Measurement of the Branching Fraction and Polarization of the τ Lepton in the Decay $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ at the Belle Experiment

Belle実験における$\bar{B} \to D^* \tau^- \bar{\nu}_\tau$崩壊の
崩壊分岐比およびτレプトン偏極度測定

Shigeki HIROSE

Graduate School of Science, Nagoya University

February, 2017

Abstract

The decay $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ is a semileptonic B meson decay with a τ lepton in the final state. In the Standard Model (SM), the decay is mediated by a virtual W boson through a tree diagram. If there exists an unknown massive boson having an enhanced coupling to the third generation fermions, it may change the branching fraction and the kinematics of the decay. Previous studies have measured a ratio of the branching fraction for $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ to that for $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$, $R(D^*) \equiv BF(\bar{B} \to D^* \tau^- \bar{\nu}_\tau)/BF(\bar{B} \to D^* \ell^- \bar{\nu}_\ell)$, where ℓ^- denotes an electron or a muon. The current world-average $R(D^*) = 0.316 \pm 0.016(\text{stat.}) \pm 0.010(\text{syst.})$ is 3.3 standard deviation (σ) away from the SM prediction. This discrepancy may be a sign of new physics (NP) beyond the SM.

The Belle experiment has accumulated the data sample containing $(7.72 \pm 0.11) \times 10^8$ $B\bar{B}$ pairs at the asymmetric-energy e^+e^- collider KEKB. To conduct measurements of $R(D^*)$ independent of the previous studies and the τ lepton polarization $P_\tau(D^*)$, which is also a sensitive variable to NP, we reconstruct the $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ signal using the hadronic τ decays $\tau^- \to \pi^- \nu_\tau$ and $\rho^- \nu_\tau$.

There are two challenges in our analysis. First, although the rest frame of τ is needed for the $P_\tau(D^*)$ measurement, the τ momentum cannot be fully determined due to two neutrinos in the final state of $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$. Using a kinematic feature of the decay, we establish the method to reconstruct a frame equivalent to the rest frame of τ and measure $P_\tau(D^*)$ correctly. The second point is the large background originating from many types of hadronic B decay modes. Due to the complicated hadronization process, it is difficult to quantitatively predict the background yield. We therefore establish a data-driven method to correctly estimate the amount of the background by reconstructing several major background modes.

Our measurement finally results in

$$\begin{aligned} R(D^*) &= 0.270 \pm 0.035^{+0.028}_{-0.025}, \\ P_\tau(D^*) &= -0.38 \pm 0.51^{+0.21}_{-0.16}, \end{aligned}$$

where the first and the second errors are the statistical and the systematic uncertainties, respectively. The result is consistent with the SM prediction. Incorporating our $R(D^*)$ result, the precision of the world-average $R(D^*)$ becomes better by about 5%, and the discrepancy from the SM prediction slightly decreases to 3.2σ. Our $P_\tau(D^*)$ result excludes the region greater than +0.5 at 90% confidence level. This is the first experimental result for $P_\tau(D^*)$ in the decay $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$. Based on our result and the previous measurements, we investigate possibilities of the NP models containing a charged Higgs boson or a leptoquark and set new constraints on the NP parameters.

Contents

Chapter 1

Introduction

In this chapter, we describe our scientific motivation and experimental strategy for the study of the decays $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$.[1] First, the Standard Model (SM) and motivations to search for new physics (NP) beyond the SM is described. Next, the theoretical calculation for the decay amplitudes is discussed. In the third section, the previous measurements in experiments are overviewed. We then present the objective of this thesis.

1.1 The Standard Model

1.1.1 Overview of the Standard Model

Elementary particles are the most fundamental elements in the Universe. The Standard Model (SM) is the theoretical framework describing interactions of elementary particles based on the gauge invariance of the $U(1)_Y \times SU(2)_L \times SU(3)_c$ group. In the SM, there are 17 elementary particles. Table 1.1 summarizes 12 fermions: six quarks and six leptons. The fermions compose three generations, where heavier fermions belong to the higher generation, except for the neutrinos which are treated as mass-less particles. An important difference between the quarks and the leptons is their interaction; only the quark interact strongly with another quark. Table 1.2 lists gauge bosons: photon, weak boson and gluon. They intermediate the electromagnetic, the weak and the strong interactions, respectively. The Higgs boson is related to origin of the mass of the elementary particles in the electroweak symmetry breaking [1, 2]. It was the last piece of the SM particles, and was finally discovered by the ATLAS [3] and the CMS [4] experiments at the large hadron collider (LHC) [5] in 2012. As of 2016, no obvious experimental evidence which disagrees with the SM has been reported.[2]

