is applied as a correction factor. The uncertainty in the correction factor is regarded as a systematic uncertainty.

(xi) The uncertainty of the efficiency of hadron identification is estimated using the inclusive $D^{*+} \to (D^0 \to K^-\pi^+)\pi^+$ sample. A correction factor is applied to the signal efficiency based on the ratio of the hadron identification efficiencies obtained for MC and data. The uncertainty in the correction factor is regarded as a systematic uncertainty.

(xii) The uncertainty of the efficiency of the K^0_S reconstruction is evaluated using the $D^{*+} \to (D^0 \to K^0_S\pi^0)\pi^+$ sample. A correction factor is applied to the signal efficiency based on the efficiency ratio obtained for MC and data. The uncertainty in the correction factor is regarded as a systematic uncertainty.

(xiii) The uncertainty of the efficiency of the π^0 detection is evaluated using the $\tau^- \to \pi^-\pi^0\nu_\tau$ sample. The ratio of the π^0 efficiency obtained for MC and data is applied to the signal efficiency as a correction factor. The uncertainty in the correction factor is regarded as a systematic uncertainty.

(xiv) The uncertainty of the efficiency of the γ detection is evaluated using the $B^+ \to (\chi_{c1} \to J/\psi\gamma)K^+$ sample: 3.0% is assigned for the $D^{*0} \to D^0\gamma$ decay mode.

(xv) The uncertainty of $\sum \epsilon \times \mathcal{B}_{\rm IMS}$ mainly arises from the uncertainties on the D^0 branching fraction, and the limited size of the signal MC sample. In addition, validation of the calculation method for the total signal efficiency shows input-output differences in the B^0 decay mode larger than expected from statistical fluctuations: the largest of these is assigned as a systematic uncertainty. These sources of uncertainty are added in quadrature.

(xvi) The number of $B\overline{B}$ pairs in the data set is measured to be $(772 \pm 11) \times 10^6$: the associated uncertainty is set to 1.4%.

(xvii) The uncertainties on the branching fractions $\mathcal{B}(\Upsilon(4S) \to B^+B^-) = (51.4 \pm 0.6)\%$ and $\mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0) = (48.6 \pm 0.6)\%$ [18] are also included. In Ref. [18], $\mathcal{B}(\Upsilon(4S) \to B\overline{B})$ is assumed to be one; therefore the correlation between $\mathcal{B}(\Upsilon(4S) \to B^+B^-)$ and $\mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0)$ is assumed to be fully negative.

The relative branching fraction between $B^0 \to X(3872)K^0$ and $B^+ \to X(3872)K^+$ is measured as

$$\frac{\mathcal{B}(B^0 \to X(3872)K^0)}{\mathcal{B}(B^+ \to X(3872)K^+)} = 1.34^{+0.47}_{-0.40}(\text{stat})^{+0.10}_{-0.12}(\text{syst}),$$

with the same sources of systematic uncertainty as for the branching fractions; some uncertainties cancel, or partially cancel, in the ratio.

Table 6.2: Summary of systematic uncertainty for the mass and the width of the relativistic Breit-Wigner lineshape.

	Source	m (MeV/c^2)	Γ_0 (MeV)
(i)	Generic BG PDF	± 0.07	± 0.38
(ii)	Mass Resolution	± 0.02	$-0.11/+0.13$
(iii)	Mass dependence of efficiency	± 0.02	$-0.08/+0.07$
(iv)	Ratio of broken-signal BG to signal	± 0.01	± 0.02
(v)	Fit bias	$-0.02/+0.00$	$-0.02/+0.00$
(vi)	D^{*0} and D^0 masses	± 0.10	$\cdots$
(vii)	D^{*0} width	$-0.01/+0.02$	± 0.02
(viii)	Broken-signal shape for $\overline{D}^{*0} \to \overline{D}^0\gamma$	± 0.00	± 0.01
(ix)	Signal ratio of $\overline{D}^{*0} \to \overline{D}^0\gamma$ to $\overline{D}^0\pi^0$	± 0.01	± 0.05
	Total	± 0.13	± 0.4

Table 6.3: Summary of systematic uncertainty for the branching fraction measurements using the relativistic Breit-Wigner lineshape.

	Source	$X(3872)K^+$ (%)	$X(3872)K^0$ (%)	Ratio(K^0/K^+) (%)
(i)	Generic BG PDF	± 8.2	± 1.4	± 6.7
(ii)	Mass Resolution	$-0.2/+0.4$	$-0.3/+0.4$	$-0.1/+0.0$
(iii)	Mass dependence of efficiency	$-2.7/+2.0$	$-2.3/+1.7$	$-0.5/+0.6$
(iv)	Ratio of broken-signal BG to signal	± 2.1	± 0.6	± 2.1
(v)	Fit bias	$-1.3/+0.0$	$-7.3/+0.0$	$-4.5/+0.0$
(vi)	D^{*0} and D^0 masses	$\cdots$	$\cdots$	$\cdots$
(vii)	D^{*0} width	± 0.0	± 0.0	± 0.0
(viii)	Broken-signal shape for $\overline{D}^{*0} \to \overline{D}^0\gamma$	± 0.1	± 0.1	± 0.1
(ix)	Signal ratio of $\overline{D}^{*0} \to \overline{D}^0\gamma$ to $\overline{D}^0\pi^0$	± 0.8	± 0.2	± 0.6
(x)	Tracking efficiency	± 2.1	± 2.4	± 0.3
(xi)	PID efficiency	± 2.9	± 2.4	± 0.4
(xii)	K^0_S reconstruction	± 0.2	± 1.0	± 0.8
(xiii)	π^0 reconstruction	± 1.9	± 1.9	$\cdots$
(xiv)	γ reconstruction	± 1.5	± 1.5	$\cdots$
(xv)	$\sum \epsilon \times \mathcal{B}_{\rm IMS}$	± 1.4	$-3.1/+2.3$	$-1.7/+0.9$
(xvi)	$N_{B\overline{B}}$	± 1.4	± 1.4	$\cdots$
(xvii)	$\mathcal{B}(\Upsilon(4S) \to B\overline{B})$	± 1.2	± 1.2	± 2.4
	Total	± 10	$-9.6/+5.7$	$-9.0/+7.6$

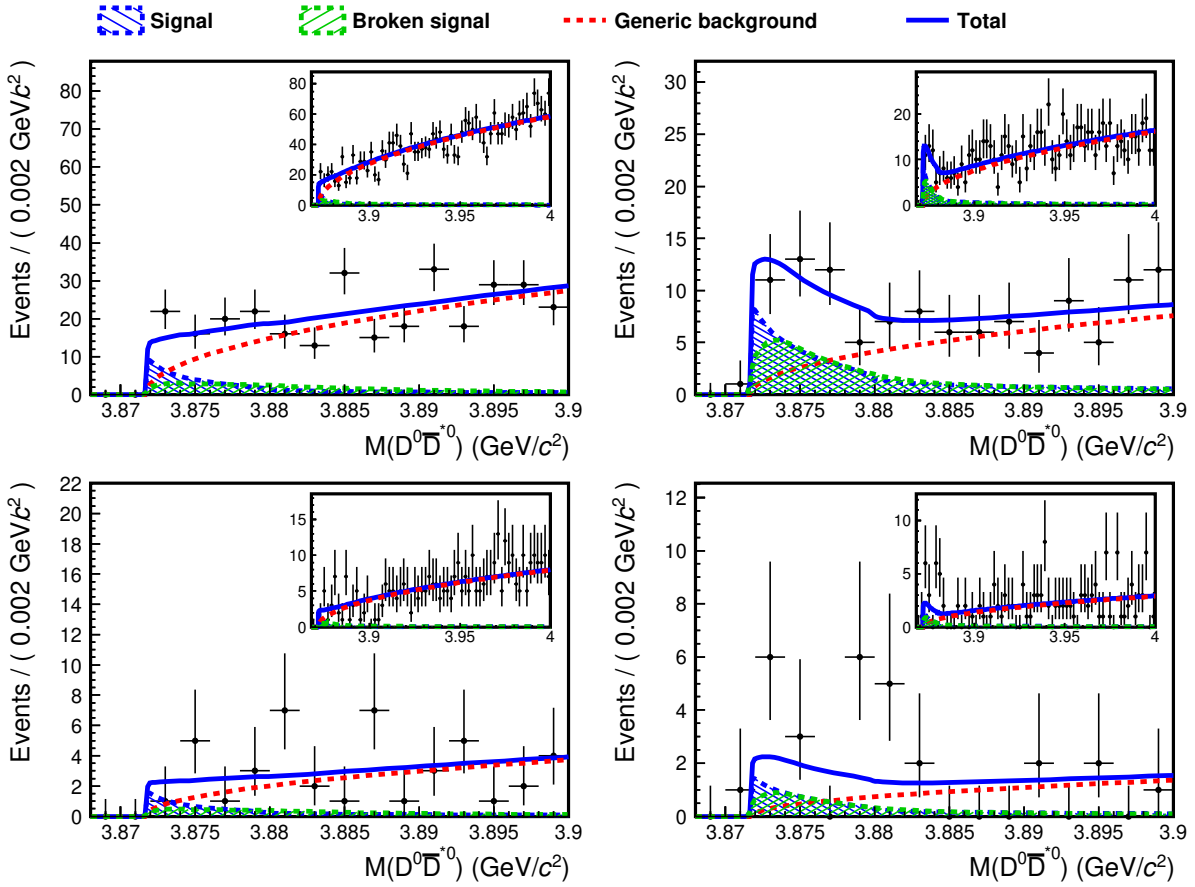


Figure 6.3: The $M(D^0\overline{D}^{*0})$ distributions with the fit result with the Flatté lineshape for $B^+ \to X(3872)K^+$ (top) and $B^0 \to X(3872)K^0$ (bottom). The left and right rows are for $\overline{D}^{*0} \to \overline{D}^0\gamma$ and $\overline{D}^{*0} \to \overline{D}^0\pi^0$, respectively. The points with error bars represent the data. The solid blue line show the total fit result. The dashed blue and green lines show the signal and broken-signal contributions, respectively. The dashed red line shows the generic background contribution.

6.2 Flatté Lineshape

The fit result for the data is shown in Fig. 6.3 and Table 6.4. The fitted g is $0.29^{+2.69}_{-0.15}$, where the uncertainty is statistical. A systematic uncertainty arising from the analysis method, e.g., the sources (i)-(xvi), are summarized in Table 6.5. The method to evaluate the uncertainty due to sources from (i) to (ix) is the same as in the measurement of the relativistic Breit-Wigner lineshape. Sources (x) to (xvii) also contribute through the constraint on the branching fraction applied in the fit. They are evaluated by performing fits after varying each parameter by $\pm 1\sigma$, and adding the resulting changes in quadrature. Regarding the fitter bias (vii), the bias causes the upper systematic uncertainty, however, its magnitude was not determined based on a study of the pseudo experiments; Details are described in Sec. 5.4.2. With the total lower systematic uncertainty, therefore, the lower limit is determined using the likelihood profile. The likelihood including the systematic uncertainties listed in Table 6.5, $L(g)$, is shown as the black solid line in Fig. 6.4. Noting that the curve is asymmetric, with a larger integral above than below the best fit value,

Table 6.4: Results using the Flatté lineshape: the fitted coupling constant g, and the signal yield.

Mode	g	$N_{\rm sig}$
Combined	$0.29^{+2.69}_{-0.15}$	$90.9^{+11.3}_{-15.9}$
$X(3872)K^+$	—	$77.9^{+9.6}_{-13.5}$
$X(3872)K^0$	—	$13.0^{+3.0}_{-2.9}$

we conservatively set the lower limit at 90% or 95% credibility $g_{\rm lower}$ from

$$\begin{aligned} \int_{g_{\rm lower}}^{g_{\rm best}} L(g)dg &= 0.8 \int_0^{g_{\rm best}} L(g)dg \quad \text{for 90\% credibility,} \\ \int_{g_{\rm lower}}^{g_{\rm best}} L(g)dg &= 0.9 \int_0^{g_{\rm best}} L(g)dg \quad \text{for 95\% credibility,} \end{aligned} \tag{6.3}$$

where $g_{\rm best}$ denotes the coupling constant at the maximum likelihood; the derivation of these requirements is described in Appendix E. The effect of fixing dg/dE_f, f_ρ, and Γ_0 to the values measured by LHCb is studied by varying each parameter by $\pm 1\sigma$. Separate curves of the relative likelihood L/L_0 for each case are also shown in Fig. 6.4, where $L = L(g)$ is the likelihood of the fit and L_0 is the likelihood of the best fit for each parameter set. The corresponding fit results and lower limits are summarized in Table 6.6. The L_0 values for the different parameter sets vary in a small range around the value for set (1): the best is favoured by only 1.2σ relative to set (1), and the worst is disfavoured by 3.4σ. The loosest lower limit is obtained for the parameter set (6), one of the disfavoured scenarios, where f_ρ is changed by $+1\sigma$. We conservatively choose this as the final lower limits for this study:

$$\begin{aligned} &g > 0.094 \text{ at 90\% credibility, and} \\ &g > 0.075 \text{ at 95\% credibility.} \end{aligned}$$

These limits correspond to

$$\begin{aligned} &E_f < -6.2 \text{ MeV at 90\% credibility, and} \\ &E_f < -5.0 \text{ MeV at 95\% credibility,} \end{aligned}$$

which is derived from $dg/dE_f = -15.11$ GeV^{-1}.

Table 6.5: Summary of systematic uncertainties for the coupling constant g of the Flatte lineshape.

	Source	g
(i)	Generic BG PDF	$< O(0.001)$
(ii)	Mass Resolution	$-0.011/+0.003$
(iii)	Mass dependence of efficiency	$-0.012/+0.024$
(iv)	Ratio of broken-signal BG to signal	$-0.007/+0.020$
(v)	Fit bias	$-0.000/+\infty$
(vi)	D^{*0} and D^0 masses	$\cdots$
(vii)	D^{*0} width	$-0.006/+0.001$
(viii)	Broken-signal shape for $\overline{D}^{*0} \to \overline{D}^0\gamma$	$-0.001/+0.002$
(ix)	Signal ratio of $\overline{D}^{*0} \to \overline{D}^0\gamma$ to $\overline{D}^0\pi^0$	$-0.000/+0.004$
(x)-(xvii)	Branching fraction	$-0.021/+0.042$
	Total	$-0.029/+\infty$

Table 6.6: Summary of the seven parameter sets used in the evaluation of lower limits on the coupling constant g, showing the g of the best fit, the g lower limits, and corresponding E_f upper limits. The parameter sets are the center values of dg/dE_f, Γ_0, and f_ρ measured at LHCb [4] (1), changing dg/dE_f by $+1\sigma$ (2), changing dg/dE_f by -1σ (3), changing Γ_0 by $+1\sigma$ (4), changing Γ_0 by -1σ (5), changing f_ρ by $+1\sigma$ (6), and changing f_ρ by -1σ (7). For the parameter set (7), no lower limit is determined, because no best fit is found in the range $g < 50$.

Parameter set	(1)	(2)	(3)	(4)	(5)	(6)	(7)
dg/dE_f (GeV^{-1})	-15.11	-14.95 $(+1\sigma)$	-15.27 (-1σ)	-15.11	-15.11	-15.11	-15.11
Γ_0/E_f	-0.19	-0.19	-0.19	-0.29 $(+1\sigma)$	-0.09 (-1σ)	-0.19	-0.19
f_ρ/E_f (GeV^{-1})	-0.25	-0.25	-0.25	-0.25	-0.25	-0.38 $(+1\sigma)$	-0.12 (-1σ)
g of best fit	0.29	0.27	0.31	0.21	0.46	0.17	> 50
g lower limit at 90% CL	> 0.143	> 0.136	> 0.151	> 0.105	> 0.212	> 0.094	—
at 95% CL	> 0.113	> 0.108	> 0.119	> 0.082	> 0.167	> 0.075	—
E_f upper limit at 90% CL (MeV)	< -9.5	< -9.0	< -10.0	< -6.9	< -14.0	< -6.2	—
at 95% CL (MeV)	< -7.6	< -7.2	< -7.9	< -5.5	< -11.1	< -5.0	—

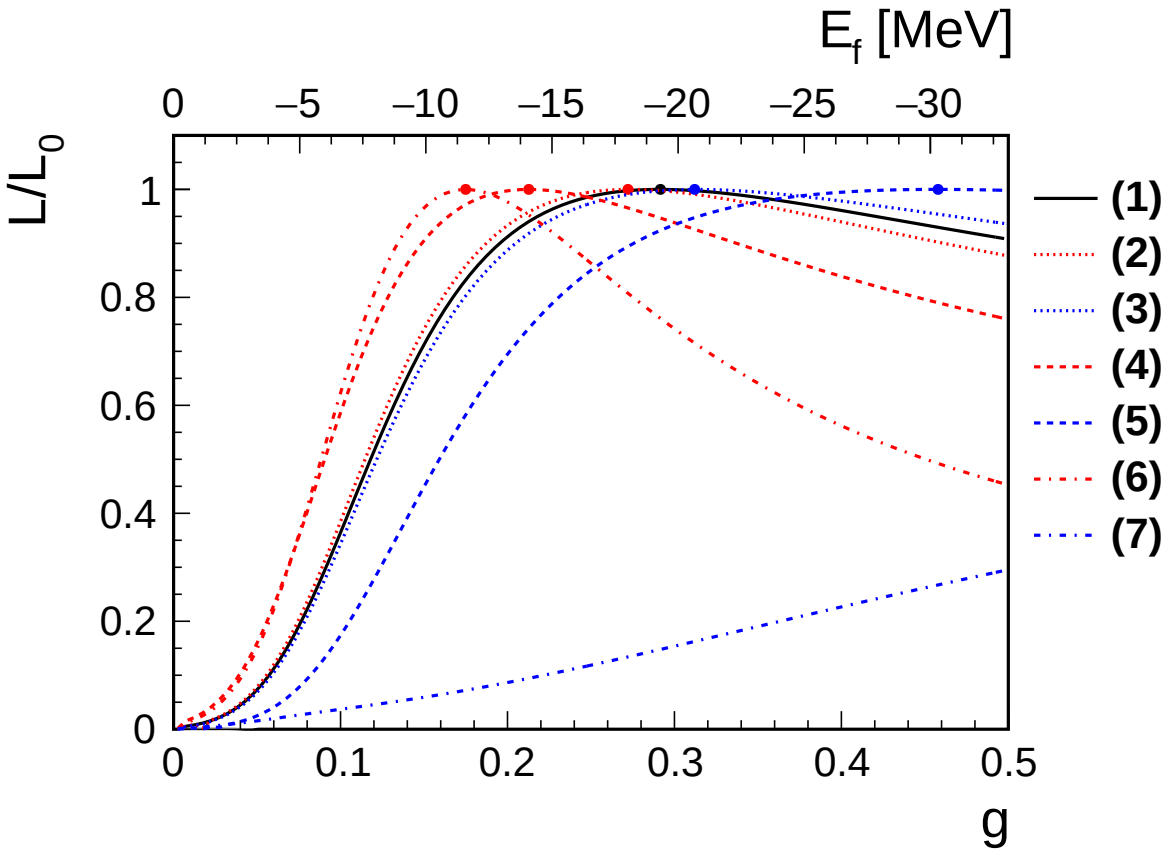


Figure 6.4: For each of seven parameter sets, the likelihood ratio L/L_0 is shown, as a function of the coupling constant g, where $L = L(g)$ is the fitted likelihood and L_0 is the likelihood of the best fit for that parameter set. The solid black line shows the parameter set (1). The red and blue dotted lines show parameter sets (2) and (3), respectively. The red and blue dashed lines show sets (4) and (5), and the red and blue dot-dashed lines show sets (6) and (7), respectively. The parameter sets are described in Table 6.6. Circles on the lines show the best fit g.

6.3 Comparison between Breit-Wigner and Flatté Lineshapes

We investigate which lineshape model best fits the $M(D^0\bar{D}^{*0})$ distribution better using a test statistic with $t = -2\ln(\mathcal{L}_{\text{B-W}}/\mathcal{L}_{\text{Flatte}})$. Here, $\mathcal{L}_{\text{B-W}}$ is the best fit likelihood for the Breit-Wigner lineshape, and $\mathcal{L}_{\text{Flatte}}$ is the best likelihood without the $R_{D\overline{D}^*}$ constraint term for the Flatte lineshape in the parameter setting (1). For data, we obtain $t = -8.5$, i.e. the Breit-Wigner lineshape is favored. Based on the t distribution obtained from pseudo experiments, the exclusion level for the Flatté lineshape is only 2.2σ; this level declines when the systematic uncertainties are taken into account. Thus, neither lineshape can be excluded.