Despite the fact that the SM is a successful theory describing a number of experimental results of the elementary particle physics, there are still several problems indicating that the SM may not be as complete as the perfect theory of the Universe. For example, as reported in Ref. [8], observations of the motion of celestial objects orbiting in a galaxy suggest the existence of dark matters (DMs). From the cosmological point of view, the DMs must have properties different from any SM particle if they are elementary particles.

[1] Throughout the thesis, charge-conjugate state is always implied.

[2] One exception is the neutrino oscillation [6, 7], which indicates a non-zero neutrino mass. However, it does not require a critical modification of the SM.

Table 1.1: Fermions in the SM. The quantities S, B, L, Q, Y and T^3 denote the spin, the baryon number, the lepton number, the electric charge with a unit of the elementary charge e, the weak hypercharge and the third component of the weak isospin, respectively.

	1-gen.	2-gen.	3-gen.	S	B	L	Q	Y	T^3
Quark	$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} c_L \\ s_L \end{pmatrix}$	$\begin{pmatrix} t_L \\ b_L \end{pmatrix}$	$\frac{1}{2}$	$\frac{1}{3}$	0	$\begin{pmatrix} +\frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}$	$+\frac{1}{6}$	$\begin{pmatrix} +\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$
	u_R	c_R	t_R	$\frac{1}{2}$	$\frac{1}{3}$	0	$+\frac{2}{3}$	$+\frac{2}{3}$	0
	d_R	s_R	b_R	$\frac{1}{2}$	$\frac{1}{3}$	0	$-\frac{1}{3}$	$-\frac{1}{3}$	0
Lepton	$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}$	$\frac{1}{2}$	0	1	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$-\frac{1}{2}$	$\begin{pmatrix} +\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$
	ν_{eR}	$\nu_{\mu R}$	$\nu_{\tau R}$	$\frac{1}{2}$	0	1	0	0	0
	e_R	μ_R	τ_R	$\frac{1}{2}$	0	1	-1	-1	0

Table 1.2: Bosons in the SM. The quantities S, Q, Y and T^3 denote the spin, the electric charge with a unit of the elementary charge e, the weak hypercharge and the third component of the weak isospin, respectively.

		S	Q	Y	T^3
Photon	γ	1	0	0	0
Weak boson	$\begin{pmatrix} W^+ \\ W^- \end{pmatrix}$	1	$\begin{pmatrix} +1 \\ -1 \end{pmatrix}$	0	$\begin{pmatrix} +1 \\ -1 \end{pmatrix}$
Weak boson	Z^0	1	0	0	0
Gluon	g	1	0	0	0
Higgs boson	$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$	0	$\begin{pmatrix} +1 \\ 0 \end{pmatrix}$	$+\frac{1}{2}$	$\begin{pmatrix} +\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$

Another example is the fact that our Universe is dominated by baryons, not by anti-baryons. The baryogenesis is not fully explained by the SM. Although the *CP*-violation described by the Kobayashi-Maskawa theory [9] is one of the sources of the baryogenesis, the amount of baryons generated is by several orders of magnitude smaller than that observed in the current Universe. To understand the nature of our Universe more deeply, it is essential to investigate NP beyond the SM.

1.1.2 Search for New Physics beyond the Standard Model

There are two types of the NP searches: direct search and indirect search. In the indirect search, quantum effects from an unknown elementary particle is measured in lower-energy phenomena compared to the mass of the new particle. In principle, the indirect search is capable of reaching the mass region much higher than initial beam energy in an experiment if an ideally large statistics of the data sample is available and physical quantities are

measured very precisely. On the other hand, the direct search investigates NP by directly producing the new elementary particle in collisions of high-energy particle beams. While the mass reach is limited by the beam energy, it is possible to find much clearer signals of the new particle compared to the indirect search. These experimental methods are complementary to each other.

A B meson, consisting of an anti-b quark and a u or a d quark and with a mass of 5.28 GeV/c^2 [10], is a good probe for indirect NP searches. With a lifetime of about 1.6 ps, it decays into various types of the final states, most of which are induced by the weak interaction. The SM can be tested by measuring decay rates and decay kinematics of the B meson precisely and comparing them with SM predictions. If an experimental result significantly deviates from the SM prediction, it indicates the existence of NP behind the phenomenon. In this thesis, we especially focus on semileptonic B meson decays, which has many advantages as a tool for testing the SM and searching for NP indirectly.

1.2 Semileptonic B Meson Decay

1.2.1 Theory of the Semileptonic B Meson Decay

Semileptonic B meson decays $\bar{B} \to X l^- \bar{\nu}_l$, where $l = e^-, \mu^-, \tau^-$ and X denotes one or more hadrons, contain a leptonic system and a hadronic transition $\bar{B} \to X$. These are the decay processes via a tree diagram mediated by the virtual W boson (denoted W^* hereafter). Among them, the exclusive decays $\bar{B} \to D^* l^- \bar{\nu}_l$, shown in Fig. 1.1 (left), have the largest branching fraction of more than 10%.

There are several interests to study $\bar{B} \to D^* l^- \bar{\nu}_l$. Firstly, it provides an environment to study low-energy QCD, which contributes to the $\bar{B} \to D^*$ transition. The b and the c quarks in the B and D^* mesons have the masses of $(4.18\ ^{+0.04}_{-0.03})$ GeV/c^2 and (1.27 $\pm$ 0.03) GeV/c^2, respectively [10]. These are significantly heavier than the QCD scale parameter $\Lambda_{QCD} \sim 200$ MeV. The QCD contribution is therefore calculated based on the heavy quark limit. The theory of the low-energy QCD is tested experimentally by measurements of $\bar{B} \to D^* l^- \bar{\nu}_l$. The experimental results are also used for determination of the theoretical parameters. Second, by estimating the hadronic contribution in $\bar{B} \to D^*$ precisely, the weak interaction via a W^* boson can be tested. In the SM, the W^* boson couples to the lepton universally over the three generation. If an unknown boson has an enhanced coupling to a specific generation, the decay rates of $\bar{B} \to D^* e^- \bar{\nu}_e$, $\bar{B} \to D^* \mu^- \bar{\nu}_\mu$ and $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ deviate from the prediction based on the lepton universality. This is a

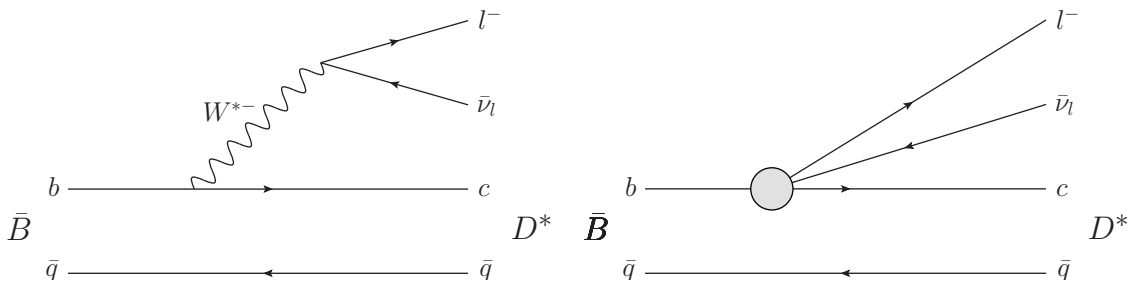


Figure 1.1: (left) First-order Feynman diagram for $\bar{B} \to D^* l^- \bar{\nu}_l$. (right) Effective interaction at low-energy scale. The quark $\bar{q}$ denotes the $\bar{u}$ or the $\bar{d}$ quark.

good test for the SM at energy scale of the electroweak symmetry breaking.