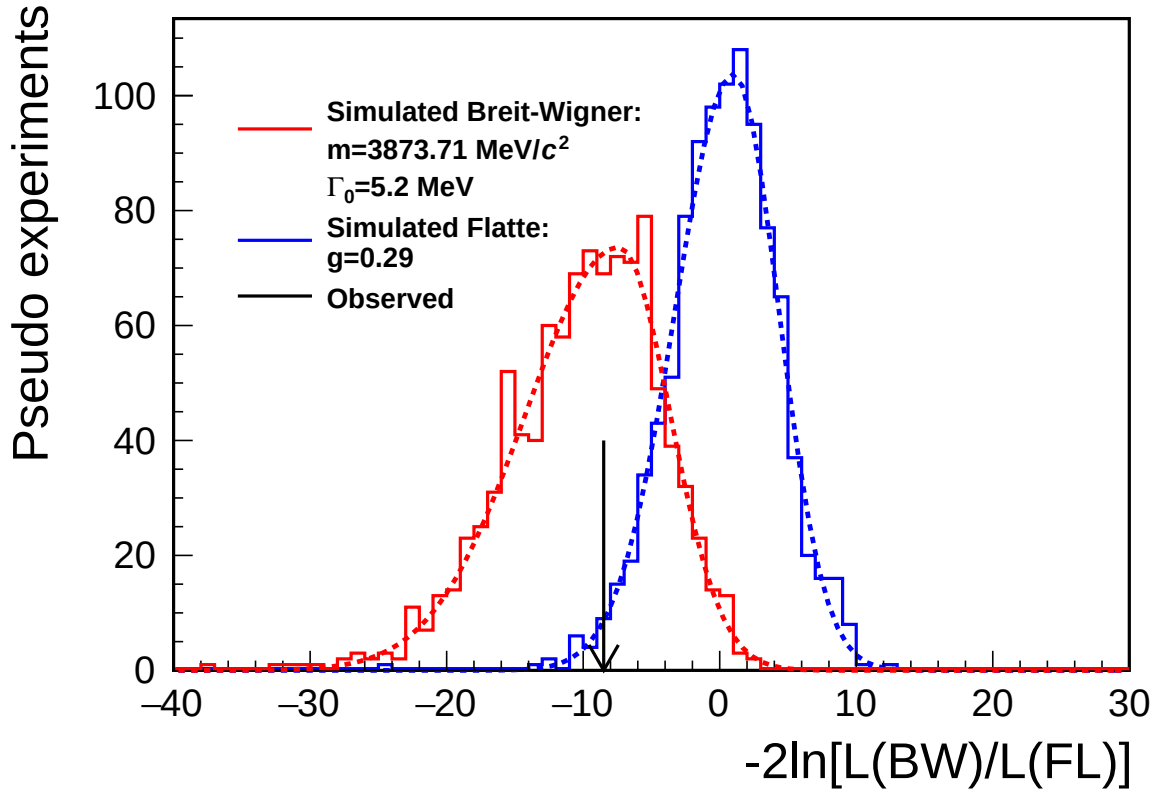


Figure 6.5: The t distributions obtained from pseudo experiments with a hypothesis of the Breit-Wigner lineshape (red) and the Flatte lineshape (blue). For each type of pseudo experiments, we use the parameters obtained from the data. The dashed curve is the fit results with an asymmetric Gaussian: ($\mu = -7.4 \pm 0.4$, $\sigma_{\text{upper}} = 3.7 \pm 0.3$, and $\sigma_{\text{lower}} = 6.9 \pm 0.3$) for the simulated Breit-Wigner hypothesis, and ($\mu = 1.0 \pm 0.3$, $\sigma_{\text{upper}} = 3.3 \pm 0.2$, and $\sigma_{\text{lower}} = 4.3 \pm 0.2$) for the simulated Flatte hypothesis. Here, μ denotes a mean, and σ_{upper} and σ_{lower} are standard deviations for the upper and lower sides, respectively. The black arrow is the t value observed for the data. Any systematic uncertainty is not taken into account in this figure.

Chapter 7

Discussion

7.1 Result of This Study

We examine the $X(3872)$ lineshape on the $D^0\overline{D}^{*0}$ decay distribution using the full data at Belle. When fitting it with the relativistic Breit-Wigner lineshape, its mass and width are measured as

$$m = 3873.71^{+0.56}_{-0.50}(\text{stat}) \pm 0.13(\text{syst})\ \text{MeV}/c^2,$$
$$\Gamma_0 = 5.2^{+2.2}_{-1.5}(\text{stat}) \pm 0.4(\text{syst})\ \text{MeV}.$$

Comparison with the previous studies using the $D^0\overline{D}^{*0}$ decay [2, 3, 56] is shown in the first and second panels from the top of Fig. 7.1. It indicates that the precision of the measurement is improved by at least 22%, and the values are in good agreement with those measured in the previous studies. The measured branching fractions are as follows:

$$\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}) = (0.97^{+0.21}_{-0.18}(\text{stat}) \pm 0.10(\text{syst})) \times 10^{-4},$$
$$\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}) = (1.30^{+0.36}_{-0.31}(\text{stat})^{+0.12}_{-0.07}(\text{syst})) \times 10^{-4}.$$

It is the first observation of the signal from B^0 decays with more than 5σ significance. The ratio of the branching fractions is determined to be

$$\frac{\mathcal{B}(B^0 \to X(3872)K^0)}{\mathcal{B}(B^+ \to X(3872)K^+)} = 1.34^{+0.47}_{-0.40}(\text{stat})^{+0.10}_{-0.12}(\text{syst}).$$

The properties of the branching fractions are in good agreement with the previous studies using the $D^0\overline{D}^{*0}$ decay (from the third to the last panels from the top of Fig. 7.1).

We compare these results with the analysis of the Breit-Wigner lineshape using the $J/\psi\pi^+\pi^-$ decay mode to date. The measured Breit-Wigner mass is significantly higher than the $D^0\overline{D}^{*0}$ threshold, while the world-average mass with the $J/\psi\pi^+\pi^-$ decay is consistent with the threshold. The measured width and $\mathcal{B}(B^0 \to X(3872)K^0)/\mathcal{B}(B^+ \to X(3872)K^+)$ are shifted from the average with the $J/\psi\pi^+\pi^-$ decay by 2.6σ and 2.0σ, respectively [45]. This study supports the trend that these properties in the $X(3872) \to D^0\overline{D}^{*0}$ decay mode differ from those in the $J/\psi\pi^+\pi^-$ decay mode as shown by the previous studies using the $D^0\overline{D}^{*0}$ decay. It means that the Breit-Wigner mass and width cannot describe the $X(3872)$ state universally.

The lineshape distortion is explained by the coupled-channel effect. In this study, we also examine the lineshape using a model considering a coupled channel effect, Flatté lineshape. In such a model, it is difficult to determine all the parameters with the narrow

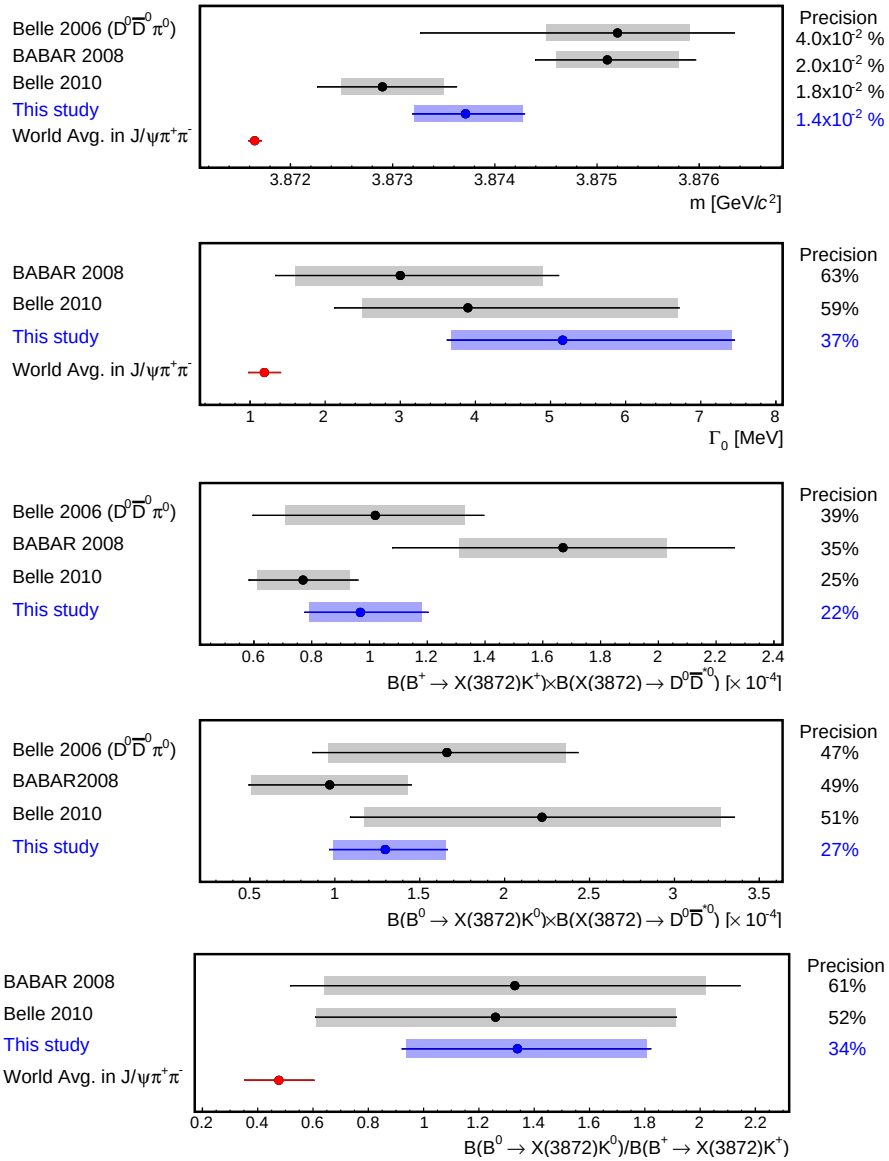


Figure 7.1: Comparison of the relativistic Breit-Wigner results obtained in this study (blue), the previous studies in the $D^0\overline{D}^{*0}$ decays (black) and the world-average values in the $J/\psi\pi^+\pi^-$ decays (red). From the top panel, results for m, Γ_0, $\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0})$, $\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0})$, and $\mathcal{B}(B^0 \to X(3872)K^0)/\mathcal{B}(B^+ \to X(3872)K^+)$ are shown. The error bar of each point is the total uncertainty. The filled area of each point shows the statistical uncertainty. The previous studies with parentheses show the results using the $X(3872)$ decay mode written in parentheses. The value on the right shows the precision of each measurement. Here, the systematic and statistical uncertainties are combined by taking the square root of the sum of squares. If the upper and lower uncertainties are different, they are averaged.

peak of the $J/\psi\pi^+\pi^-$ mode due to the scaling behavior and the detector resolution, as reported by the LHCb analysis [4]. Thus, based on the analysis with the $J/\psi\pi^+\pi^-$ decay at LHCb, we established the method to measure the coupling constant to the $D\overline{D}^*$ channel g as the undetermined parameter. As a result, we find that the fitted value of g is in a region that is relatively insensitive to the underlying value. We determine its lower limit to be

$$g > 0.094 \text{ at } 90\% \text{ credibility.}$$

$$g > 0.075 \text{ at } 95\% \text{ credibility.}$$

It shows the coupled-channel effect cannot be ignored in the $X(3872)$ lineshape analysis. This limit at 90% credibility corresponds to an upper limit $E_f < -6.2$ MeV, which is slightly more stringent than the LHCb measurement, -270 MeV $< E_f < -2.0$ MeV at 90% credibility [4]. This suggests that analysis using $D^0\overline{D}^{*0}$ can indeed complement the study of the $J/\psi\pi^+\pi^-$ mode in this framework. The limit includes the solution $E_f = -7.2$ MeV assumed in the scattering amplitude analysis at LHCb. There is still uncertainty in the pole positions of the scattering amplitude, because the limit is not so stringent.

Both lineshapes fit the invariant mass distribution obtained from the data. Finally, we examine which lineshape model fits the invariant mass distribution. Based on a likelihood ratio from the fits, the Breit-Wigner lineshape is favored, but the Flatté lineshape is not significantly excluded.

Analysis of a large dataset will be important, because the statistical uncertainty dominates in both of the lineshape measurements. It would be accomplished by adding more D^0 modes to reconstruct and using higher statistical data. Such an analysis is possible in the Belle II experiments. Moreover, we need a more sensitive analysis, especially for the Flatté lineshape measurements. The dominant sources of systematic uncertainty in our measurement and their necessary improvements are as follows.

- Systematic uncertainty due to the fitter bias: Because the bias decreases as data size increases, the data size needs to be improved.

- Systematic uncertainty due to f_ρ and Γ_0: Because the parameters contribute differently to lineshapes for both $J/\psi\pi^+\pi^-$ and $D^0\bar{D}^{*0}$ decays[*1], and the relative branching fraction $\mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)/\mathcal{B}(X(3872) \to D^0\overline{D}^{*0})$, it is necessary to improve all of their precision.

Therefore, improving the statistic of the $X(3872) \to D^0\bar{D}^{*0}$ decay sample is essential to suppress these uncertainties, too. Especially for a measure of the latter uncertainty, a simultaneous fit of the $J/\psi\pi^+\pi^-$ and $D^0\overline{D}^{*0}$ decay modes is also useful to handle all of the information effectively. For the $J/\psi\pi^+\pi^-$ sample, an exclusive $B^+ \to (X(3872) \to J/\psi\pi^+\pi^-)K^+$ sample at LHCb is the most appropriate from viewpoints of statistic. Here, the exclusive B decay is a key to applying the constraint of the relative branching fraction. Furthermore, in the future, resonance energy scans with a high-resolution beam are planned for the PANDA experiment [76]. The resolution is several ten times better than the mass resolution at LHCb. It will provide detailed lineshapes of the $X(3872)$ lineshape in the $J/\psi\pi^+\pi^-$ decays, which was hidden by the detector

[*1] If f_ρ is varied, the lineshapes in both $J/\psi\pi^+\pi^-$ and $D^0\bar{D}^{*0}$ decays barely vary, whereas $\mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)/\mathcal{B}(X(3872) \to D^0\overline{D}^{*0})$ aries in proportion to f_ρ. If Γ_0 is varied, not only the lineshape widths in both $J/\psi\pi^+\pi^-$ and $D^0\bar{D}^{*0}$ decays but also $\mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-)/\mathcal{B}(X(3872) \to D^0\overline{D}^{*0})$ vary.

resolution at LHCb. These could fully determine the lineshape in the coupled-channel framework, and greatly contribute to determining the internal structure.

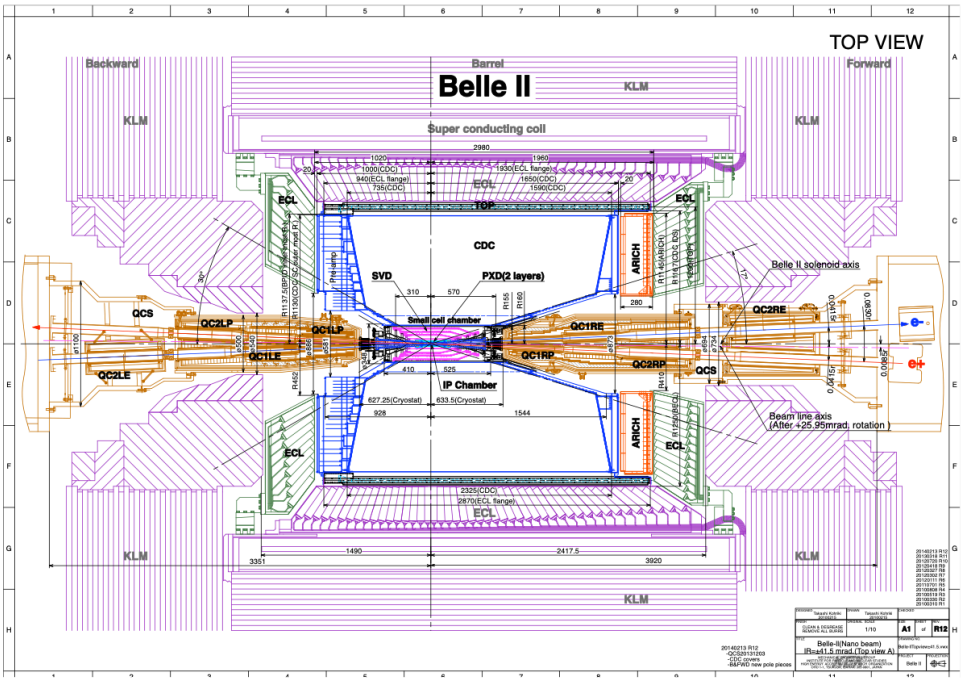


Figure 7.2: Top view of the Belle II detector [79].

7.2 Prospect for the Belle II Experiment

The Belle II experiment [77] is an advanced B-factory experiment using an energy-asymmetric e^+e^- collider, SuperKEKB [78]. Our goal is to advance our understanding of particle physics broadly, e.g., search for physics beyond the Standard Model, precise measurements of electroweak interaction parameters, and exploration of properties of the strong interaction; The planned physics projects are summarized in Ref. [79]. Collision data began to be collected with the Belle II detector in 2019, aiming to acquire 50 times more data than Belle.

The configuration of the Belle II detector is shown in Fig. 7.2. It consists of seven sub-detectors: a two-layer silicon pixel detector (PXD), a four-layer double-sided silicon strip detector (SVD) and a 56-layer central drift chamber (CDC), a time-of-propagation counter in the barrel region of the detector (TOP), an aerogel ring-imaging Cherenkov counter in the endcap region (ARICH), an electromagnetic calorimeter comprised of CsI(Tl) crystals (ECL), and a K_L^0 and muon detector (KLM). All the sub-detectors were upgraded to tolerate the high background environment associated with an increase of instantaneous luminosity, improve the resolutions of the vertex position and momentum measurements, improve the particle identification performance, and so on.

At Belle II, photons and π^0's can be reconstructed as well as the Belle detector [59]. It provides the best environment for the analysis in the $B \to K(X(3872) \to D^0\overline{D}^{*0})$ decay. As the first milestone of this program, we searched $B \to K(X(3872) \to J/\psi\pi^+\pi^-)$ with the initial data collected with the Belle II detector. Furthermore, the tools for the Belle II distributed computing system were developed to perform physics analysis efficiently with this system, which is important for all physics analyses at Belle II. The subsequent subsections describe summaries of the achievements.

7.2.1 Revisiting $X(3872)$ at Belle II

We searched $B \to K(X(3872) \to J/\psi\pi^+\pi^-)$ using 62.8 fb^{-1} data collected at a center-of-mass system energy of 10.58 GeV with the Belle II detector at the SuperKEKB accelerator in 2019. This subsection describes the event reconstruction and selection, the validation using the control sample, and the result of the $X(3872)$ search.

Event Reconstruction and Selection

We analyze only events that passed the following preselection. The number of good tracks with $p_T < 0.2$ GeV/c, $|dr| < 2.0$ cm and $|dz| < 4.0$ cm is greater than three, where p_T, dr and dz are a transverse momentum, a signed distance of the POCA in the r–ϕ plane and that along the z-axis, respectively.

For the track reconstruction except K^0_S, well-measured tracks, $|dr| < 1.0$ cm and $|dz| < 3.0$ cm, are used. In addition, charged track candidates are selected using the likelihood $\mathcal{L}_h$ for a particle hypothesis h, given based on particle identification information of all sub-detectors. Moreover, for charged kaon candidates, the kaon identification likelihood ratio, $\mathcal{L}_K/(\mathcal{L}_\pi + \mathcal{L}_K + \mathcal{L}_p + \mathcal{L}_e + \mathcal{L}_\mu + \mathcal{L}_d)$, is required to be greater than 0.1. For charged lepton candidates, the lepton identification likelihood ratio, $\mathcal{L}_l/(\mathcal{L}_\pi + \mathcal{L}_K + \mathcal{L}_p + \mathcal{L}_e + \mathcal{L}_\mu + \mathcal{L}_d)$, is required to be greater than 0.5 ($l = e, \mu$). Only for electron candidates, the energies of the bremsstrahlung photons in the vicinity of the tracks of the electron candidates are recovered for the momentum calculation.

The candidates of K^0_S mesons are reconstructed from a pair of oppositely-charged tracks, for which the pion mass is assumed. The candidates are selected by the criteria described in Ref. [80]. The $\pi^+\pi^-$ invariant mass is required to be in the range of 4σ from the peak mean, 490 MeV/$c^2 < M(\pi^+\pi^-) <$ 506 MeV/c^2.

The candidates of J/ψ mesons are reconstructed from a pair of oppositely-charged lepton tracks. We select the J/ψ candidates within around 3σ from the peak mean for the $\mu^+\mu^-$ decay mode, i.e., 3.070 GeV/$c^2 < M(\mu^+\mu^-) <$ 3.117 GeV/c^2. For the e^+e^- decay mode, 3.065 GeV/$c^2 < M(e^+e^-) <$ 3.117 GeV/c^2 is required to consider the tail component. In order to improve the ΔE and $M_{\rm bc}$ resolutions, a mass-constrained fit of J/ψ candidates is applied.

The candidates of B mesons are reconstructed in the decay modes of $B^+ \to J/\psi\pi^+\pi^-K^+$ and $B^0 \to J/\psi\pi^+\pi^-K^0_S$. The B candidates with $M_{\rm bc} > 5.27$ GeV/c^2 and $|\Delta E| < 0.02$ GeV are selected. To reduce more background in the $X(3872)$ signal region, a requirement of the $\pi^+\pi^-$ invariant mass is useful [31]. The same requirement as [31], $M(\pi^+\pi^-) - M(l^+l^-\pi^+\pi^-) + m_{J/\psi} > -0.150$ GeV/c^2, is adopted only for the later $X(3872)$ analysis. Here, $M(\pi^+\pi^-)$, $M(l^+l^-\pi^+\pi^-)$ and $m_{J/\psi}$ mean the $\pi^+\pi^-$ invariant mass, the $l^+l^-\pi^+\pi^-$ invariant mass and the nominal J/ψ mass, respectively. In order to suppress background contributed from continuum events, we also require $R2 < 0.4$, where $R2$ is the normalized Fox-Wolfram moment [81]. After reconstructing the $B \to KJ/\psi\pi^+\pi^-$ decay, the best candidate selection is used to reduce the multiplicity of the B candidates. It is performed by selecting one candidate with the lowest $|\Delta E|$ per one event.

Control Sample Study using $B \to \psi(2S)K$

The control sample study is performed as a validation analysis for the rediscovery work. The $J/\psi\pi^+\pi^-$ distribution for data is shown in Fig. 7.3. In order to extract the $\psi(2S)$ events, we perform an Unbinned Extended Maximum-likelihood fit. A triple Gaussian

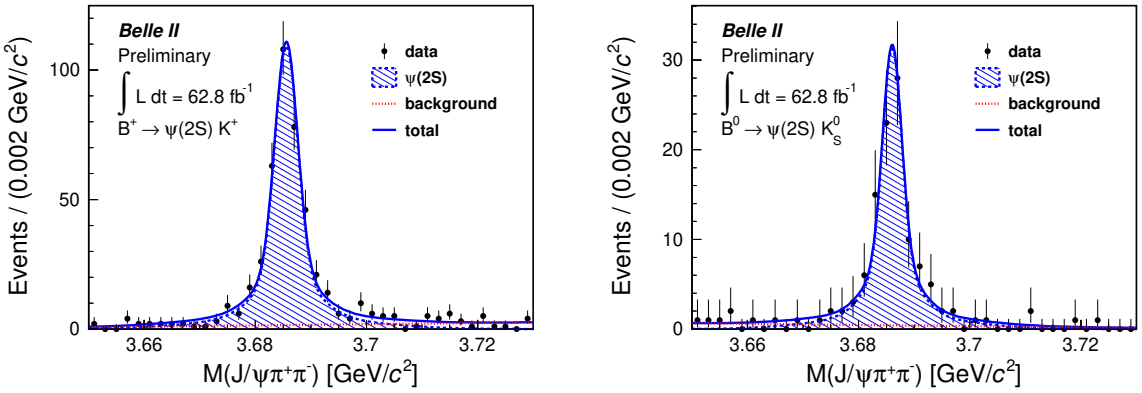


Figure 7.3: The $M(J/\psi\pi^+\pi^-)$ distributions on the $\psi(2S)$ signal region with the real data. Here, the solid blue line shows the total fit result. The striped blue area and the dotted red line represent the signal and background contributions, respectively.