Since the W boson mass of 80.4 GeV/c^2 [10] is much heavier than the B meson mass, the decay is represented by the effective interaction with the four-fermion vertex shown in Fig. 1.1 (right). The corresponding effective Lagrangian is represented by

$$\mathcal{L} = -2\sqrt{2}G_F V_{cb}[\bar{c}\gamma_\mu(1-\gamma^5)b][\bar{l}\gamma^\mu(1-\gamma^5)\nu_l], \tag{1.1}$$

where V_{cb} denotes the Cabibbo-Kobayashi-Maskawa matrix element for the $b \to c$ transition, and

$$G_F = \frac{g_2^2}{4\sqrt{2}M_W^2} \tag{1.2}$$

is the Fermi constant. The parameters g_2 and M_W denote the gauge coupling constant in $SU(2)_L$ and the W boson mass. The fermion field $\psi = c, b, l, \nu_l$ is the Dirac spinor $\psi^T = (\psi_L, \psi_R)^T$ and $\bar{\psi} = \psi^\dagger\gamma^0$. The subscripts L and R denote the fermion chirality. Since we adopt the Weyl representation for γ^μ ($\mu = 0, 1, 2, 3$),

$$\gamma_5\psi_L = +\psi_L, \tag{1.3}$$
$$\gamma_5\psi_R = -\psi_R, \tag{1.4}$$

where $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$.

A general form of the decay amplitude for $\bar{B} \to D^* l^- \bar{\nu}_l$ is given by [11, 12]

$$\mathcal{M}_{\rm SM}^{\lambda_l,\lambda} = \frac{G_F}{\sqrt{2}} V_{cb} \sum_{\lambda_W} \eta_{\lambda_W} H_{\lambda_W}^{\lambda} L_{\lambda_W}^{\lambda_l}, \tag{1.5}$$

where H and L are the hadronic and the leptonic amplitudes, respectively. Helicities of l and W^* are represented by $\lambda_l = \pm$ and $\lambda_W = \pm, 0, s$, respectively, where s denotes the pseudo-scalar state. The states $\lambda_W = \pm, 0$ and $\lambda_W = s$ have the metric factor $\eta_{\lambda_W} = 1$ and $\eta_{\lambda_W} = -1$, respectively. The suffix λ is the helicity of the daughter meson: $\lambda = s$ for D and $\lambda = \pm, 0$ for D^*.

The leptonic amplitudes are calculated based on the spin rotation from the W^* boson to the leptonic system composed of τ^- and $\bar{\nu}_\tau$. According to the calculation in Ref. [12],

$$L_\pm^+ = \pm\sqrt{2}m_l v \sin\theta_l, \tag{1.6}$$
$$L_0^+ = 2m_l v \cos\theta_l, \tag{1.7}$$
$$L_s^+ = -2m_l v, \tag{1.8}$$
$$L_\pm^- = \sqrt{2q^2}v(1 \pm \cos\theta_l), \tag{1.9}$$
$$L_0^- = -2\sqrt{q^2}v \sin\theta_l, \tag{1.10}$$

where

$$v = \sqrt{1 - \frac{m_l^2}{q^2}}, \tag{1.11}$$

and m_l denotes the charged lepton mass. The angle θ_l is defined as the direction of the charged lepton momentum with respect to the $\bar{B}$ momentum in the rest frame of W^*. Using the four momenta of B and D^*, the square of the momentum transfer is defined as

$$q^2 \equiv (p_B - p_{D^*})^2, \tag{1.12}$$
$$= (E_B - E_{D^*})^2 - (\vec{p}_B - \vec{p}_{D^*})^2, \tag{1.13}$$

where $E_{B(D^*)}$ and $\vec{p}_{B(D^*)}$ denote the energy and the three momentum of $B(D^*)$, respectively. When the helicity states $\lambda_W = s$ and $\lambda_l = -$ are taken, the helicity of $\bar{\nu}_l$ is automatically determined to be $-$. To prohibit a right-handed neutrino or a left-handed anti-neutrino, the amplitude L_s^- vanishes.

The hadronic amplitudes are more complicated than the leptonic amplitudes due to the low-energy QCD contribution. Reference [13] reports

$$H_\pm^\pm = (m_B + m_{D^*})A_1(q^2) \mp \frac{2m_B}{m_B + m_{D^*}}|\mathbf{p}|V(q^2), \tag{1.14}$$

$$H_0^0 = \frac{1}{2m_{D^*}\sqrt{q^2}} \times \left[(m_B^2 - m_{D^*}^2 - q^2)(m_B + m_{D^*})A_1(q^2) - \frac{4m_B^2|\mathbf{p}|^2}{m_B + m_{D^*}}A_2(q^2)\right], \tag{1.15}$$