Table 7.1: $B \to \psi(2S)K$ signal yields and branching fraction for data.

	$B^+ \to \psi(2S)K^+$	$B^0 \to \psi(2S)K^0$
Signal yield with data	409.07 ± 23.4	104.30 ± 11.5
Signal efficiency ϵ (%)	22.69 ± 0.16	17.40 ± 0.17
Branching fraction	6.21 ± 0.36	6.31 ± 0.70
World-average Branching fraction [18]	$(6.24 \pm 0.20) \times 10^{-4}$	$(5.8 \pm 0.5) \times 10^{-4}$
Measured / World average	0.995 ± 0.066	1.08 ± 0.15

with a common mean is used as a PDF for $\psi(2S)$ contributions, where parameters except the mean and scaling factor of the standard deviations are fixed to the values determined with the signal MC sample. The reason why the Breit-Wigner function is not used is that the total width of $\psi(2S)$ is almost zero, and thus the lineshape is determined by the resolution only. For combinatorial background contributions, a first-order Chebyshev function is used. The yield obtained by the fit is shown in Table 7.1. To ensure the signal efficiency in the data, which is essential for the later search of $B \to X(3872)K$ decay, we compare the measured branching fraction with the world average,

$$\mathcal{B}(B \to \psi(2S)K) = \frac{N_{\rm sig}}{2N_{B\overline{B}}\mathcal{B}(\Upsilon(4S) \to B\overline{B})\ \mathcal{B}(\psi(2S) \to J/\psi\pi^+\pi^-)\mathcal{B}(J/\psi \to l^+l^-)\epsilon f_K}, \tag{7.1}$$

where $N_{\rm sig}$ is signal yields for data and $N_{B\overline{B}}$ is the number of B-meson pairs in the data sets, measured as $(68.2 \pm 0.01) \times 10^6$. The symbol $\mathcal{B}$ denotes branching fractions in the parenthesis: $\mathcal{B}(\Upsilon(4S) \to B^+B^-) = (51.4 \pm 0.6)\%$, $\mathcal{B}(\Upsilon(4S) \to B^0\overline{B}^0) = (48.6 \pm 0.6)\%$ and $\mathcal{B}(\psi(2S) \to l^+l^-) = (11.03 \pm 0.05)\%$. ϵ means the signal efficiency determined using the signal MC sample. f_K is a fraction of the kaon decay mode: $f_K = 1$ for B^+ decay mode and $f_K = 0.346$ for B^0 decay mode. The measured decay branching fractions are

$$\mathcal{B}(B^+ \to \psi(2S)K^+) = (6.21 \pm 0.36) \times 10^{-4},$$
$$\mathcal{B}(B^0 \to \psi(2S)K^0) = (6.31 \pm 0.70) \times 10^{-4},$$

where uncertainty is statistical only. They are in good agreement with the world average within the statistical uncertainty. It indicates little difference in the signal efficiency between data and MC in $B \to J/\psi\pi^+\pi^-K$ decay.

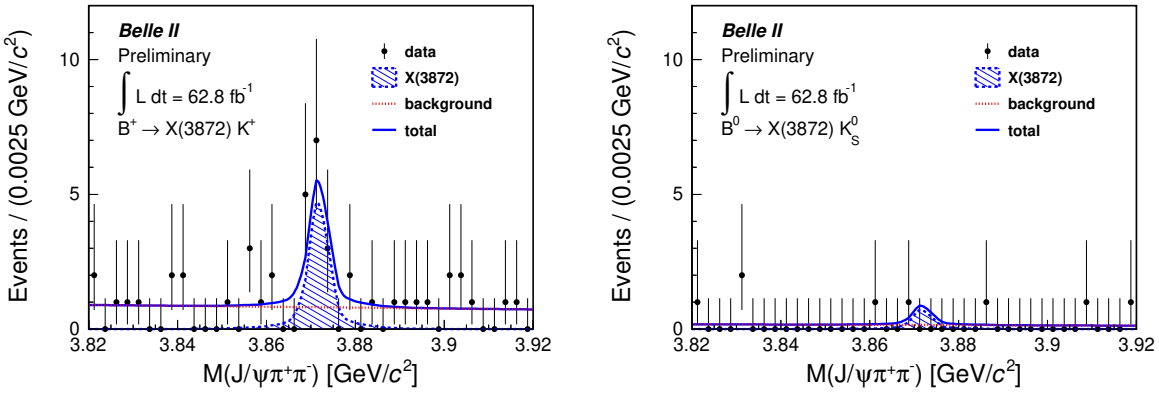


Figure 7.4: The $M(J/\psi\pi^+\pi^-)$ distributions on the $X(3872)$ signal region with the real data. Here, the solid blue line shows total fit result. The striped blue area and the dotted red line represent the signal and background contributions, respectively.

Search of $B \to X(3872)K$ decay

Since little difference in the signal efficiency between data and MC in $B \to J/\psi\pi^+\pi^-K$ decay is found, we can derive the $B \to X(3872)K$ yield expected for the data from the signal efficiency and the branching fraction measured using the Belle full data [31],

$$\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-) = (8.63 \pm 0.97) \times 10^{-6},$$
$$\mathcal{B}(B \to X(3872)K) \times \mathcal{B}(X(3872) \to J/\psi\pi^+\pi^-) = (4.3 \pm 1.3) \times 10^{-6}.$$

The signal efficiency is 22.9% and 17.5% from B^+ and B^0 decay using the signal MC sample, respectively. The expected signal yield is 19.8 for 62.8 fb^{-1}, and its statistical significance is expected to exceed 4.3σ.

For data, the observed $M(J/\psi\pi^+\pi^-)$ distributions on the $X(3872)$ signal region are shown in Fig. 7.4. To extract the $X(3872)$ signal, an Unbinned Simultaneous Extended Maximum-likelihood fit is performed. For signal contributions, a histogram PDF generated with the signal MC sample assuming the world average mass of $X(3872)$ [18] and the Breit-Wigner width of the LHCb measurement [72] is used. A first-order Chebyshev function is used as a PDF of the background contribution. To combine B^0 mode, in the simultaneous fit, the ratio of the number of the signals between the B^0 and B^+ mode is fixed by the ratio of the expected signal yields per 1 fb^{-1} with assuming $\mathcal{B}(B^0 \to X(3872)K^0)/\mathcal{B}(B^+ \to X(3872)K^+) = 0.50$ [31].

The fitted signal yield is 14.4 ± 4.6, where uncertainty is statistical only. The statistical significance is determined from the log-likelihood ratio $-2\ln(\mathcal{L}_0/\mathcal{L})$, where $\mathcal{L}_0$ and $\mathcal{L}$ are the fit likelihood without and with the signal component, respectively. Taking into account of the number of degrees of freedom, the significance is determined to be 4.6σ. Furthermore, the possibility that the peak is caused by background coming from B meson decay is verified by examining the $M(J/\psi\pi^+\pi^-)$ distributions on $M_{\rm bc}$ and ΔE sidebands. No peaking background on the $X(3872)$ signal region for all the sidebands is found. Therefore, we conclude that we get the evidence of $X(3872)$ signal at Belle II.

7.2.2 Belle II Distributed Computing System

At the Belle II experiment, massive collision data is collected to broadly advance our understanding of particle physics. At the end of the data taking, an order of 100 PB disk

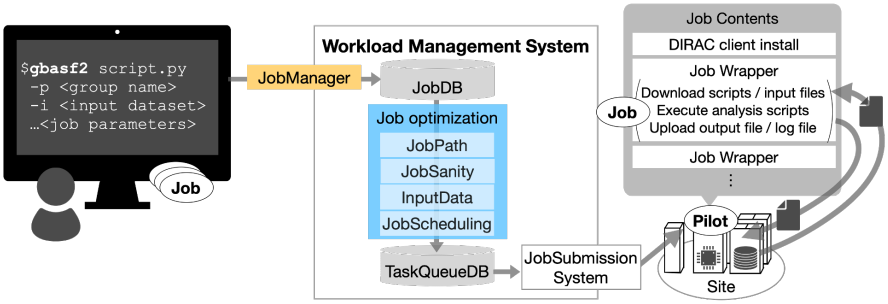


Figure 7.5: Workflow of usual analysis job submission on the DC system. In the gbasf2 command, script.py shows the analysis script. The blue box shows an executor, and the job optimization is performed using four executors. The orange box and cylinder show a service and a database (DB), respectively.

storage and tens of thousands of CPU cores are required [82]. To store and process the massive data with the resources, a worldwide distributed computing (DC) system is utilized. This system is also used for physics analysis. To perform physics analysis efficiently with this system, it is important that jobs are processed quickly without interfering with the analysis activities. A job that does not run properly due to issues, hereafter a "failed job", is one of the causes of the efficiency reduction. To suppress failed jobs, a syntax checker and a scout job framework were developed. This subsection describes overviews of the Belle II DC system and physics analysis, the two developed features, and their total performance.

Belle II DC System and Physics Analysis

The Belle II DC system consists of several pieces of software [82, 83]. The core one is DIRAC interware [84]. It provides a complete grid solution that interconnects end-users and heterogeneous computing resources. It includes a Workload Management System (WMS), a Data Management System, support of analysis job execution, and so on. Its extension, BelleDIRAC, has been developed to meet the requirements of the experiment, e.g., an automatic system that generates jobs for producing simulation samples and processing raw data [85].

Processes for all aspects of the data-processing chain can be performed by the Belle II Analysis Software Framework (basf2) [86]. To analyze data distributed over computing sites, a client tool to support basf2 job execution on the DC system (gbasf2) and a set of client tools to manage submitted jobs and output data are provided. End-users can submit a group of basf2 jobs by a single command specifying input data sets and the other job parameters. In detail, the jobs are stored in JobDB when the command is executed, as shown in Fig. 7.5. The parameters are quickly analyzed by several tasks called executors, e.g., site assignment based on input data, etc. What is done by all of the executors is henceforth referred to as job optimization. The jobs are registered in TaskQueueDB by the last executor. Gradually, WMS submits pilot jobs to the computing elements (CE) of the sites where input data is hosted. When the pilot job is submitted to the worker node and executed, the end-user job is pulled, input data is downloaded from the storage element (SE), and the basf2 process is executed. This workflow indicates that even if there is a problem with an end-user job, the job spends a few minutes on the worker node to authenticate and download input data.

The failed jobs have potential to prevent efficient analysis in two points. First, worker

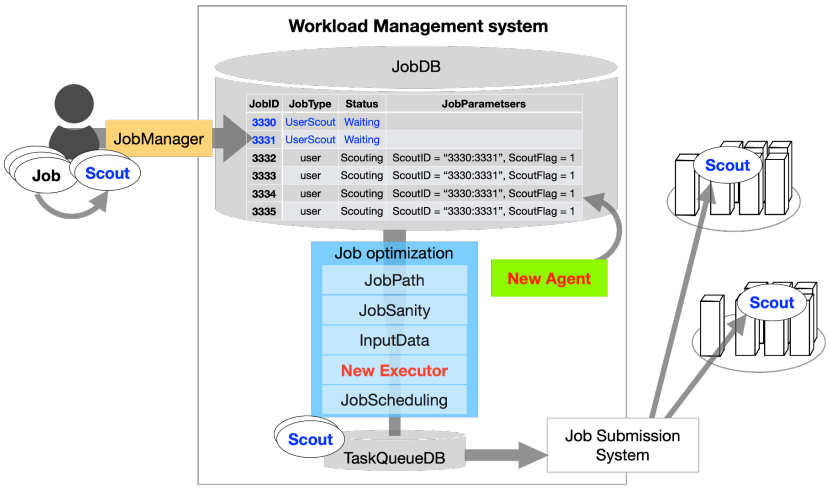


Figure 7.6: Design and workflow of the scout job framework

nodes are unnecessarily occupied for at least a few minutes due to failed jobs. Second, when many jobs are submitted at once and they fail quickly, access to the central system, CE, and SE is concentrated for a short time. It often triggers system trouble and reduces system uptime. In addition, solving the trouble becomes a load on the operation side. In 2019, 9.6 million analysis jobs were executed, but about ten percent of them failed. The main reason was not problems in the DC system or computing resources but basf2 or gbasf2 termination, which is caused by syntax errors, improper job parameter settings specified by the end-user, failure to upload output files, etc. At Belle II, therefore, python syntax checker and scout job framework were introduced as countermeasures against such failed jobs.

Developments to Suppress Failed Jobs

We have developed two features to suppress the failed jobs.

First, the python syntax checker is added to gbasf2. It detects syntax errors at the language level of analysis scripts and prevents jobs with simple errors from storing the job in the system.

Next, since the syntax checker is not enough to detect complicated syntax errors or the improper settings of job parameters, a scout job framework was developed. The targets of the framework are job groups with a large number of jobs. Its concept is that original jobs and a small number of test jobs (henceforth referred to as "main jobs" and "scout jobs", respectively) are stored in JobDB at the same time, and the main jobs are submitted to sites only when scout jobs are successful. To achieve this functionality, The following three components were developed.

- A mechanism to generate scout jobs in the client tool.
- A new executor to temporarily hold back the main jobs from registering in TaskQueueDB.
- A new agent, which is to perform actions periodically, to monitor the status of scout jobs and to take action to resume the remaining job optimization of main jobs only when the scouting is judged to be successful.

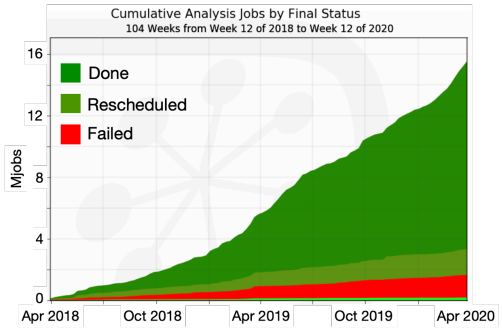
After introducing the framework, the workflow of the job execution is shown in Fig. 7.6. Details of their design are described in Ref. [87].

Performance

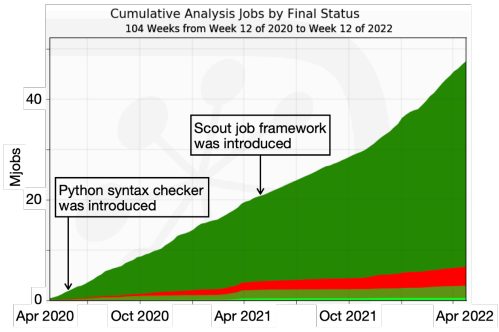
The python syntax checker and the scout job framework were introduced to the production environment in early May 2020 and late April 2021, respectively. Total performance was evaluated by comparing the percentage of failed jobs between two periods: two years before both introductions, Apr. 2018 – Mar. 2020, and those after the syntax checker was introduced first, Apr. 2020 – Mar. 2022. Fig. 7.7 shows the cumulative number of analysis jobs by job status in the two periods. Comparing these two plots, the total number of executed jobs in the latter period is three times larger than in the former period, because analysis activity increased as data accumulated. The percentage of failed jobs is 9.5% and 7.9% in the former and the latter period, respectively. That results in a relative improvement of 17%. Under a more severe environment in which there are more people and opportunities to use the system, the situation could be improved.

For evaluation of the performance of the scout job framework, 245 thousand job groups executed in the period from April 2021 to March 2022 are analyzed. Of these job groups, the 45.2 thousand groups were the target of this framework. 6.5% of them, 2930 groups, were deemed problematic, and stopped before the submission. If such problematic jobs had been submitted, the number of failed jobs would have increased by more than 12% relative to the current results. The result concludes that this framework works well enough.

As a result, the percentage of failed jobs was reduced compared to before these features were implemented. It enables us to prevent failed jobs from causing system problems and waste of computing resources.



(a) Results for two years before both features were introduced.



(b) Results for two years after the syntax checker was introduced.

Figure 7.7: Cumulative number of analysis jobs by job status as a function of time and pie charts of the job failure reasons.

Chapter 8

Conclusion

In this thesis, the $X(3872)$ lineshape in the $D^0\overline{D}^{*0}$ decays is examined using exclusive decays $B \to D^0\overline{D}^{*0}K$ in a data sample of 772×10^6 $B\overline{B}$ pairs collected with the Belle detector at the KEKB asymmetric-energy e^+e^- collider.

In this work, we evaluate the lineshape with two models. The first is the relativistic Breit-Wigner lineshape, commonly used for resonance states. Through the evaluation with this model, we ensure the analysis method and examine the tendency to yield the higher mass, the larger width, and the higher relative branching fraction between $B^0 \to X(3872)K^0$ and $B^+ \to X(3872)K^+$ in the $X(3872) \to D^0\overline{D}^{*0}$ decay mode only. The second is the Flatte model, in which the Breit-Wigner model is extended to account for the coupled-channel effect. It can describe not only resonance states but also a bound state and a threshold cusp; it explains the lineshape discrepancy between the decays to the $J/\psi\pi^+\pi^-$ and the $D^0\overline{D}^{*0}$ final states. However, this model has the scaling behavior that the lineshape does not change in the vicinity of the threshold when all parameters are simultaneously scaled. Because of this behavior and the too-narrow width of the lineshape compared with the mass resolution, the analysis using $J/\psi\pi^+\pi^-$ decay was performed at LHCb, however, all parameters could not be determined. The analysis in the $D^0\overline{D}^{*0}$ decay is expected to be useful in pinning down the scaling behavior. Therefore, the issue of each lineshape model regarding the analysis of the $D^0\overline{D}^{*0}$ decays is investigated.

In this study, we developed an analysis method that is independent of the lineshape model so that it can be applied to other lineshape models. In addition, D^0 decay modes used in the reconstruction are added to improve statistical sensitivity. The lineshape dependence of the fit mode for incorrectly reconstructed signal events is additionally considered to reduce systematic uncertainty.

As a result of the data analysis, the mass, width and branching fractions of the relativistic Breit-Wigner lineshape are determined to be

$$m = 3873.71^{+0.56}_{-0.50}(\text{stat}) \pm 0.13(\text{syst})\ \text{MeV}/c^2,$$

$$\Gamma_0 = 5.2^{+2.2}_{-1.5}(\text{stat}) \pm 0.4(\text{syst})\ \text{MeV},$$

$$\mathcal{B}(B^+ \to X(3872)K^+) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}) = (0.97^{+0.21}_{-0.18}(\text{stat}) \pm 0.10(\text{syst})) \times 10^{-4},$$

$$\mathcal{B}(B^0 \to X(3872)K^0) \times \mathcal{B}(X(3872) \to D^0\overline{D}^{*0}) = (1.30^{+0.36}_{-0.31}(\text{stat})^{+0.12}_{-0.07}(\text{syst})) \times 10^{-4},$$

$$\frac{\mathcal{B}(B^0 \to X(3872)K^0)}{\mathcal{B}(B^+ \to X(3872)K^+)} = 1.34^{+0.47}_{-0.40}(\text{stat})^{+0.10}_{-0.12}(\text{syst}).$$

The signal from B^0 decays is firstly observed with more than 5σ significance. These results are consistent with the previous study using the $D^0\overline{D}^{*0}$ decays, and our measurement

precision is relatively improved by at least 22%. The measured Breit-Wigner mass is significantly higher than the $D^0\overline{D}^{*0}$ threshold, while the world-average mass with the $J/\psi\pi^+\pi^-$ decay is consistent with the threshold. The measured width and $\mathcal{B}(B^0 \to X(3872)K^0)/\mathcal{B}(B^+ \to X(3872)K^+)$ are shifted from the average with the $J/\psi\pi^+\pi^-$ decay by 2.6σ and 2.0σ, respectively [45]. This study supports the trend that these properties in the $X(3872) \to D^0\overline{D}^{*0}$ decay mode differ from those in the $J/\psi\pi^+\pi^-$ decay mode, as shown by the previous studies.

For the Flatté lineshape, we established the method to measure the coupling constant to the $D\overline{D}^*$ channel g as the undetermined parameter based on the LHCb result using the $J/\psi\pi^+\pi^-$ decay. As a result, the fitted g is in the region that is insensitive to exact values due to the limited size of our data sample. The lower limit is determined to be

$$g > 0.075 \text{ at } 95\% \text{ credibility.}$$

It is a more stringent lower limit than the previous measurement at LHCb. This suggests that analysis using $D^0\overline{D}^{*0}$ can indeed complement the study of the $J/\psi\pi^+\pi^-$ mode in this framework.

As a prospect, these measurements are currently dominated by the uncertainty due to the limited size of the data sample, and they will be improved by applying this analysis at the successor experiment, Belle II. Furthermore, for the analysis with the Flatté lineshape, the sensitivity can be improved more by performing a simultaneous fit between samples of $B \to (X(3872) \to D^0\overline{D}^{*0})K$ at Belle II and $B^+ \to (X(3872) \to J/\psi\pi^+\pi^-)K^+$ at LHCb. Such an analysis could fully determine the lineshape in the coupled-channel framework, and greatly contribute to determining the internal structure.

Acknowledgments

Firstly, this research could never have been completed without much help from many people. I would like to express my sincere gratitude to all of them.

I would like to express my deepest appreciation to my supervisor, Prof. Toru Iijima, for giving me a great opportunity to be involved in this research and for encouraging me to accomplish this research from a broad perspective. Thanks to this opportunity, I have gained a wide range of experience from physics to data analysis and computing.

I am extremely grateful to Dr. Yuji Kato for his tremendous support in the technical aspects of data analysis and computing. He gave me basic knowledge and a lot of practical suggestions for this research. Without his support, this research would never have been completed.

I would like to extend my sincere thanks to Dr. Kiyoshi Tanida for the many consultations related to the physical motivation of this analysis and the hadron physics analysis. With his help, I developed this analysis.

I gratefully acknowledge all of the discussions with all members of N-lab. The members of the Belle and Belle II group at N-lab, Dr. Kenji Inami, Dr. Kodai Matsuoka, Dr. Keisuke Yoshihara, and Dr. Qidong Zhou, continuously support my research. Especially for my main analysis, computational resources at N-lab were essential. We would like to thank Dr. Kenji Inami, Mr. Yuki Sue, and the staffs of Fujitsu Japan Limited, who manage them. When I gave the public presentations, Prof. Makoto Tomoto, Dr. Yasuyuki Horii, and the other members of our laboratory gave me much useful advice. Ms. Mieko Miyake, our secretary, always helped me with various procedures for my research activities.

I would like to thank the members of the Belle II Japanese hadron group, Prof. Kenkichi Miyabayashi, Prof. Tatsuro Matsuda, Dr. Makoto Takizawa, Dr. Mizuki Sumihama, Dr. Masayuki Niiyama, and Dr. Yue Ma, for their constructive suggestions based on their diverse experience in the analysis of the hadron physics at Belle.

In preparing a journal publication, many Belle collaborators supported me. For example, the internal referees, Dr. Pasha Pakhlov, Dr. Bruce Yabsley, and Dr. Klemens Lautenbach, gave us precise advice and helped us to finalize our analysis. Prof. Peter Krizan helped us with correcting our English writing. Prof. Yoshihide Sakai and Dr. Mikihiko Nakao gave me valuable comments and suggestions for the publication.

For my research of the Belle II DC system, the great support of Prof. Takanori Hara, Dr. Ikuo Ueda, Dr. Hideki Miyake, and Dr. Michel Hernandez Villanueva was essential. Because I was an absolute beginner in computing, it must not have been easy to teach me. Thanks to all of their supports, I was able to take the first step toward the challenge of software development and accomplish it.

Since 2022, I have been working in the research group for Hadron Nuclear Physics at JAEA. I am very grateful to Dr. Hiroyuki Sako, Ms. Ai Mashiko, and the other members of the group for all their help and support. Thanks to them, I was able to lead a fulfilling research life.

This work is supported by Grant-in-Aid for JSPS Fellows (No. 19J23314).

Appendix A

Optimization of Event Selection

This appendix describes the optimization of the signal event selection. In optimizing the selection criteria, we require a loose signal selection for the candidates at the reconstruction. Hereafter, this is referred to as preselection. In this chapter, we describe the preselection and the selection optimization.

A.1 Preselection

This section describes the preselection of each particle candidates.

Tracks are required to have a transverse distance between the point of closest approach and the interaction point (dr) of less than 1 cm, and a distance along the beam axis between them (dz) of less than 4 cm. Since the signal events do not include any electrons, well-identified electrons are eliminated by requiring that electron-ID is smaller than 0.95. Among the remaining charged track candidates, those with $\mathcal{L}_\pi/(\mathcal{L}_\pi + \mathcal{L}_K) > 0.1$ are considered as pion candidates, and those with $\mathcal{L}_K/(\mathcal{L}_\pi + \mathcal{L}_K) > 0.1$ are used as kaon candidates. For the kaon candidates from the B^+ decay, a tighter requirement of kaon identification, $\mathcal{L}_K/(\mathcal{L}_\pi+\mathcal{L}_K) > 0.6$, is used to reduce the multiplicity of the B^+ candidates.

Neutral pion candidates are required that both photons satisfies $E9E25 > 0.8$, and an energy $E_{\rm Lab} > 30$ MeV for the barrel region or $E_{\rm Lab} > 50$ MeV for the endcap region. Only those with the $\gamma\gamma$ invariant mass of 0.1124 GeV/c^2 $< M(\gamma\gamma) <$ 0.1575 GeV/c^2 is retained for the later analysis. In order to improve the mass resolution of the $D^0\overline{D}^{*0}$ invariant mass, a mass-constrained fit is applied.

Candidates of K^0_S mesons are seleccted by requireing the $\pi^+\pi^-$ invariant mass to be consistent with K^0_S nominal mass [18] within 15 MeV/c^2. A mass- and vertex-constrained fit are applied.

For D^0 candidates, the reconstructed invariant mass is required to be within 39 MeV/c^2 of the nominal mass [18] if a π^0 is included; otherwise, it is required to be within 19 MeV/c^2. A mass- and vertex-constrained fit is applied. For photon candidates for D^{*0} reconstruction, γ candidates with $E9E25 > 0.8$ and $E_{\rm lab} > 70$ MeV are required. To reduce the number of candidates, we retain the candidates with $M(D^0\overline{D}^{*0})$ of less than 4.0 GeV/c^2 only.

The beam-energy constraint mass, $M_{\rm bc}$, and the difference of the energy in the center-of-mass system between the beam and the B candidate, ΔE, are used to identify the signal B candidates. The events which satisfy $M_{\rm bc} > 5.2$ GeV/c^2 and $|\Delta E| < 50.0$ MeV are selected. Events with more than 10000 candidates of D^0, D^{*0}, $X(3872)$ or B in the reconstruction are discarded.

A.2 Optimization of the Selection

Selection criteria are optimized so that Figure-Of-Merit (FOM) is improved while the signal efficiency keeps as high as possible. There are a total of ten variables to be optimized. We divided them into two groups: variables with high separation power and the others. For the former, the selection criteria are determined to be the point in the FOM plateau region. For the latter, the criteria are set so that the FOM increases slightly, but the efficiency is greater than 90%.

- Variables with high separation power:
  - $M(\pi^+\pi^-)$ for the K^0_S selection,
  - $M(\gamma\gamma)$ for the π^0 selection,
  - Reconstructed D^0 mass for the D^0 selection,
  - Difference of the reconstructed mass between D^{*0} and D^0 for the D^{*0} selection,
  - ΔE for the B selection,
  - $M_{\rm bc}$ for the B selection.
- The others:
  - γ energy in the laboratory system for the selection of γ from D^{*0},
  - π^0 energy in the laboratory system ($E_{\rm lab}$) for the π^0 selection,
  - π^0 momentum in the center-of-mass system ($P_{\rm cms}$) for the π^0 selection,
  - χ^2 probability of the vertex- and mass-constraint fit on the D^0 candidates.

In the following, the results of the selection optimization for each intermediate state are described.

Neutral kaons

Neutral kaons in the signal event have two sources: D^0 decays and B^0 decays. The selection criterion for each source is optimized. Figure A.1 shows the FOM as a function of a criterion of $M(\pi^+\pi^-)$. Regardless of the source, we can maximize FOMs by requiring $M(\pi^+\pi^-)$ to be consistent with the nominal mass within 7.0 MeV/c^2.

photons

For γ from D^{*0} decays, $E_{\rm lab}$ is optimized with requirements of the D^0, D^{*0}, and B selections described in the later sub-subsections. The reason is that the FOM is efficiently improved compared to the γ optimization performed before the optimization of the D^0, D^{*0}, and B selection. The distribution and FOM as a function of the criterion are shown in Fig. A.2. FOM is slightly improved by the requirement of $E_{\rm lab} > 90$ MeV, so we use the requirement.

Neutral pions

In this analysis, π^0's in the signal event have two sources, one is from D^0 decay, and the second one is from D^{*0} decays. The selection criteria of these variables are optimized separately since the momentum and energy are different for each source.

Figures A.3–A.4 show the $E_{\rm lab}$ and $P_{\rm cms}$ distributions for each source and the FOM as a function of the criterion. We use criteria of $E_{\rm lab}$ and $P_{\rm cms}$, which can increase the FOM without decreasing the signal efficiency: $E_{\rm lab} > 0.15$ GeV and $P_{\rm cms} > 0.10$ GeV/c for the π^0 from D^0 decay, and $E_{\rm lab} < 0.20$ GeV and $P_{\rm cms} < 0.10$ GeV/c for the π^0 from D^{*0} decay. The FOM as a function of the $M(\gamma\gamma)$ criterion with the other optimized selections is shown in Fig. A.5. Regardless of the source, the $\gamma\gamma$ invariant mass is required to be consistent with a nominal π^0 mass within 12 MeV/c^2.

Candidates of D^0 mesons

In case the D^0 candidate is reconstructed without π^0, the mass resolution is better and the background level is smaller than those with π^0. Therefore, the criteria are optimized by dividing the sample into two parts. Figure A.6 shows the FOM as a function of the criterion of the reconstructed D^0 mass after the other optimized selections are required. To maximize FOM, the invariant mass is required to be consistent with a nominal D^0 mass within 8.5 MeV/c^2 for candidates without π^0 and 16 MeV/c^2 for candidates with π^0. Figure A.7 shows the FOM as a function of the criterion of the χ^2 probability of the constraint fit. It is required to be greater than 0.0001, regardless of the reconstruction mode.

Candidates of D^{*0} mesons

Since the mass resolution changes depending on the reconstruction D^{*0} mode, the selection is separately optimized. The FOM as a function of the criterion with the other optimized selections is shown in Fig. A.8. The mass difference between D^{*0} and D^0 is required to be within 9.0 MeV/c^2 of the nominal value for $D^{*0} \to D^0\gamma$ and 2.0 MeV/c^2 of the nominal value for $D^{*0} \to D^0\pi^0$.

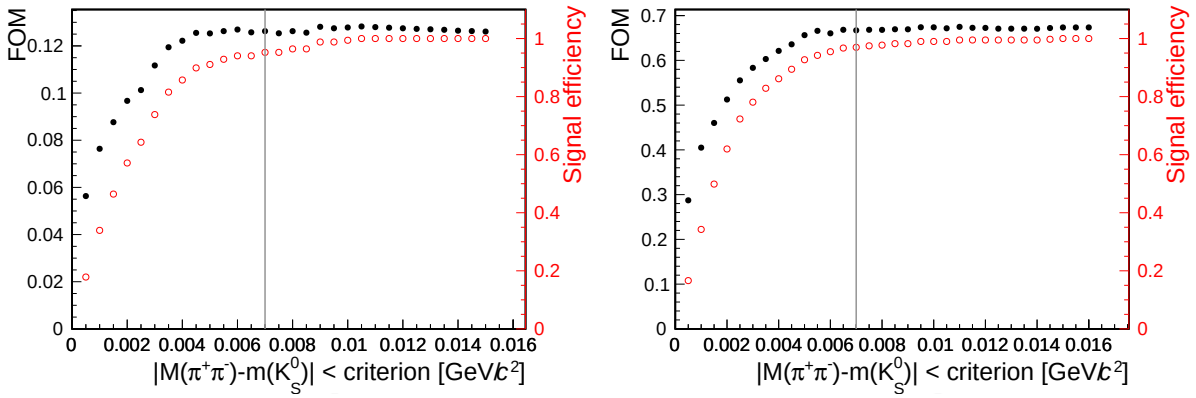
Candidates of B mesons

The criteria of $M_{\rm bc}$ and ΔE are optimized for each of the D^{*0} and B modes. The FOMs as a function of the criterion are shown in Fig. A.9–A.10. For ΔE, FOMs are maximized when $|\Delta E| < 12$ MeV is required, regardless of the B and D^{*0} decay modes. For $M_{\rm bc}$, regardless of the B decay mode, FOMs are maximized when $M_{\rm bc}$ is required to agree with the nominal B mass [18] within 6.0 MeV/c^2 for $D^{*0} \to D^0\pi^0$ and 4.5 MeV/c^2 for $D^{*0} \to D^0\gamma$.

Continuum Suppression

For suppression of continuum events, we use a FastBDT classifier [67]; Details are described in Appendix B. A requirement of the FastBDT output is optimized by FOM, as shown in Fig. A.11. In what follows, the FastBDT output is required to be greater than 0.15, regardless of whether it is the B^+ mode or the B^0 mode. Also, we checked if the sensitivity could be increased by binning with that output as in Ref. [88] because the

classification power is not very strong. The improvement of the FOM was only a few percent; therefore, we use it for the cut rather than binning.



(a) K^0_S candidates from the D^0 decays (b) K^0_S candidates from the B decays

Figure A.1: FOM (closed black circles) and the signal efficiency (opened red circles) as a function of the criterion of $M(\pi^+\pi^-)$. The vertical gray line shows the optimized criterion.

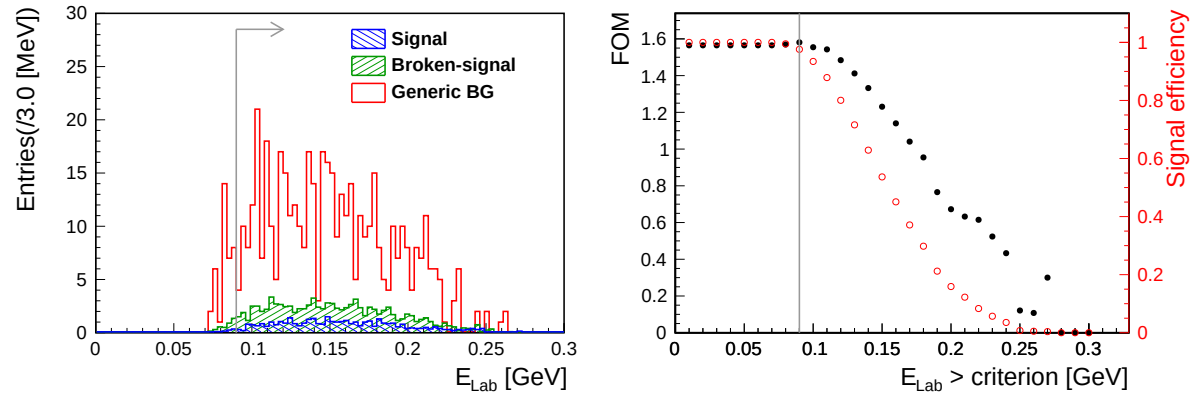
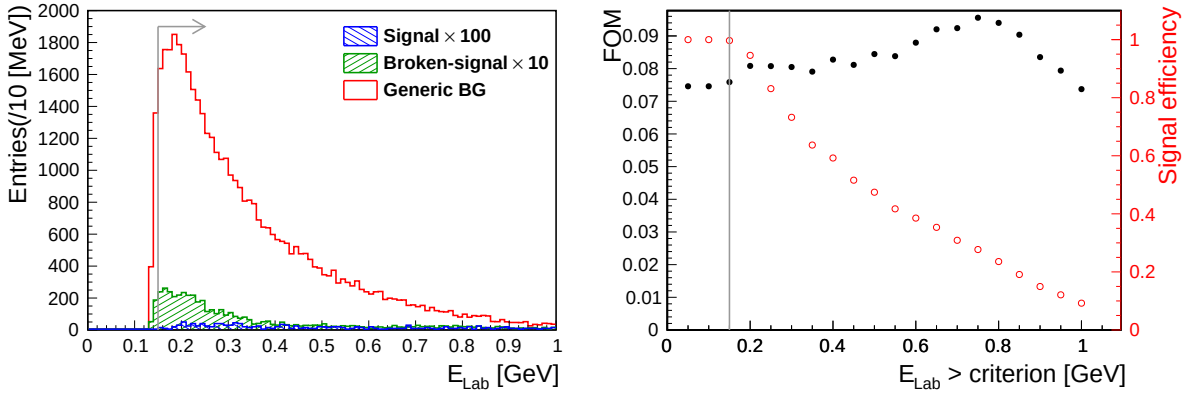
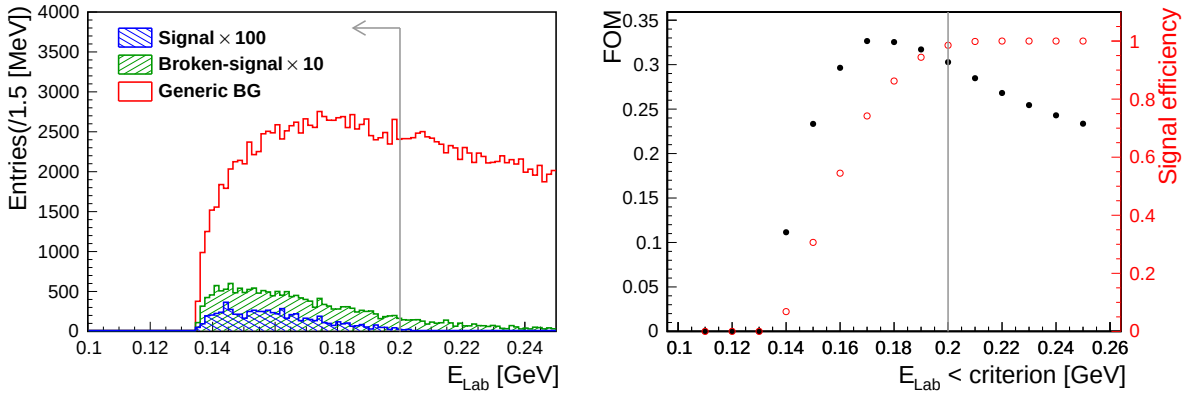


Figure A.2: (Left) The distribution of $E_{\rm lab}$ for γ from D^{*0}. The blue, green, and red histograms show distributions for signal, broken-signal, and generic background, respectively. (Right) FOM (closed black circles) and the signal efficiency (opened red circles) as a function of the criterion of $E_{\rm lab}$. In each panel, the vertical gray line shows the optimized criterion.

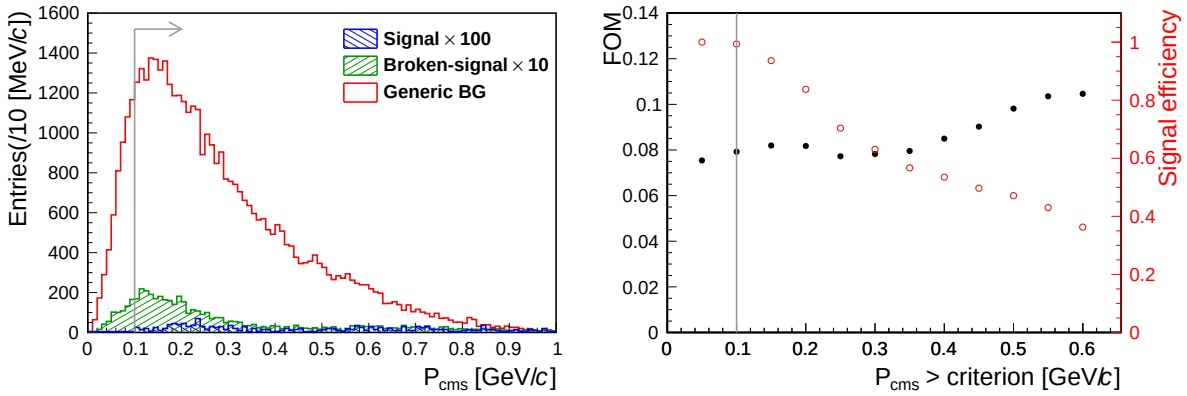


(a) π^0 candidates from the D^0 decays

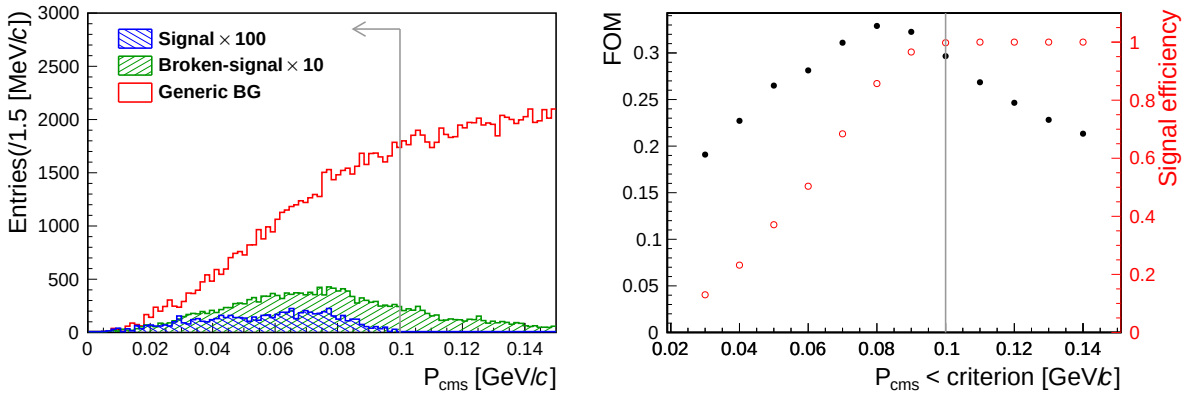


(b) π^0 candidates from the D^{*0} decays

Figure A.3:(Left) The distribution of $E_{\rm lab}$ of the π^0 candidates. The blue, green, and red histograms show distributions for signal, broken-signal, and generic background, respectively. (Right) FOM (closed black circles) and the signal efficiency (opened red circles) as a function of the criterion of $E_{\rm lab}$. In each panel, the vertical gray line shows the optimized criterion.

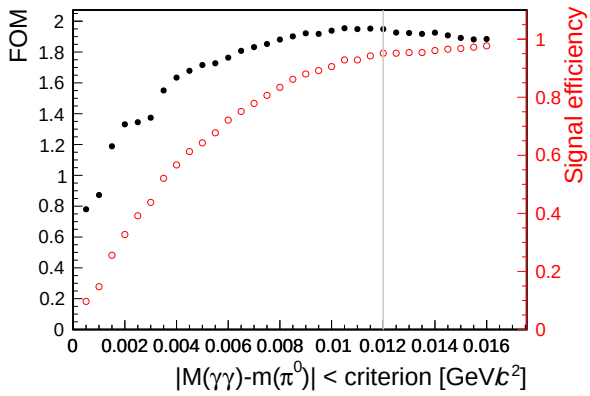


(a) π^0 candidates from the D^0 decays

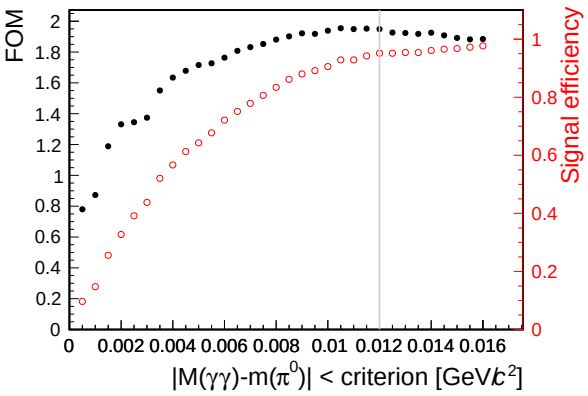


(b) π^0 candidates from the D^{*0} decays

Figure A.4: (Left) The distribution of $P_{\rm cms}$ for the π^0 candidates. The blue, green, and red histograms show distributions for signal, broken-signal, and generic background, respectively. (Right) FOM (closed black circles) and the signal efficiency (opened red circles) as a function of the $P_{\rm cms}$ criterion. In each panel, the vertical gray line shows the optimized criterion.

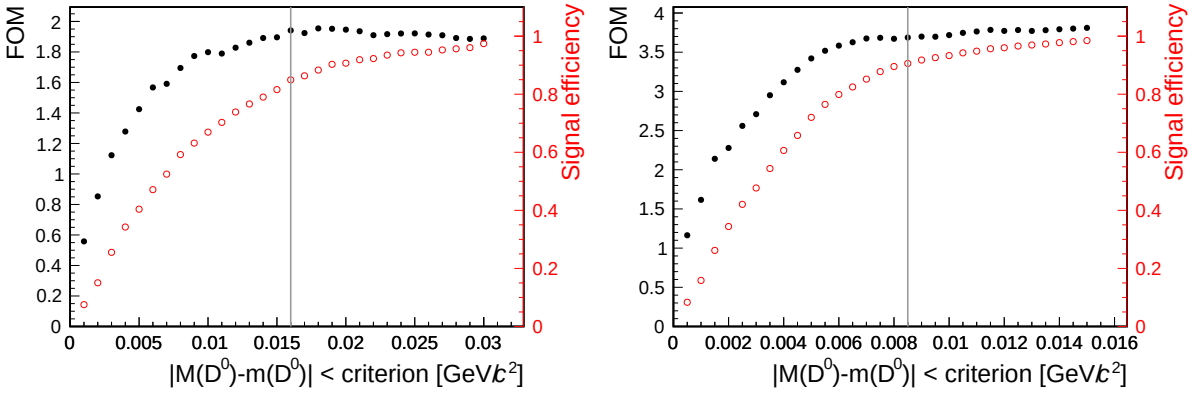


(a) π^0 candidates from the D^0 decays



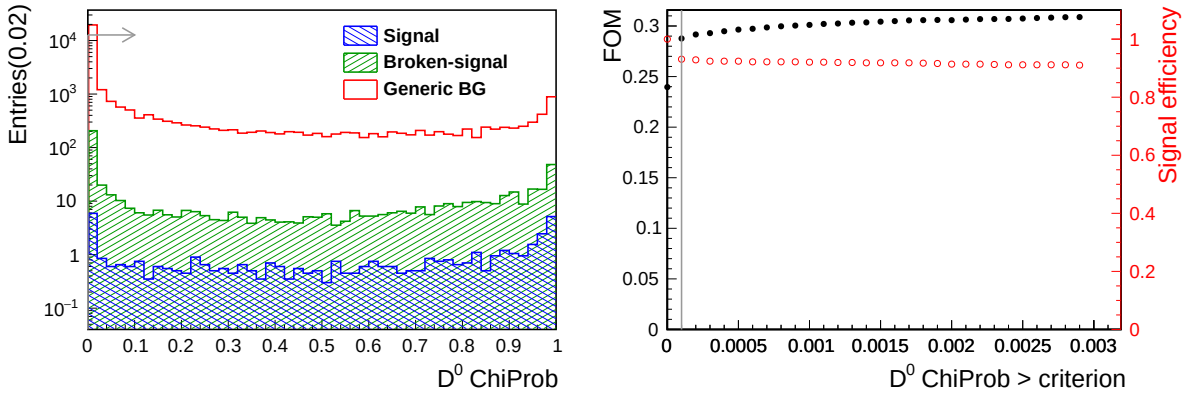
(b) π^0 candidates from the D^{*0} decays

Figure A.5: FOM (closed black circles) and the signal efficiency (opened red circles) as a function of the criterion of the $\gamma\gamma$ invariant mass with the optimized selections other than $M(\gamma\gamma)$. In each panel, the vertical gray line shows the optimized criterion.

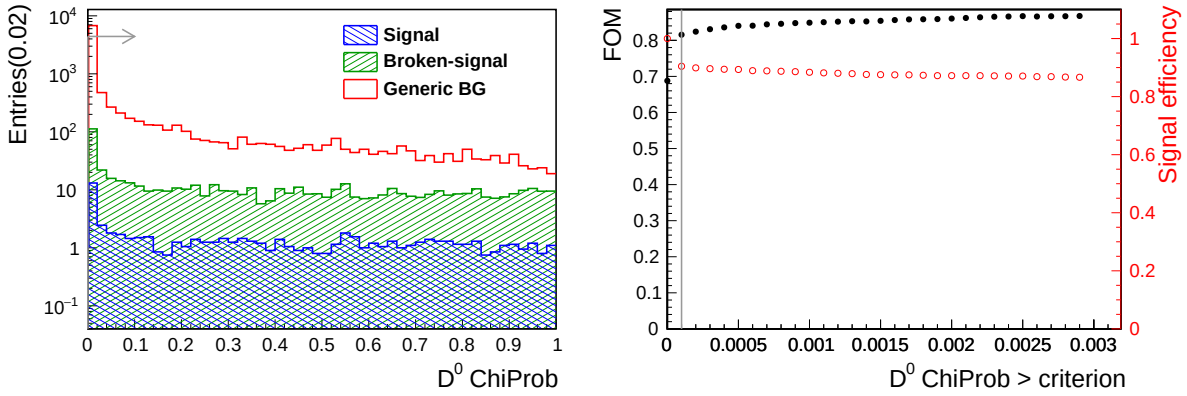


(a) D^0 candidates reconstructed with π^0 (b) D^0 candidates reconstructed without π^0

Figure A.6: FOM (closed black circles) and the signal efficiency (opened red circles) as a function of the criterion of the reconstructed D^0 mass with the optimized selection other than the reconstructed D^0 mass. In each panel, the vertical gray line shows the optimized criterion.

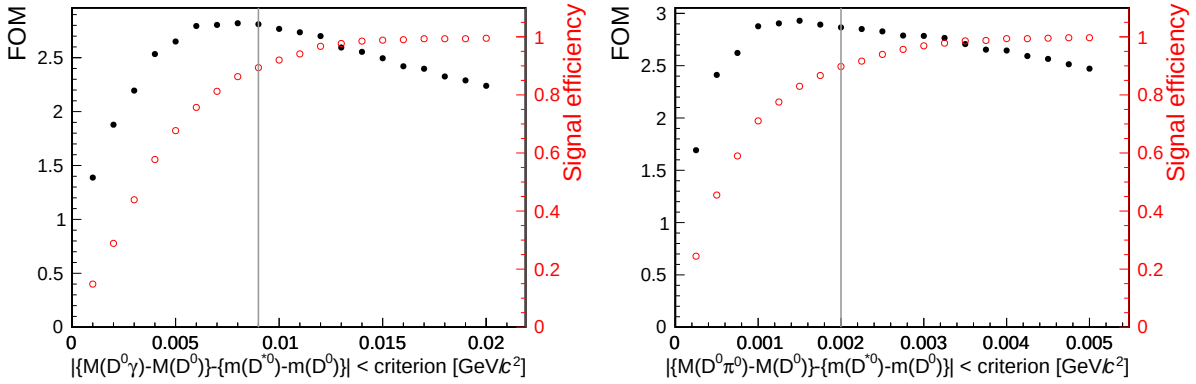


(a) D^0 candidates reconstructed with π^0



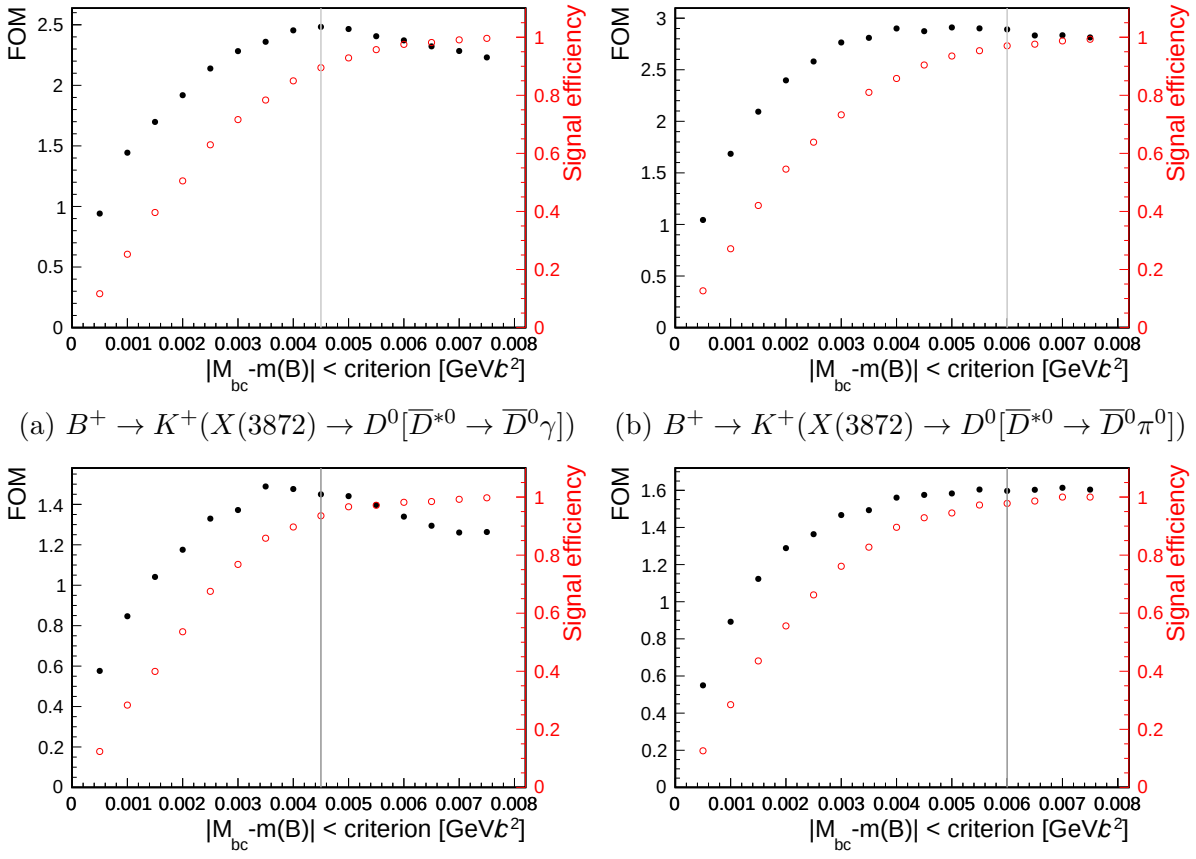
(b) D^0 candidates reconstructed without π^0

Figure A.7: (Left) The distribution of the χ^2 probability of the vertex- and mass-constraint fit on the D^0 candidates. The blue, green, and red histograms show distributions for signal, broken-signal, and generic background, respectively. (Right) FOM (closed black circles) and the signal efficiency (opened red circles) as a function of the criterion of the χ^2 probability.



(a) Decay mode of $D^{*0} \to D^0\gamma$ (b) Decay mode of $D^{*0} \to D^0\pi^0$

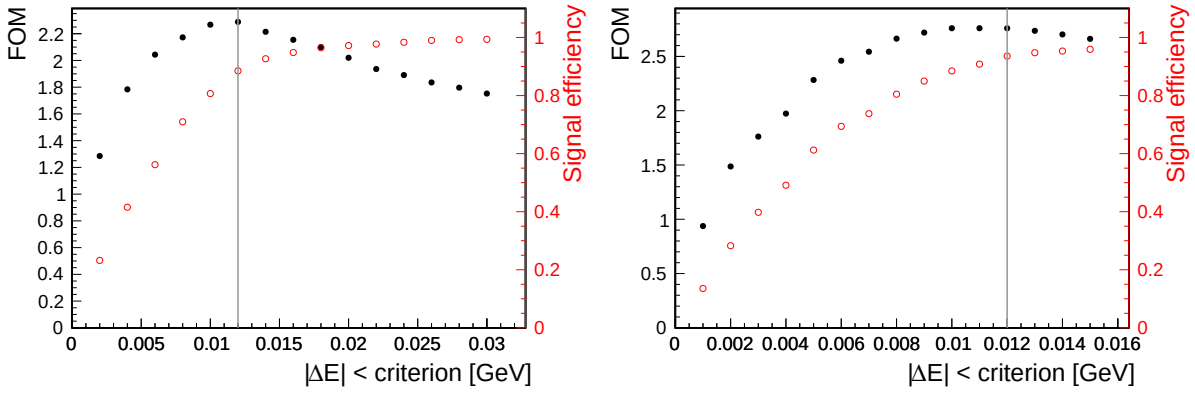
Figure A.8: FOM (closed black circles) and the signal efficiency (opened red circles) as a function of the criterion of the difference of the reconstructed mass between $\overline{D}^{*0}$ and $\overline{D}^0$ with the optimized selection other than the mass difference. In each panel, the vertical gray line shows the optimized criterion.



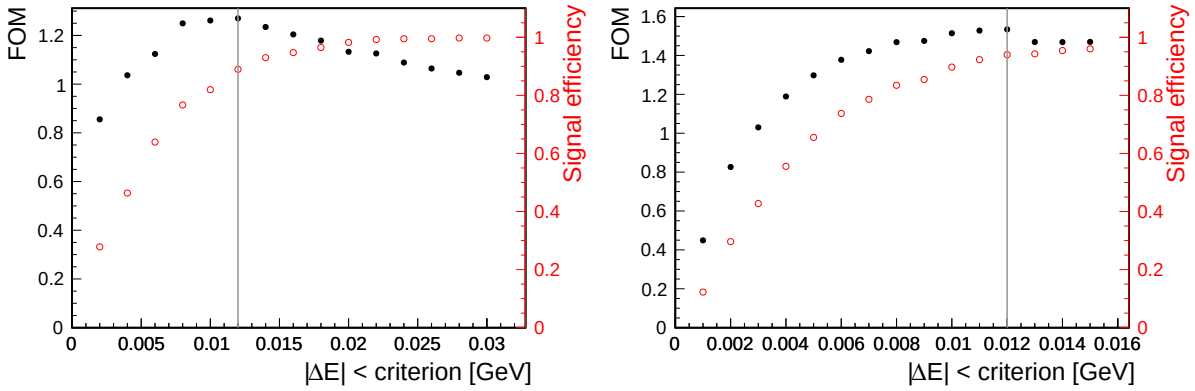
(a) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (b) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

(c) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (d) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

Figure A.9: FOM (closed black circles) and the signal efficiency (opened red circles) as a function of the criterion of $M_{\rm bc}$. In each panel, the vertical gray line shows the optimized criterion.

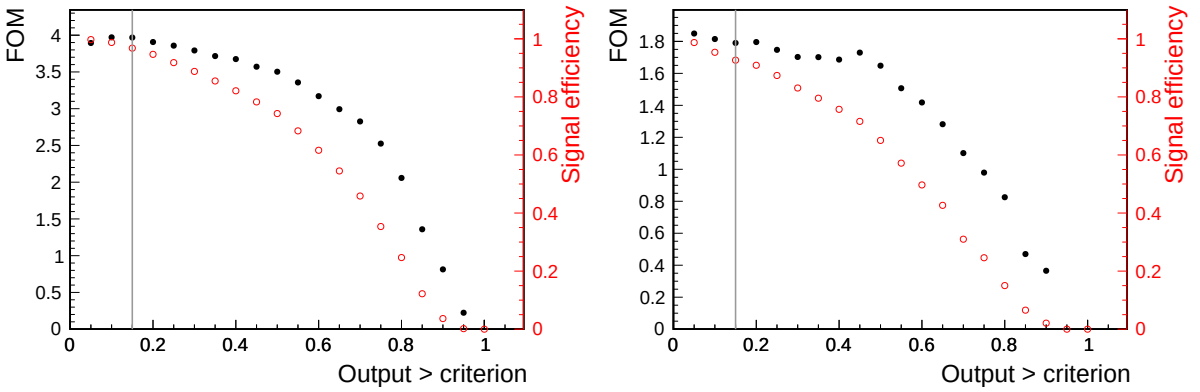


(a) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (b) $B^+ \to K^+(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$



(c) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma])$ (d) $B^0 \to K^0(X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0])$

Figure A.10: FOM (closed black circles) and the signal efficiency (opened red circles) as a function of the criterion of ΔE. In each panel, the vertical gray line shows the optimized criterion.



(a) The B^+ decay mode (b) The B^0 decay mode

Figure A.11: FOM (closed black circles) and the signal efficiency (opened red circles) as a function of the FastBDT output criterion. In each panel, the vertical gray line shows the optimized criterion.

Appendix B

Multivariate Analysis for Continuum Suppression

About 30% of the generic background comes from the continuum event. For the continuum suppression, the FastBDT classifier is used. It is trained on MC samples containing 1.00×10^5 signal events and 3.31×10^6 continuum events, which are independent of any other analysis samples. Thirty variables are used as input: KSFW variables (et, mm2, hso00, hso04, hso10, hso14, hso20, hso22, hso24, hoo0, hoo1, hoo2, hoo3, hoo4) [89], CleoConeCS (1st–9th) [90], magnitude of the signal B thrust axis (thrustBm), magnitude of the ROE thrust axis (thrustOm), cosine of the angle between thrust axis of the signal B and thrust axis of ROE (cosTBTO), cosine of the angle between thrust axis of the signal B and z-axis (cosTBz), and reduced Fox-Wolfram moment $R2$ [81]. Figures B.1–B.2 show the distributions of the input variables. For each distribution, there is a difference in the distribution of the signal events and the continuum events. $R2$ has the largest difference between them, and it is the most important variable of the classification.

To check for overlearning, a test sample is used to check the performance, e.g. output distribution and the receiver operating characteristic (ROC) curves, as shown in Figs. B.3 and B.4. Here, the test sample is a part of the analysis sample and contains 2.00×10^5 signal events and 2.05×10^6 continuum events. Compared with those performances for the training sample, the performances agree with each other's samples. It means that the classifier is not overtrained. Furthermore, the performance is better than that of $R2$ only.

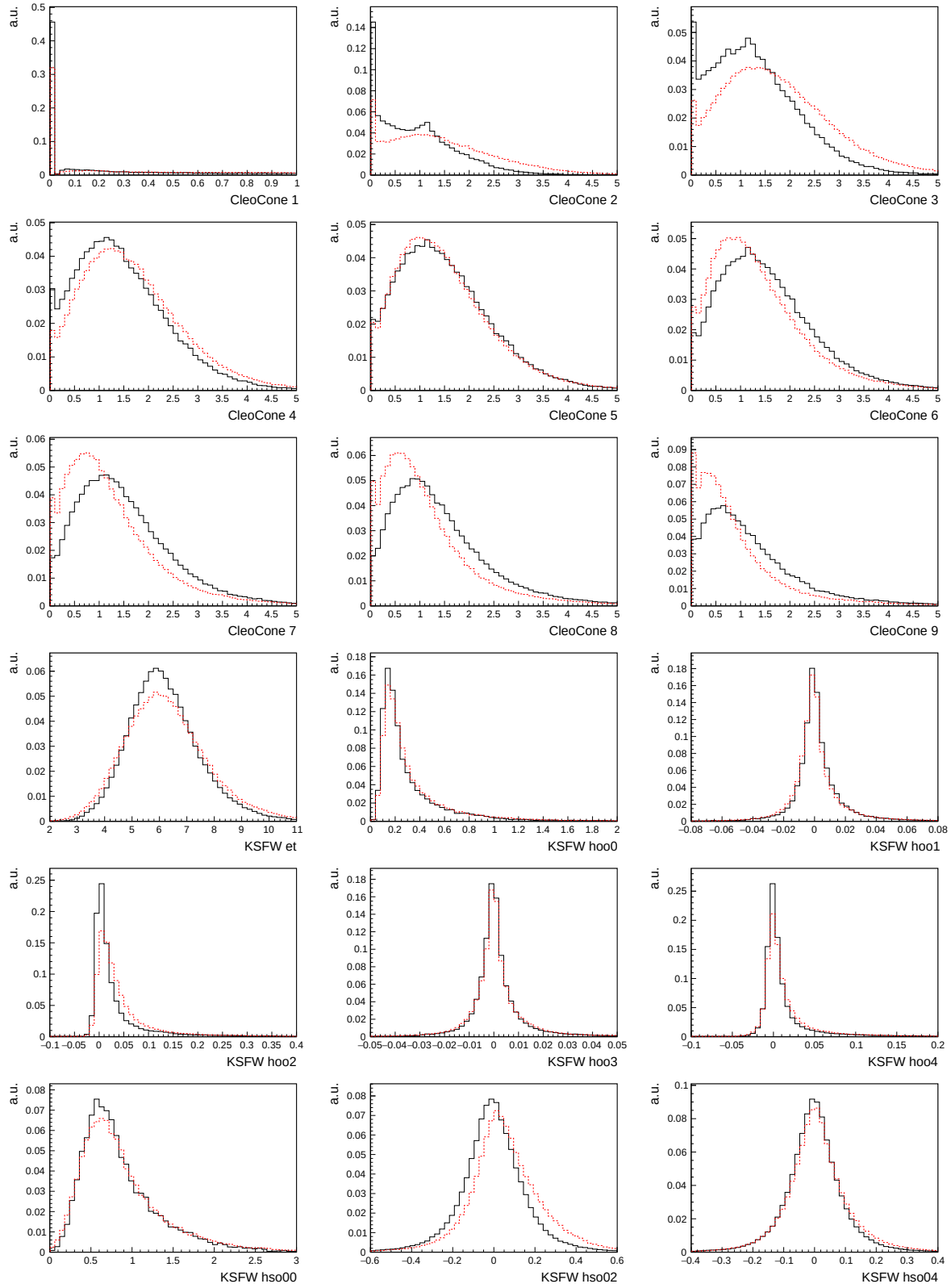


Figure B.1: The distributions of the input variables with the training samples. In each panel, the solid black line is for the signal events, and the dashed red line is for the continuum events. To compare easily, all distributions are normalized by the number of events.

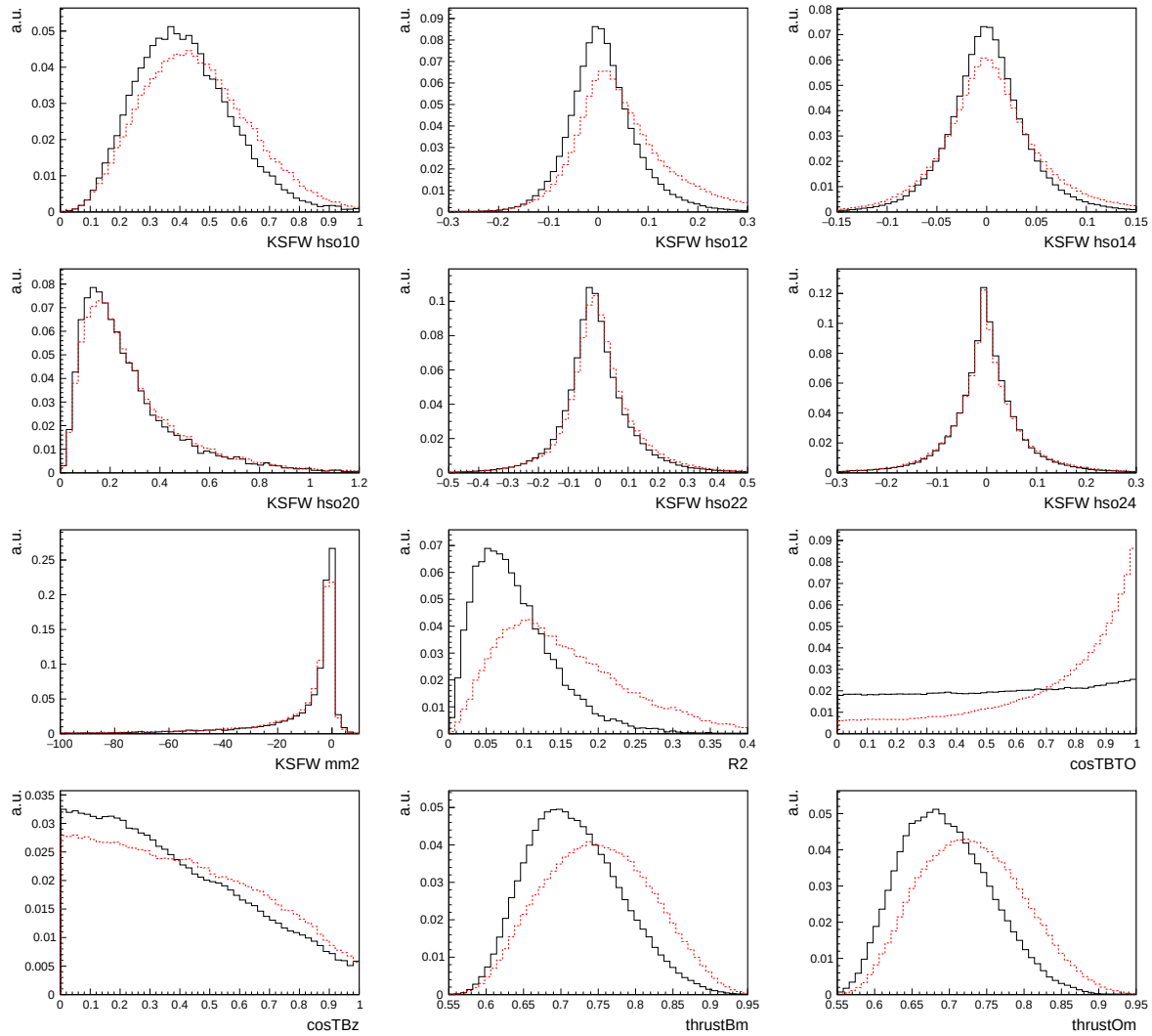
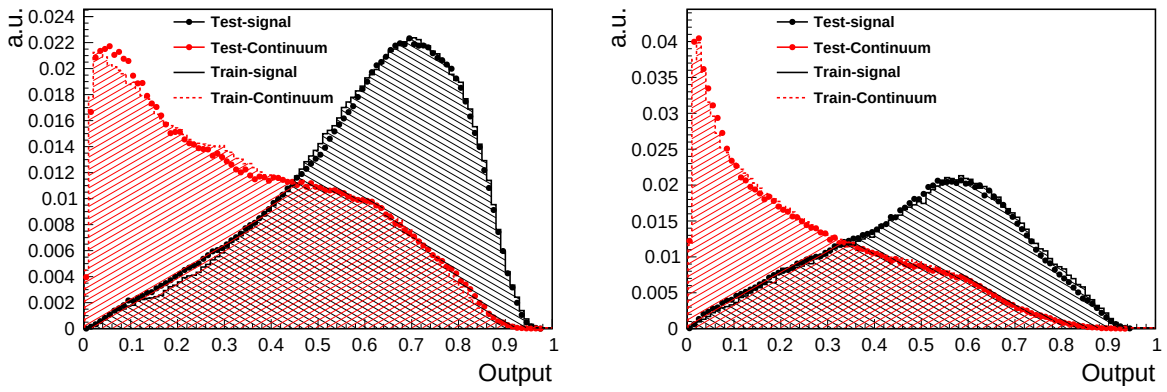
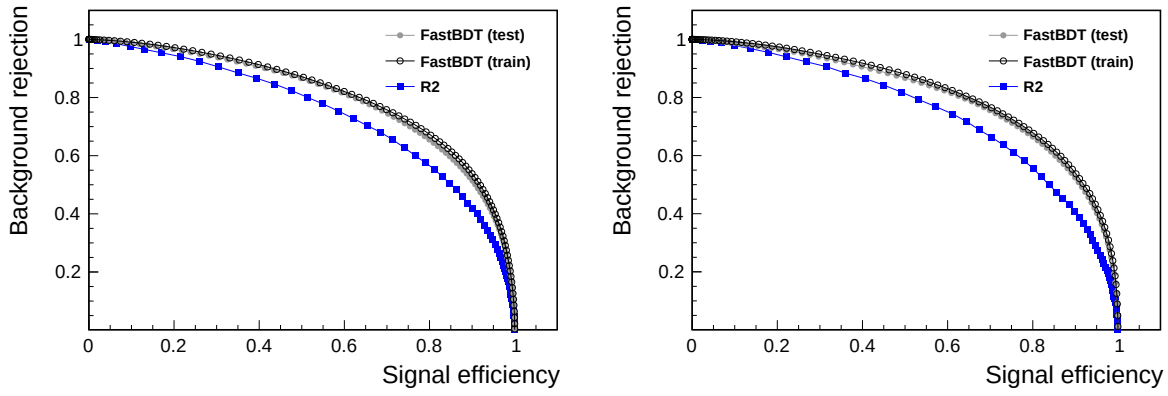


Figure B.2: Continuation of Fig. B.1.



(a) $B^+ \to X(3872)K^+$ (b) $B^0 \to X(3872)K^0$

Figure B.3: The distributions of the FastBDT output in the $X(3872)$ signal region. In each panel, the black and red points are the distributions of the signal and continuum events with the test sample, respectively. The striped black and red area are the distributions of the signal and continuum events with the training sample, respectively.



(a) $B^+ \to X(3872)K^+$ (b) $B^0 \to X(3872)K^0$

Figure B.4: The ROC curves of the FastBDT output (closed gray circles for the test sample and opened black circles for the training sample) and $R2$ (blue squares).

Appendix C

Parameterization of the $D^0\overline{D}^{*0}$ Invariant Mass Resolution

This appendix describes the parameterization of the $D^0\overline{D}^{*0}$ invariant mass resolution for the signal and broken-signal contributions. Before describing the parameterizations, the tendency of the mass resolution as a function of true mass is explained using the formulation.

C.1 Dependence of Mass Resolution on the $D^0\overline{D}^{*0}$ Invariant Mass

The dependence of mass resolution on the $D^0\overline{D}^{*0}$ invariant mass is examined by the propagation of resolution of the D^0 and D^{*0} invariant mass. The $D^0\overline{D}^{*0}$ invariant mass is obtained by the D^0 and $\overline{D}^{*0}$ three-momenta in the rest frame of $X(3872)$ as follows:

$$\begin{aligned} M(D^0\overline{D}^{*0}) &= E(D^0) + E(\overline{D}^{*0}) \\ &= m_{D^0}\sqrt{1+\frac{\mathbf{p}(D^0)^2}{m_{D^0}^2}} + m_{\overline{D}^{*0}}\sqrt{1+\frac{\mathbf{p}(\overline{D}^{*0})^2}{m_{\overline{D}^{*0}}^2}} \\ &\sim m_{D^0} + m_{\overline{D}^{*0}} + \frac{\mathbf{p}(D^0)^2}{2m_{D^0}} + \frac{\mathbf{p}(\overline{D}^{*0})^2}{2m_{\overline{D}^{*0}}}, \end{aligned} \tag{C.1}$$

where $\mathbf{p}$ is the three-momentum, and the approximation is used because the momentum is very small in the vicinity of the $D^0\overline{D}^{*0}$ threshold. The resolution can be expressed as follows by propagating from the resolution of the three-momenta,

$$\begin{aligned} \sigma M(D^0\overline{D}^{*0}) &= \sqrt{(\frac{\partial M(D^0\overline{D}^{*0})}{\partial|\mathbf{p}(D^0)|})^2(\sigma|\mathbf{p}(D^0)|)^2 + (\frac{\partial M(D^0\overline{D}^{*0})}{\partial|\mathbf{p}(\overline{D}^{*0})|})^2(\sigma|\mathbf{p}(\overline{D}^{*0})|)^2} \\ &= \sqrt{(\frac{\mathbf{p}(D^0)}{m_{D^0}})^2(\sigma|\mathbf{p}(D^0)|)^2 + (\frac{\mathbf{p}(\overline{D}^{*0})}{m_{\overline{D}^{*0}}})^2(\sigma|\mathbf{p}(\overline{D}^{*0})|)^2} \\ &\sim \sqrt{\frac{M(D^0\overline{D}^{*0}) - m_{D^0} - m_{D^{*0}}}{2m}(\sigma|\mathbf{p}|)^2} = \sqrt{\frac{Q}{2m}(\sigma|\mathbf{p}|)^2}, \end{aligned} \tag{C.2}$$

where the approximation from the second line to the third line is assumed that the momentum resolutions and masses of $\overline{D}^{*0}$ and D^0 are the same. Therefore, the resolution

increases roughly in proportion to the squared root of the Q value, which is the mass difference between the $D^0\overline{D}^{*0}$ invariant and the threshold.

C.2 Signal Reconstructed Correctly

The resolution PDF is defined by $M_{\rm diff}$ as follows,

$$\begin{aligned} f_{\rm res}(M_{\rm diff}) =& [c \cdot f_{\rm gauss}(M_{\rm diff}, \mu, \sigma_{\rm gauss}) + (1-c) \cdot f_{\rm CB}(M_{\rm diff}, \mu, \sigma_{\rm CB}, n, \alpha)] \\ & \times f_{\text{turn-on}}(M_{\rm diff}), \end{aligned} \tag{C.3}$$

where $f_{\rm gauss}$ and $f_{\rm CB}$ represent a Gaussian PDF and a Crystal Ball PDF, respectively. μ is a common mean, σ is a standard deviation, n is a value of the exponent of the power law, α is a transition point between the Gaussian and the power-law tail, and c is the fraction; Details are described in Sec. 5.1.1.

In parameterizing those parameters, it is difficult to determine at once due to the correlation among the parameters. Therefore, they are sequentially investigated using the following procedure.

1. The distributions are fitted with $f_{\rm res}$, which all parameters are floated. The correlation between $\sigma_{\rm CB}$ and $\sigma_{\rm gauss}$ is fitted with a first-order polynomial function (Fig. C.1). The intercept is almost consistent with zero, so we ignored it, and $\sigma_{\rm CB}/\sigma_{\rm gauss}$ is obtained as follows

$$\sigma_{\rm CB}/\sigma_{\rm gauss} = \begin{cases} 2.095 & \text{for } X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma] \text{ mode,} \\ 1.737 & \text{for } X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0] \text{ mode.} \end{cases}$$

2. The distributions are fitted with $f_{\rm res}$ fixed $\sigma_{\rm CB}/\sigma_{\rm gauss}$, and c as a function of the input mass is plotted in Fig. C.2. The relation fluctuates because the resolution function contains the multiple tail components, e.g., a tail on one side of the Crystal Ball function and a wide Gaussian (or Crystal Ball function), despite the small asymmetry of the distribution. Therefore, this relation is tentatively fitted and determined with a constant in this step,

$$c = \begin{cases} 0.2200 & \text{for } X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma] \text{ mode,} \\ 0.6931 & \text{for } X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0] \text{ mode.} \end{cases}$$

   It is re-investigated after determining the other parameters.

3. The distributions are fitted with $f_{\rm res}$, which $\sigma_{\rm CB}/\sigma_{\rm gauss}$, and c are fixed, and n as a function of the input mass is plotted in Fig. C.3. Since the dependence is very small, it can be approximated by a constant and determined as follows by fit:

$$n = \begin{cases} 134.12 & \text{for } X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma] \text{ mode,} \\ 127.18 & \text{for } X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0] \text{ mode.} \end{cases}$$

4. The distributions are fitted with $f_{\rm res}$ fixed $\sigma_{\rm CB}/\sigma_{\rm gauss}$, c and n, and α as a function of the input mass is plotted in Fig. C.4. The tail component decreases as the input mass increases. It is parameterized by fitting with a first-order polynomial function,

$$\alpha = \begin{cases} 18.0 - 4.99\ m_{X(3872)} & \text{for } X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma] \text{ mode,} \\ 67.1 - 17.6\ m_{X(3872)} & \text{for } X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0] \text{ mode.} \end{cases}$$

It is not a problem that the first-order polynomial function does not reproduce α as a function of the input mass completely because the final function with parameterized α can reproduce the $M_{\rm diff}$ distribution well.

5. The distributions are fitted with $f_{\rm res}$ fixed $\sigma_{\rm CB}/\sigma_{\rm gauss}$, c, n and α, and $\sigma_{\rm gauss}$ as a function of the input mass is plotted in Fig. C.5. In principle, the resolution increases roughly in proportion to the squared root of Q value, which is the mass difference between the $D^0\overline{D}^{*0}$ invariant and the threshold,

$$\sigma M(D^0\overline{D}^{*0}) \sim \sqrt{\frac{Q}{2m}(\sigma|\mathbf{p}|)^2}. \tag{C.4}$$

The relation follows the threshold function like Eq. (C.4). To take into account the effect of the higher-order terms, a constant is introduced in the threshold function. It is parameterized by

$$\sigma_{\rm gauss} = \begin{cases} 2.6043 \times 10^{-3}\sqrt{m_{X(3872)} - (m_{D^0} + m_{D^{*0}}) + 2.559 \times 10^{-4}}\ {\rm GeV}/c^2 \\ \qquad\qquad \text{for } X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma] \text{ mode}, \\ 4.343 \times 10^{-3}\sqrt{m_{X(3872)} - (m_{D^0} + m_{D^{*0}}) + 1.86 \times 10^{-4}}\ {\rm GeV}/c^2 \\ \qquad\qquad \text{for } X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0] \text{ mode}. \end{cases}$$

6. As with step 2, the distributions are fitted with $f_{\rm res}$ fixed $\sigma_{\rm CB}/\sigma_{\rm gauss}$, n, α, and $\sigma_{\rm gauss}$ floated c, and c as a function of the input mass is plotted in Fig. C.6. The relation is determined by fitting with a constant

$$c = \begin{cases} 0.1852 & \text{for } X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma] \text{ mode}, \\ 0.7003 & \text{for } X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0] \text{ mode}. \end{cases}$$

There is a small structure near the threshold, but the impact on the lineshape after convolution is negligible.

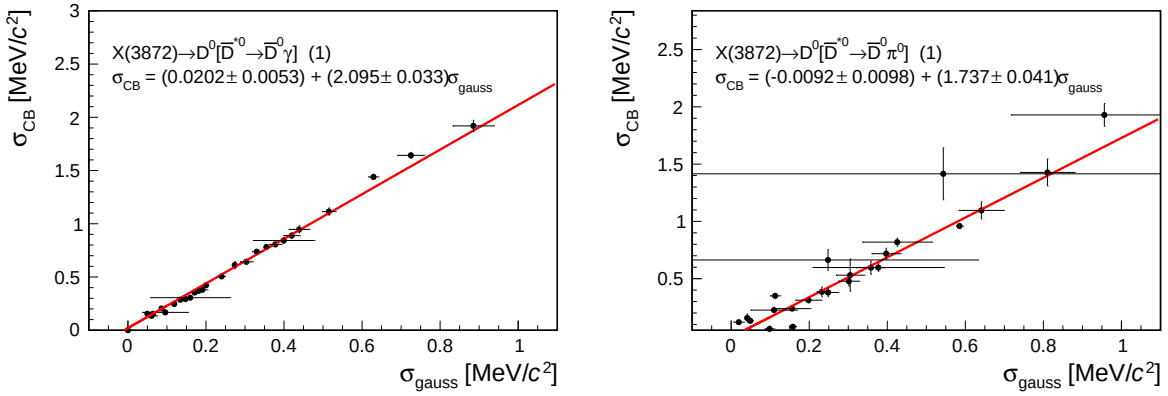


Figure C.1: The correlation between $\sigma_{\rm CB}$ and $\sigma_{\rm gauss}$ with the zero-width signal MC samples for $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$ (left) and $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$ (right). In each correlation plot, the results with failed fits are stripped out. The red line shows the fit result with a first-order polynomial function.

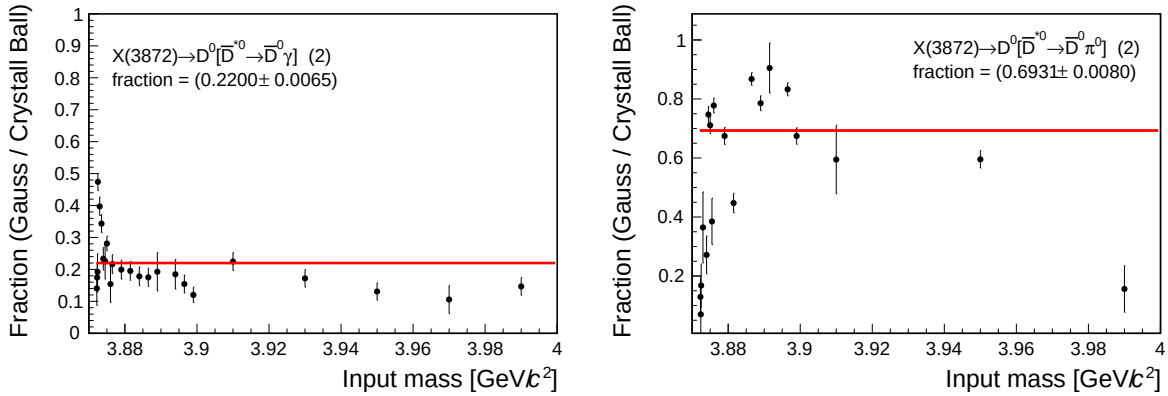


Figure C.2: The fraction c as a function of the input mass with the zero-width signal MC samples for $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$ (left) and $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$ (right).In each correlation plot, the results with failed fits are stripped out. The red line shows the fit result with a constant.

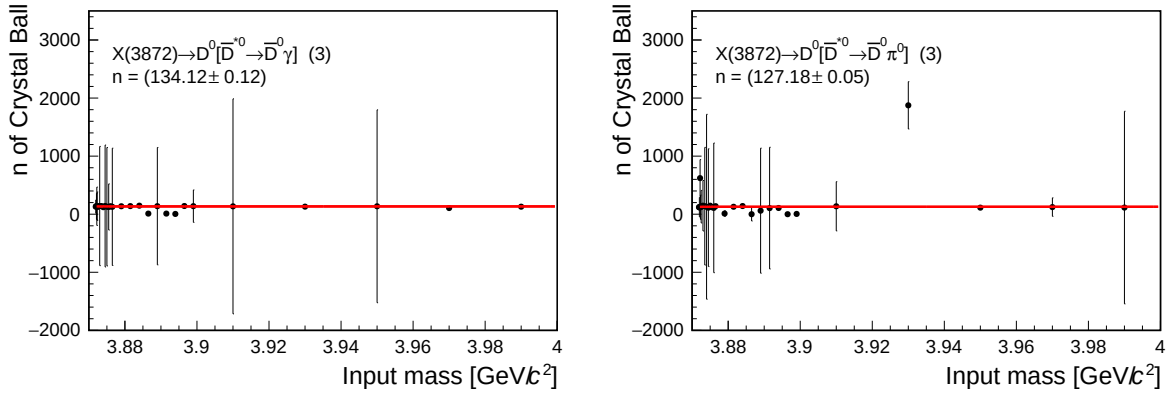


Figure C.3: The value of the exponent of the power law n as a function of the input mass with the zero-width signal MC samples for $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$ (left) and $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$ (right). In each correlation plot, the results with failed fits are stripped out. The red line shows the fit result with a constant.

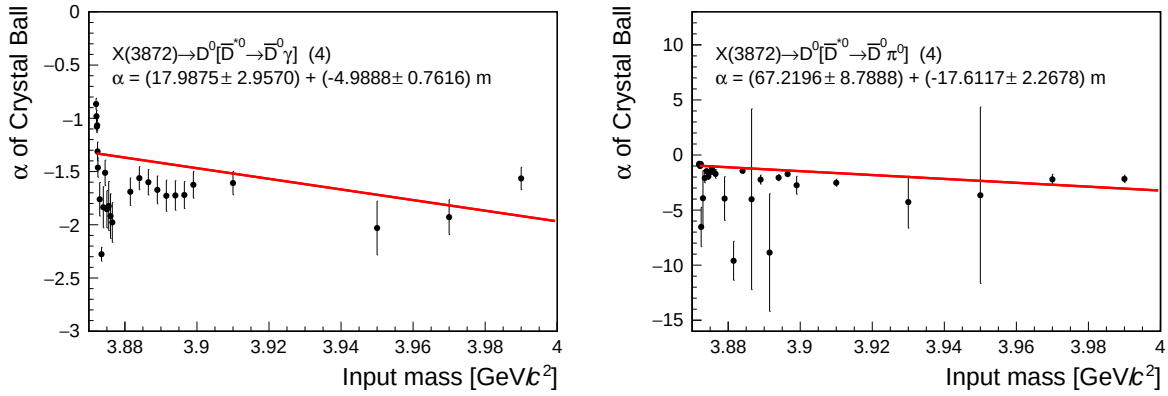


Figure C.4: The transition point between the Gaussian and the power-law tail α as a function of the input mass with the zero-width signal MC samples for $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$ (left) and $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$ (right). In each correlation plot, the results with failed fits are stripped out. The red line shows the fit result with a first-order polynomial function.

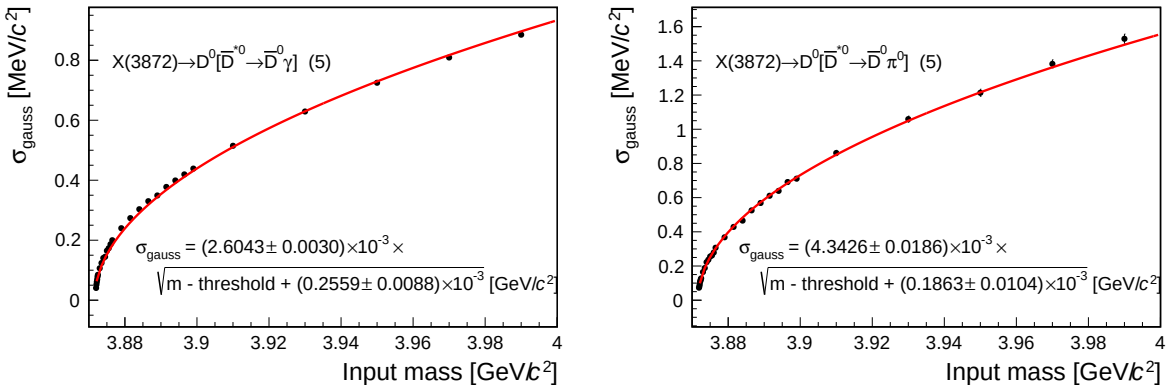


Figure C.5: The standard deviation of the Gaussian σ_{gauss} as a function of the input mass with the zero-width signal MC samples for $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$ (left) and $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$ (right). In each correlation plot, the results with failed fits are stripped out. The red line shows the fit result with a threshold function.

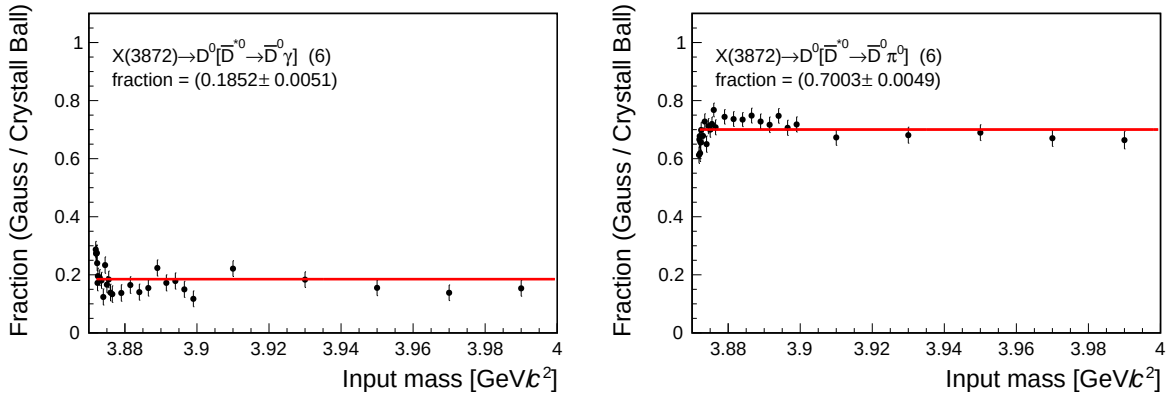


Figure C.6: The fraction c as a function of the input mass with the zero-width signal MC samples after fixing the other parameters for $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\gamma]$ (left) and $X(3872) \to D^0[\overline{D}^{*0} \to \overline{D}^0\pi^0]$ (right). The red line shows the fit result with a constant.

C.3 Signal Reconstructed Incorrectly for $\overline{D}^{*0} \to \overline{D}^0\pi^0$

The broken-signal resolution is parameterized by the input mass as

$$\begin{aligned} f_{\rm res}^{\rm broken}(M) =& \big(c_1 \cdot f_{\rm gauss}(M,\mu,\sigma_{\rm core}) + c_2 \cdot f_{\rm gauss}(M,\mu,\sigma_{\rm middle}) \\ &+ (1-c_1-c_2)\cdot f_{\rm gauss}(M,\mu,\sigma_{\rm tail})\big) \times f_{\text{turn-on}}(M), \end{aligned} \tag{C.5}$$

Here, $f_{\rm gauss}$ is a Gaussian PDF. $f_{\text{turn-on}}$ is the same definition as Eq. (5.3). The p_0 in $f_{\text{turn-on}}$ is smeared by the ratio of the standard deviation between the signal and the broken-signal. $\sigma_{\rm core}$, $\sigma_{\rm middle}$ and $\sigma_{\rm tail}$ are standard deviations for the Gaussian of the core, middle and tail components, respectively. c_1–c_2 is the fractions of the core and middle components. The five parameters of $\sigma_{\rm core}$, $\sigma_{\rm middle}$, $\sigma_{\rm tail}$, c_1 and c_2 are determined with the following procedure.

1. The $M(D^0\overline{D}^{*0})$ distributions for broken-signals are fitted with $f_{\rm res}^{\rm broken}$ which all parameters are floated. The correlation between $\sigma_{\rm core}$ and $\sigma_{\rm middle}$ and that between $\sigma_{\rm core}$ and $\sigma_{\rm tail}$ are plotted in Fig. C.7. Each correlation is fitted with a first-order polynomial function, and the relation is determined as follows:

$$\sigma_{\rm middle}/\sigma_{\rm core} = 2.97,$$

$$\sigma_{\rm tail}/\sigma_{\rm core} = 11.09.$$

   Here, the intercepts are always ignored because they are ideally zero.

2. The distributions are fitted with $f_{\rm res}^{\rm broken}$ fixed $\sigma_{\rm middle}/\sigma_{\rm core}$ and $\sigma_{\rm tail}/\sigma_{\rm core}$. The fractions as a function of the input mass are plotted in Fig. C.8 and fitted with a threshold function that converges to a constant for high mass,

$$c_1 = 0.2032 \times [1 - 1.0180 \cdot \exp[206.2(m_{D^0} + m_{D^{*0}} - m_{X(3872)})],$$

$$c_2 = 0.6006 \times [1 + 0.442 \cdot \exp[1386(m_{D^0} + m_{D^{*0}} - m_{X(3872)})].$$

   These relations are re-investigated after determining $\sigma_{\rm core}$, because the relations tend to depend on $\sigma_{\rm core}$, especially near the threshold.

3. The distributions are fitted with $f_{\rm res}^{\rm broken}$ fixed $\sigma_{\rm middle}/\sigma_{\rm core}$, $\sigma_{\rm tail}/\sigma_{\rm core}$, c_1, and c_2. $\sigma_{\rm core}$ as a function of the input mass is plotted in Fig. C.9. It is fitted with the threshold function using square root like Eq. (C.4), and the relation is obtained as

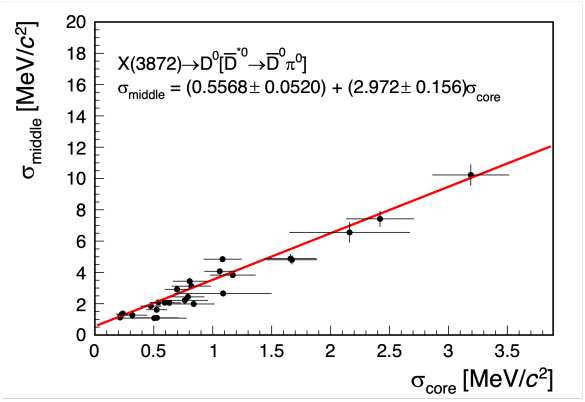
$$\sigma_{\rm core} = 9.137 \times 10^{-3}\sqrt{m_{X(3872)} - (m_{D^0} + m_{D^{*0}}) + 1.340 \times 10^{-3}}\ {\rm GeV}/c^2.$$

4. The distributions are fitted with $f_{\rm res}^{\rm broken}$ fixed $\sigma_{\rm middle}/\sigma_{\rm core}$, $\sigma_{\rm tail}/\sigma_{\rm core}$, and $\sigma_{\rm core}$ floated c_1 and c_2. The fractions as a function of the input mass are plotted in Fig. C.10 and fitted with the same threshold function as in step 2. The relations are obtained as

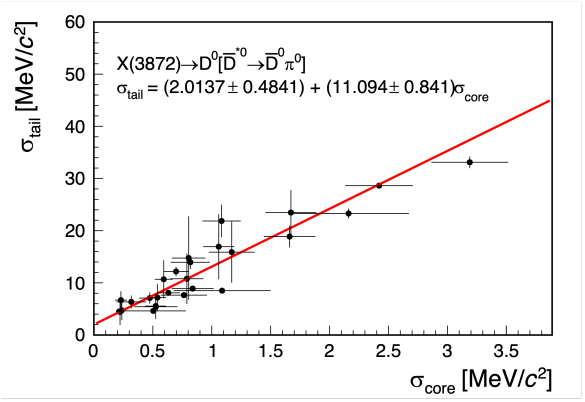
$$c_1 = 0.1896 \times [1 - 1.0557 \cdot \exp[524(m_{D^0} + m_{D^{*0}} - m_{X(3872)})],$$

$$c_2 = 0.5999 \times [1 + 0.425 \cdot \exp[1430(m_{D^0} + m_{D^{*0}} - m_{X(3872)})].$$

   Near the threshold, c_1 is less than zero, and c_2 is greater than one in these formulas, even though c_1 and c_2 must be values between zero and one. In such a case, we always reset it to zero or one.

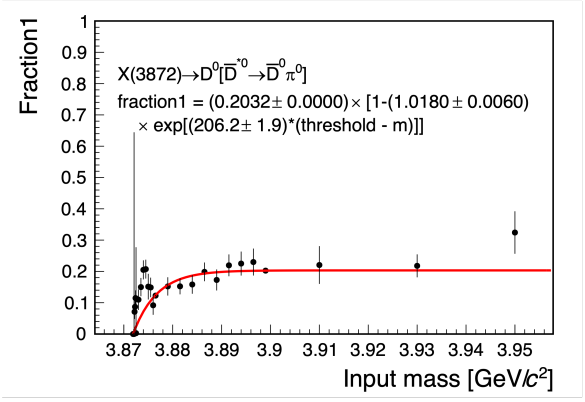


(a) The correlation between $\sigma_{\rm core}$ and $\sigma_{\rm middle}$.

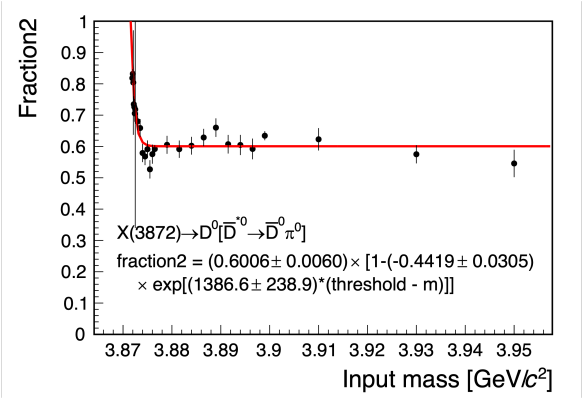


(b) The correlation between $\sigma_{\rm core}$ and $\sigma_{\rm tail}$.

Figure C.7: The correlation between $\sigma_{\rm core}$ and the standard deviation of the other Gaussian with the zero-width signal MC samples. The red line shows the fit result with a first-order polynomial function.



(a) The fraction of the core component c_1.



(b) The fraction of the middle component c_2.

Figure C.8: The fractions as a function of the input mass with the zero-width signal MC samples. The red line shows the fit result with the threshold function.

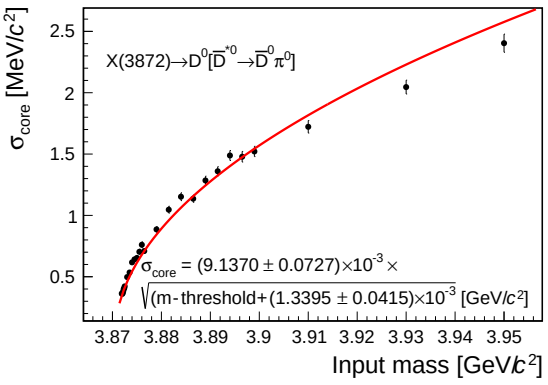
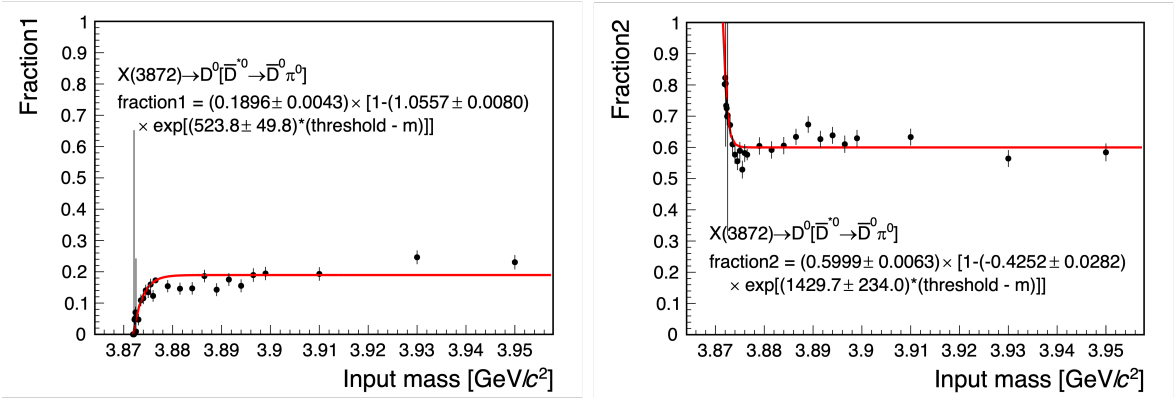


Figure C.9: The standard deviation of the core component σ_{core} as a function of the input mass with the zero-width signal MC samples. The red line shows the fit result with the threshold function using a square root.



(a) The fraction of the core component c_1. (b) The fraction of the middle component c_2.

Figure C.10: The fractions as a function of the input mass after fixing $\sigma_{\text{middle}}/\sigma_{\text{core}}$, $\sigma_{\text{tail}}/\sigma_{\text{core}}$, and σ_{core} with the zero-width signal MC samples. The red line shows the fit result with the threshold function.

Appendix D

Mass Resolution Difference between Data and MC

In this appendix, we investigate the difference in the $M(D^0\overline{D}^{*0})$ resolution in the MC and data, which is one source of the systematic uncertainty.

There is no perfect control sample, such as a resonance decaying into $D^0\overline{D}^{*0}$, except for $X(3872)$ itself. Therefore, examining the resolution of ΔE with $B^+ \to \overline{D}^{*0}\pi^+\pi^-\pi^+$ sample is used as an alternative. The reason for using $B^+ \to \overline{D}^{*0}\pi^+\pi^-\pi^+$ is that it has a similar decay topology to $B^+ \to [X(3872) \to \overline{D}^{*0}D^0]K$ and thus can cover low D^{*0} momentum like D^{*0} coming from the $X(3872)$ decays.

The reason for using ΔE is that the ΔE resolution change is similar as that of the $M(D^0\overline{D}^{*0})$ resolution when the track momentum and/or the cluster energy are smeared. Ratios of the resolution with smearing to that without smearing are shown in Fig. D.1, where three types of smearing are tried; smear cluster energy by 2% (red circles), smear cluster energy by 5% (green circles), and both smear cluster energy by 2% and track momentum by two times the scaling factors used in `smear_trk` (blue circles). `smear_trk` is a module that smears the helix parameters in the MC sample so as to get a better agreement between data and MC for the mass resolution of the hadrons reconstructed with only charged particles. We check the resolution ratio for ΔE, $M(D^0\gamma$ or $\pi^0)$, and $M(D^0)$ using signal MC samples of $B^+ \to \overline{D}^{*0}\pi^+\pi^-\pi^+$, and $M(D^0\overline{D}^{*0})$ using the default samples of $B^+ \to [X(3872) \to \overline{D}^{*0}D^0]K$. It can be confirmed that the behavior of the ΔE resolution reproduces that of the $M(D^0\overline{D}^{*0})$ resolution.

The ΔE resolution is evaluated using the background MC sample and the Belle full data. About the $B^+ \to \overline{D}^{*0}\pi^+\pi^-\pi^+$ selection, we require the same selection criteria for D^0, D^{*0}, and the final state particles (K^+, K^0_S, π^+, π^0, γ) as for $B^+ \to [X(3872) \to \overline{D}^{*0}D^0]K$. As with the $X(3872)$ analysis, a mass-constrained fit are applied for the π^0 candidates, and a mass- and vertex-constrained fit are applied for the K^0_S and D^0 candidates. In addition, only candidates with low momentum D^{*0}, which is less than 1.6 GeV/c, are used to select the same momentum regions as D^{*0} in $B^+ \to [X(3872) \to \overline{D}^{*0}D^0]K$. For B selection, $M_{\rm bc}$ is required to be within 4.5 MeV/c^2 of nominal B mass. For continuum suppression, $R2$ and $\cos TBTO$ are required to be less than 0.2 and 0.8, respectively. To reduce combinatorial background, we requier the number of tracks in a single event $N_{\rm trk} < 17$ if D^0 is reconstructed with a π^0; otherwise, we require $N_{\rm trk} < 15$ and $\chi^2 > 0.01$ of the π^0 mass-constraint fit.

The ΔE distribution for $\overline{D}^{*0} \to \overline{D}^0\gamma$ and $\overline{D}^{*0} \to \overline{D}^0\pi^0$ are shown in Figs. D.2 and D.3, respectively. Here, only for $\overline{D}^{*0} \to \overline{D}^0\gamma$, the $\overline{D}^0$ reconstruction mode is restricted

to $K^+\pi^-$ to reduce the huge background. The resolution is evaluated by fitting the distribution with a double-Gaussian function for the signal component and a second-order Chebyshev function for the background component. First, the standard deviations of the two Gaussians are determined using the ΔE distribution of the background MC sample (left panels in Figs. D.2–D.3). When fitting the data distribution (right panels), the standard deviations are fixed to the fit result for the MC sample, but a common scaling factor to the standard deviations is floated. The scaling factors are obtained as 1.01 ± 0.10 for $\overline{D}^{*0} \to \overline{D}^0\gamma$ and 1.08 ± 0.13 for $\overline{D}^{*0} \to \overline{D}^0\pi^0$. It indicates that the $M(D^0\overline{D}^{*0})$ resolution of MC is consistent with that of data within a precision of 13%.

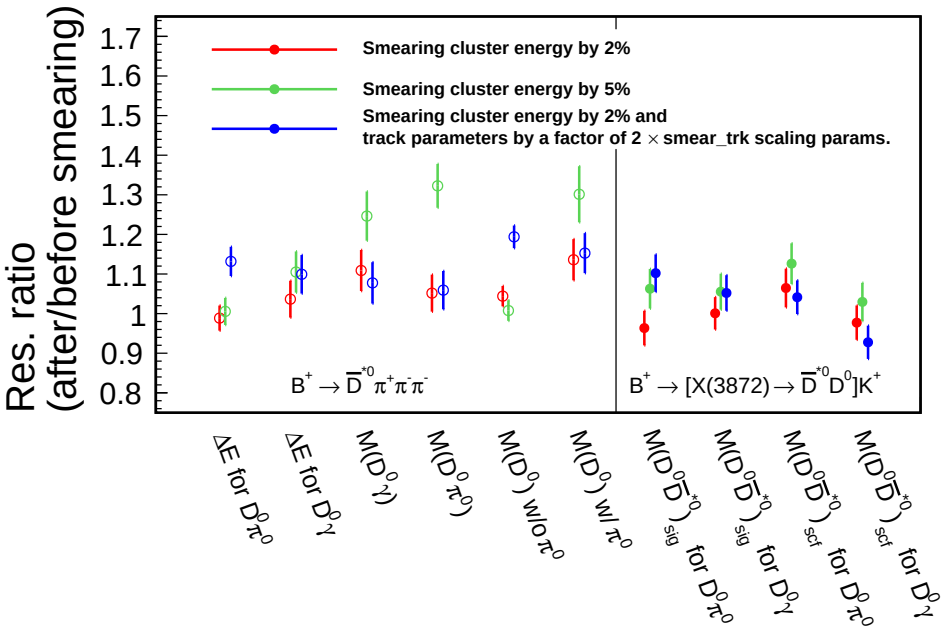
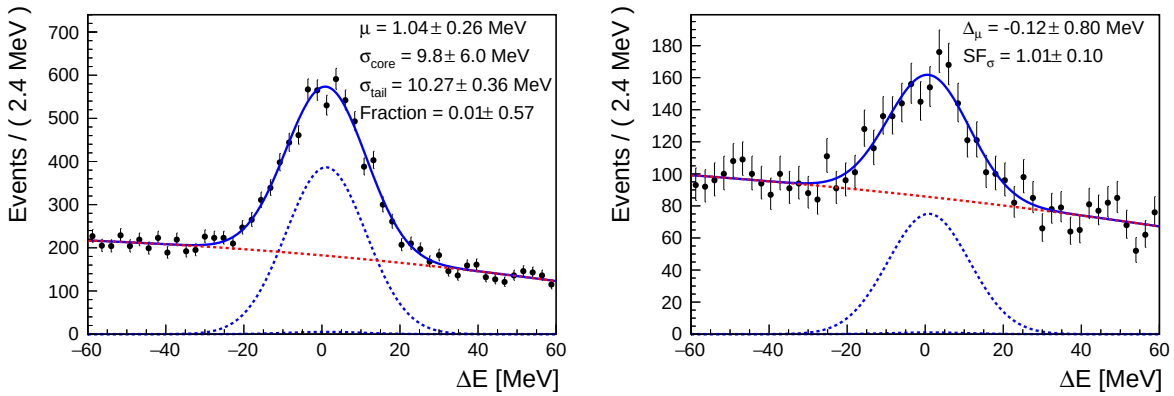
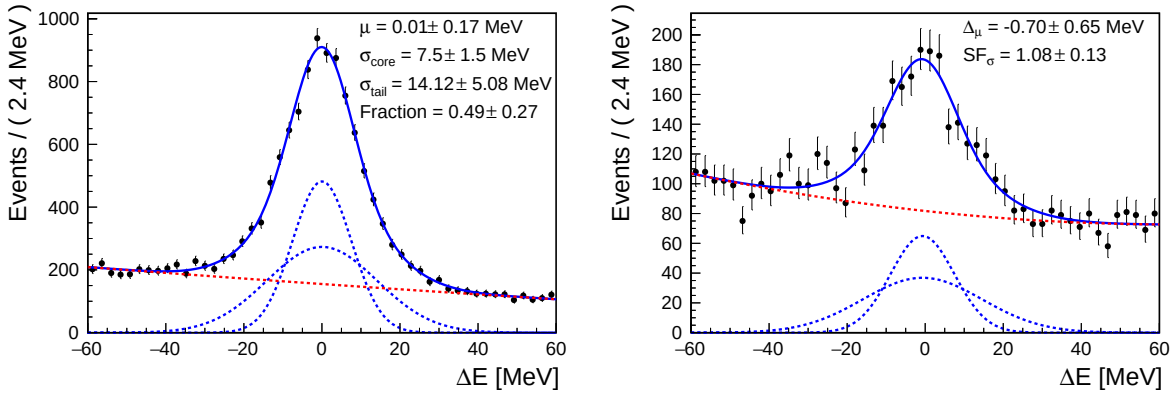


Figure D.1: Ratios of the resolutions with smearing to that without smearing. The six from the left show the result for ΔE of $\overline{D}^{*0} \to \overline{D}^0\pi^0$, that of $\overline{D}^{*0} \to \overline{D}^0\gamma$, $M(\overline{D}^0\gamma)$, $M(\overline{D}^0\pi^0)$, $M(D^0)$ using $\overline{D}^0$ mode with π^0 and that without π^0 for signal MC samples of $B^+ \to \overline{D}^{*0}\pi^+\pi^-\pi^+$. The four from the right show the $M(D^0\overline{D}^{*0})$ of signal candidates for $\overline{D}^{*0} \to \overline{D}^0\pi^0$, that for $\overline{D}^{*0} \to \overline{D}^0\gamma$, $M(D^0\overline{D}^{*0})$ of broken-signal candidates for $\overline{D}^{*0} \to \overline{D}^0\pi^0$, and that for $\overline{D}^{*0} \to \overline{D}^0\gamma$. For each point, three types of smearing are attempted; smear cluster energy by 2% (red circles), smear cluster energy by 5% (green circles), and smear cluster energy by 2% and track momentum by two times the scaling factors used in `smear_trk` (blue circles).



(a) The background MC sample (b) The real data sample

Figure D.2: The ΔE distributions of $B^+ \to [\overline{D}^{*0} \to \overline{D}^0\gamma]\pi^+\pi^-\pi^+$. In each panel, the solid blue line is the total fit result. The dashed blue represents each Gaussian in the double Gaussian for the signal contributions. The dashed red line shows a second-order Chebyshev function as the background component. In the left panel, μ, $\sigma_{\rm core}$, and $\sigma_{\rm tail}$ denote a common mean, a standard deviation of the core Gaussian, and a standard deviation of the tail Gaussian, respectively. In the right panel, Δ_μ and SF_μ denote the difference in μ between data and MC and the scale factor of the σ's for data, respectively.



(a) The background MC sample (b) The real data sample

Figure D.3: The ΔE distributions of $B^+ \to [\overline{D}^{*0} \to \overline{D}^0\pi^0]\pi^+\pi^-\pi^+$. The notations are the same as in Fig. D.2.

Appendix E

Derivation of a Lower Limit for the Coupling Constant of the Flatté lineshape

In principle, a lower limit at certain credibility is set by

$$\alpha = \int_{\theta_{\rm lower}}^{\infty} p(\theta|\boldsymbol{x})d\theta, \tag{E.1}$$

where α denotes the credibility, $p(\theta|\boldsymbol{x})$ is the posterior probability of the desired theoretical parameter θ, and $\boldsymbol{x}$ is the data. For the Flatté lineshape study, θ is the coupling constant g. Following common practice, we use a uniform prior probability of g, and then we get the following requirement in this study.

$$\int_{g_{\rm lower}}^{\infty} \mathcal{L}dg = \alpha \int_{0}^{\infty} \mathcal{L}dg, \tag{E.2}$$

where $g_{\rm lower}$ is the lower limit. On the other hand, we use the following different equations to derive the lower limit in the Flatté lineshape study:

$$\int_{g_{\rm lower}}^{g_{\rm best}} \mathcal{L}dg = \beta \int_{0}^{g_{\rm best}} \mathcal{L}dg, \tag{E.3}$$

where $g_{\rm best}$ is the g at the maximum likelihood. In this appendix, we prove that this lower limit derivation formula gives us a conservative lower limit at $(\beta+1)/2$ credibility when the $\mathcal{L}$ function is asymmetric about $g = g_{\rm best}$ and the right-handed area is larger than the left-handed area.

First, $\int_{g_{\rm best}}^{\infty}$ is added to both sides of Eq. (E.3).

$$\begin{aligned} \int_{g_{\rm lower}}^{g_{\rm best}} \mathcal{L}dg + \int_{g_{\rm best}}^{\infty} \mathcal{L}dg &= \beta \int_{0}^{g_{\rm best}} \mathcal{L}dg + \int_{g_{\rm best}}^{\infty} \mathcal{L}dg \\ \int_{g_{\rm lower}}^{\infty} \mathcal{L}dg &= \beta \int_{0}^{g_{\rm best}} \mathcal{L}dg + \int_{g_{\rm best}}^{\infty} \mathcal{L}dg \end{aligned} \tag{E.4}$$

This right-hand side can be transformed as

$$\begin{aligned}
&\beta \int_0^{g_{\rm best}} \mathcal{L} dg + \int_{g_{\rm best}}^{\infty} \mathcal{L} dg \\
&\quad = -(1-\beta) \int_0^{g_{\rm best}} \mathcal{L} dg + \int_0^{g_{\rm best}} \mathcal{L} dg + \int_{g_{\rm best}}^{\infty} \mathcal{L} dg \\
&\quad = \int_0^{\infty} \mathcal{L} dg - (1-\beta) \int_0^{g_{\rm best}} \mathcal{L} dg,
\end{aligned} \tag{E.5}$$

According to the asymmetry of the $\mathcal{L}$ function, we obtain the following relationship:

$$\int_0^{g_{\rm best}} \mathcal{L} dg < \frac{1}{2} \int_0^{\infty} \mathcal{L} dg. \tag{E.6}$$

Based on Eq. (E.6), Eq. (E.5) is given by

$$\int_0^{\infty} \mathcal{L} dg - (1-\beta) \int_0^{g_{\rm best}} \mathcal{L} dg > \frac{1+\beta}{2} \int_0^{\infty} \mathcal{L} dg, \tag{E.7}$$

where $1-\beta$ is always positive. Thus, Eq. (E.4) is written as

$$\int_{g_{\rm lower}}^{\infty} \mathcal{L} dg > \frac{1+\beta}{2} \int_0^{\infty} \mathcal{L} dg. \tag{E.8}$$

Compared with Eq. (E.2), this relation means that the credibility of the lower limit is greater than $(\beta+1)/2$. Therefore, we can get the conservative lower limit of 90% (95%) credibility for $\beta = 0.8$ (0.9) from Eq. (E.3).

Reference

[1] S.-K. Choi et al. (Belle Collaboration), "Observation of a Narrow Charmoniumlike State in Exclusive $B^{\pm} \to K^{\pm}\pi^{+}\pi^{-}J/\psi$ Decays", Phys. Rev. Lett. **91**, 262001 (2003).

[2] T. Aushev et al. (Belle Collaboration), "Study of the $B \to X(3872)(\to D^{*0}\bar{D}^{0})K$ decay", Phys. Rev. D **81**, 031103 (2010).

[3] B. Aubert et al. (*BABAR* Collaboration), "Study of resonances in exclusive B decays to $\overline{D}^{(*)}D^{(*)}K$", Phys. Rev. D **77**, 011102 (2008).

[4] R. Aaij et al. (LHCb Collaboration), "Study of the lineshape of the $\chi_{c1}(3872)$ state", Phys. Rev. D **102**, 092005 (2020).

[5] M. Gell-Mann, "A schematic model of baryons and mesons", Phys. Lett. **8**, 214 (1964).

[6] G. Zweig, "An SU(3) model for strong interaction symmetry and its breaking. Version 2", in *DEVELOPMENTS IN THE QUARK THEORY OF HADRONS. VOL. 1. 1964 - 1978*, edited by D. B. Lichtenberg and S. P. Rosen (Feb. 1964), pp. 22–101.

[7] T. Nakano and K. Nishijima, "Charge Independence for V-particles", Prog. Theor. Phys. **10**, 581 (1953).

[8] R. J. Jaffe, "Multiquark hadrons. I. Phenomenology of $Q^2\overline{Q}^2$ mesons", Phys. Rev. D **15**, 267 (1977).

[9] R. L. Jaffe, "Multiquark hadrons. II. Methods", Phys. Rev. D **15**, 281 (1977).

[10] J. .-. Ader, J. .-. Richard, and P. Taxil, "Do narrow heavy multiquark states exist?", Phys. Rev. D **25**, 2370 (1982).

[11] H. J. Lipkin, "New possibilities for exotic hadrons — anticharmed strange baryons", Phys. Lett. B **195**, 484 (1987).

[12] R. L. Jaffe, "Perhaps a Stable Dihyperon", Phys. Rev. Lett. **38**, 195 (1977).

[13] N. A. Törnqvist, "From the deuteron to deusons, an analysis of deuteronlike meson-meson bound states", Z. Phys. C - Particles and Fields **61**, 525 (1994).

[14] H. Fritzsch and P. Minkowski, "Ψ-resonances, gluons and the Zweig rule", Nouv Cim A **30**, 393 (1975).

[15] R. Jaffe and K. Johnson, "Unconventional states of confined quarks and gluons", Phys. Lett. B **60**, 201 (1976).

[16] P. G. O. Freund and Y. Nambu, "Dynamics of the Zweig-Iizuka Rule and a New Vector Meson below 2 GeV/c^2", Phys. Rev. Lett. **34**, 1645 (1975).

[17] D. Horn and J. Mandula, "Model of mesons with constituent gluons", Phys. Rev. D **17**, 898 (1978).

[18] R. L. Workman et al. (Particle Data Group), "Review of Particle Physics", Prog. Theor. Exp. Phys. **2022**, 083C01 (2022).

[19] S. Godfrey and N. Isgur, "Mesons in a relativized quark model with chromodynamics", Phys. Rev. D **32**, 189 (1985).

[20] T. Barnes, S. Godfrey, and E. S. Swanson, "Higher charmonia", Phys. Rev. D **72**, 054026 (2005).

[21] LHCb Collaboration, "Exotic hadron naming convention", arXiv:2206.15233 (2022).

[22] S. L. Olsen, T. Skwarnicki, and D. Zieminska, "Nonstandard heavy mesons and baryons: Experimental evidence", Rev. Mod. Phys. **90**, 015003 (2018).

[23] A. Hosaka et al., "Exotic hadrons with heavy flavors: X, Y, Z, and related states", Prog. Theor. Exp. Phys. **2016**, 062C01 (2016).

[24] V. M. Abazov et al. (D0 Collaboration), "Observation and Properties of the $X(3872)$ Decaying to $J/\psi\pi^+\pi^-$ in $p\overline{p}$ Collisions at $\sqrt{s} = 1.96$ TeV", Phys. Rev. Lett. **93**, 162002 (2004).

[25] B. Aubert et al. (BABAR Collaboration), "Study of the $B^- \to J/\psi K^-\pi^+\pi^-$ decay and measurement of the $B^- \to X(3872)K^-$ branching fraction", Phys. Rev. D **71**, 071103 (2005).

[26] D. Acosta et al. (CDF II Collaboration), "Observation of the Narrow State $X(3872) \to J/\psi\pi^+\pi^-$ in $\overline{p}p$ Collisions at $\sqrt{s} = 1.96$ TeV", Phys. Rev. Lett. **93**, 072001 (2004).

[27] R. Aaij et al. (LHCb Collaboration), "Observation of $X(3872)$ production in pp collisions at $\sqrt{s} = 7$ TeV", Eur. Phys. J. C **72**, 1972 (2012).

[28] M. Ablikim et al. (BESIII Collaboration), "Observation of $e^+e^- \to \gamma X(3872)$ at BESIII", Phys. Rev. Lett. **112**, 092001 (2014).

[29] R. Aaij et al. (LHCb Collaboration), "Determination of the $X(3872)$ Meson Quantum Numbers", Phys. Rev. Lett. **110**, 222001 (2013).

[30] R. Aaij et al. (LHCb Collaboration), "Quantum numbers of the $X(3872)$ state and orbital angular momentum in its $\rho^0 J/\psi$ decay", Phys. Rev. D **92**, 011102 (2015).

[31] S.-K. Choi et al. (Belle Collaboration), "Bounds on the width, mass difference and other properties of $X(3872) \to \pi^+\pi^- J/\psi$ decays", Phys. Rev. D **84**, 052004 (2011).

[32] LHCb Collaboration, "Observation of sizeable ω contribution to $\chi_{c1}(3872) \to pi^+\pi^- J/\psi$ decays", arXiv:2204.12597 (2022).

[33] K. Abe et al. (Belle Collaboration), 2005.

[34] P. del Amo Sanchez et al. (*BABAR* Collaboration), "Evidence for the decay $X(3872) \to J/\psi\omega$", Phys. Rev. D **82**, 011101 (2010).

[35] M. Ablikim et al. (BESIII Collaboration), "Study of $e^+e^- \to \gamma\omega J/\psi$ and Observation of $X(3872) \to \omega J/\psi$", Phys. Rev. Lett. **122**, 232002 (2019).

[36] N. A. Törnqvist, "Isospin breaking of the narrow charmonium state of Belle at 3872 MeV as a deuson", Phys. Lett. B **590**, 209 (2004).

[37] E. S. Swanson, "Short range structure in the $X(3872)$", Phys. Lett. B **588**, 189 (2004).

[38] C.-Y. Wong, "Molecular states of heavy quark mesons", Phys. Rev. C **69**, 055202 (2004).

[39] D. Gamermann and E. Oset, "Isospin breaking effects in the $X(3872)$ resonance", Phys. Rev. D **80**, 014003 (2009).

[40] M. Takizawa and S. Takeuchi, "$X(3872)$ as a hybrid state of charmonium and the hadronic molecule", Prog. Theor. Exp. Phys. **2013**, 093D01 (2013).

[41] C. Hanhart et al., "Reconciling the $X(3872)$ with the near-threshold enhancement in the $D^0\bar{D}^{*0}$ final state", Phys. Rev. D **76**, 034007 (2007).

[42] E. Braaten and M. Lu, "Line shapes of the $X(3872)$", Phys. Rev. D **76**, 094028 (2007).

[43] D. V. Bugg, "How resonances can synchronize with thresholds", J. Phys. G: Nucl. Part. Phys. **35**, 075005 (2008).

[44] T. Hyodo and M. Niiyama, "QCD and the strange baryon spectrum", Progress in Particle and Nuclear Physics **120**, 103868 (2021).

[45] Our own average using the most recent measurements from Belle and *BABAR* [31, 71].

[46] S. Flatté, "Coupled-channel analysis of the $\pi\eta$ and $K\bar{K}$ systems near $K\bar{K}$ threshold", Phys. Lett. B **63**, 224 (1976).

[47] G. S. Adams et al. (CLEO Collaboration), "Amplitude analyses of the decays $\chi_{c1} \to \eta\pi^+\pi^-$ and $\chi_{c1} \to \eta^{'}\pi^+\pi^-$", Phys. Rev. D **84**, 112009 (2011).

[48] M. Ablikim et al. (BES Collaboration), "Resonances in $J/\psi \to \phi\pi^+\pi^-$ and ϕK^+K^-", Phys. Lett. B **607**, 243 (2005).

[49] G. Bonvicini et al. (CLEO Collaboration), "Dalitz plot analysis of the $D^+ \to \pi^-\pi^+\pi^+$ decay", Phys. Rev. D **76**, 012001 (2007).

[50] A. Garmash et al. (Belle Collaboration), "Evidence for Large Direct CP Violation in $B^\pm \to \rho(770)^0 K^\pm$ from Analysis of Three-Body Charmless $B^\pm \to K^\pm\pi^\pm\pi^\pm$ Decays", Phys. Rev. Lett. **96**, 251803 (2006).

[51] T. Ishikawa et al., "Resonance-like structure near the ηd threshold in the $\gamma d \to \pi^0\eta d$ reaction", Phys. Rev. C **104**, L052201 (2021).

[52] Y. S. Kalashnikova and A. V. Nefediev, "Nature of $X(3872)$ from data", Phys. Rev. D **80**, 074004 (2009).

[53] J. R. Taylor, *Scattering Theory* (Wiley, New York, 1972).

[54] W. R. Frazer and A. W. Hendry, "S-Matrix Poles Close to Threshold", Phys. Rev. **134**, B1307 (1964).

[55] V. Baru et al., "Flatté -like distributions and the $a_0(980)/f_0(980) mesons$", Eur. Phys. J. A **23**, 523 (2005).

[56] G. Gokhroo et al. (Belle Collaboration), "Observation of a Near-Threshold $D^0\bar{D}^0\pi^0$ Enhancement in $B \to D^0\bar{D}^0\pi^0 K$ Decay", Phys. Rev. Lett. **97**, 162002 (2006).

[57] S. Kurokawa and E. Kikutani, "Overview of the KEKB accelerators", Nucl. Instrum. Methods Phys. Res., Sect A **499**, and other papers included in this volume., 1 (2003).

[58] T. Abe et al., "Achievements of KEKB", Prog. Theor. Exp. Phys. **2013**, 03A001 (2013).

[59] A. Abashian et al. (Belle Collaboration), "The Belle detector", Nucl. Instrum. Methods Phys. Res., Sect A **479**, 117 (2002).

[60] A. J. Bevan et al., "The Physics of the B Factories", Eur. Phys. J. C **74**, 3026 (2014).

[61] Z. Natkaniec et al., "Status of the Belle silicon vertex detector", Nucl. Instrum. Methods Phys. Res. Sect. A **560**, 1 (2006).

[62] D. J. Lange, "The EvtGen particle decay simulation package", Nucl. Instrum. Methods Phys. Res., Sect A **462**, 152 (2001).

[63] T. Sjöstrand, "High-energy-physics event generation with PYTHIA 5.7 and JETSET 7.4", Computer Physics Communications **82**, 74 (1994).

[64] R. Brun et al., *GEANT3*, CERN Report No. DD/EE/84-1, 1987.

[65] E. Nakano, "Belle PID", Nucl. Instrum. Methods Phys. Res., Sect A **494**, 402 (2002).

[66] M. Feindt and U. Kerzel, "The NeuroBayes neural network package", Nucl. Instrum. Methods Phys. Res., Sect A **559**, 190 (2006).

[67] T. Keck, "FastBDT: A Speed-Optimized Multivariate Classification Algorithm for the Belle II Experiment", Comput. Softw. Big Sci. **1**, 2 (2017).

[68] T. Skwarnicki, PhD thesis (Institute of Nuclear Physics, Krakow, 1986).

[69] R. Brun, others Fons Rademakers, et al., *root-project/root: v6.18/02*, version v6-18-02, Aug. 2019.

[70] L. Demortier and L. Lyons, "Everything you always wanted to know about pulls", CDF/ANAL/PUBLIC/5776.

[71] B. Aubert et al. (*BABAR* Collaboration), "Study of $B \to X(3872)K$, with $X(3872) \to J/\psi\pi^+\pi^-$", Phys. Rev. D **77**, 111101 (2008).

[72] R. Aaij et al. (LHCb Collaboration), "Study of the $\psi_2(3823)$ and $\chi_{c1}(3872)$ states in $B^+ \to (J/\psi\pi^+\pi^-)K^+$ decays", J. High Energy Phys. **2020**, 123 (2020).

[73] Till Moritz Karbach and Maximilian Schlupp, "Constraints on Yield Parameters in Extended Maximum Likelihood Fits", arXiv:1210.7141 (2012).

[74] H. Albrecht et al. (ARGUS Collaboration), "Search for hadronic $b \to u$ decays", Phys. Lett. B **241**, 278 (1990).

[75] C. Hanhart, Y. S. Kalashnikova, and A. V. Nefediev, "Lineshapes for composite particles with unstable constituents", Phys. Rev. D **81**, 094028 (2010).

[76] G. Barucca et al. (PANDA Collaboration), "Precision resonance energy scans with the PANDA experiment at FAIR", Eur. Phys. J. A **55**, 42 (2019).

[77] T. Abe et al., "Belle II Technical Design Report", arXiv:1011.0352, KEK Report 2010-1 (2010).

[78] K. Akai, K. Furukawa, and H. Koiso, "SuperKEKB collider", Nucl. Instrum. Methods Phys. Res., Sect A **907**, 188 (2018).

[79] E. Kou et al., "The Belle II Physics Book", Prog. Theor. Exp. Phys. **2019**, 123C01 (2019).

[80] F. Fang et al. (Belle Collaboration), "Measurement of Branching Fractions for $B \to \eta_c K^{(*)}$ Decays", Phys. Rev. Lett. **90**, 071801 (2003).

[81] G. C. Fox and S. Wolfram, "Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes", Phys. Rev. Lett. **41**, 1581 (1978).

[82] T. Hara et al., "Computing at the Belle II experiment", J. Phys. Conf. Ser. **664**, 012002 (2015).

[83] K. Yuji et al., "Overview of the Belle II Computing", PoS(KMI2017)024 (2017).

[84] F. Stagni et al., *DIRACGrid/DIRAC: v6r20p15*, version v6r20p15, Oct. 2018.

[85] H. Miyake et al., "Belle II production system", J. Phys. Conf. Ser. **664**, 052028 (2015).

[86] T. Kuhr et al., "The Belle II Core Software", Comput. Softw. Big Sci. **3**, 1 (2018).

[87] H. Hikari et al., "Development of a Scout Job Framework for Improving Efficiency of Analysis Jobs in Belle II Distributed Computing System", PoS(ISGC2022)029 (2022).

[88] R. Aaij et al. (LHCb collaboration), "Search for the lepton flavour violating decay $\tau^- \to \mu^-\mu^+\mu^-$", J. High Energy Phys. **2015**, 121 (2015).

[89] S. H. Lee et al. (Belle Collaboration), "Evidence for $B^0 \to \pi^0\pi^0$", Phys. Rev. Lett. **91**, 261801 (2003).

[90] D. M. Asner et al. (CLEO Collaboration), "Search for exclusive charmless hadronic B decays", Phys. Rev. D **53**, 1039 (1996).