$$H_s^0 = \frac{2m_B|\mathbf{p}|}{\sqrt{q^2}}A_0(q^2), \tag{1.16}$$

where

$$|\mathbf{p}| = \frac{1}{2m_B}\sqrt{m_B^4 + m_{D^*}^4 + q^4 - 2(m_B^2 m_{D^*}^2 + m_{D^*}^2 q^2 + q^2 m_B^2)}, \tag{1.17}$$

and $m_{B(D^*)}$ denotes the $B(D^*)$ meson mass. The hadronic form factors (FFs) $A_0(q^2)$, $A_1(q^2)$, $A_2(q^2)$ and $V(q^2)$ describe the hadronic interaction in the $B \to D^*$ transition; $V(q^2)$ originates from the vector term (first term) and the remaining from the axial-vector term (second term) in the hadronic matrix elements

$$\langle D^*|\gamma^\mu(1-\gamma^5)|B\rangle = \langle D^*|\gamma^\mu|B\rangle - \langle D^*|\gamma^\mu\gamma^5|B\rangle\,. \tag{1.18}$$

These FFs are associated by the relations [13]

$$h_{A_1}(w) = A_1(q^2)\frac{1}{R_{D^*}}\frac{2}{w+1}, \tag{1.19}$$

$$A_0(q^2) = \frac{R_0(w)}{R_{D^*}}h_{A_1}(w), \tag{1.20}$$

$$A_2(q^2) = \frac{R_2(w)}{R_{D^*}}h_{A_1}(w), \tag{1.21}$$

$$V(q^2) = \frac{R_1(w)}{R_{D^*}}h_{A_1}(w), \tag{1.22}$$

where

$$w = \frac{m_B^2 + m_{D^*}^2 - q^2}{2m_B m_{D^*}}, \tag{1.23}$$

$$R_{D^*} = \frac{2\sqrt{m_B m_{D^*}}}{m_B + m_{D^*}}. \tag{1.24}$$

The quantities $R_0(w)$, $R_1(w)$ and $R_2(w)$ are essentially the ratios of the FFs $A_0(q^2)$, $V(q^2)$ and $A_2(q^2)$, respectively, to $A_1(q^2)$. By substituting Eqs. 1.19 to 1.22 for Eqs. 1.14 to 1.16, the helicity amplitudes are expressed by the products of $h_{A_1}(w)$ and the other parts.

Theoretically, it is the simplest to calculate dynamics inside the hadrons at $w = 1$, which is the minimum value of w corresponding to the maximum q^2. At this condition, all the energy taken from the mass difference between B and D^* is transferred to the leptonic system, and D^* also becomes static in the rest frame of $\bar{B}$. The heavy quarks b and c are thus approximately treated as static quarks. Expanding the FFs around $w = 1$ and adopting theoretical constraints from the heavy quark effective theory (HQET), $h_{A_1}(W)$, $R_1(w)$, $R_2(w)$ and $R_0(w)$ are represented as [13, 14]

$$\begin{aligned} h_{A_1}(w) &= h_{A_1}(1)\left[1 - 8\rho^2 z + (53\rho^2 - 15)z^2 - (231\rho^2 - 91)z^3\right], && (1.25)\\ R_1(w) &= R_1(1) - 0.12(w-1) + 0.05(w-1)^2, && (1.26)\\ R_2(w) &= R_2(1) + 0.11(w-1) - 0.06(w-1)^2, && (1.27)\\ R_0(w) &= R_0(1) - 0.11(w-1) + 0.01(w-1)^2, && (1.28) \end{aligned}$$

where

$$z = \frac{\sqrt{w+1} - \sqrt{2}}{\sqrt{w+1} + \sqrt{2}}. \quad (1.29)$$

Here, ρ^2 is a coefficient for the first-order term of the Taylor expansion of $h_{A_1}(w)$ around $w = 1$. In Eq. 1.25, ρ^2 determines the dependence of $h_{A_1}(w)$ on w. The most recent lattice-QCD calculation gives [15][3]

$$h_{A_1}(1) = 0.906 \pm 0.004(\text{stat.}) \pm 0.012(\text{syst.}). \quad (1.30)$$

It should be noted that $h_{A_1}(1)$ only scales the overall factor of the hadronic amplitudes and never changes decay kinematics. We use the values of ρ^2, $R_1(1)$ and $R_2(1)$ extracted from measurements of the kinematic distributions in $\bar{B} \to D^* e^- \bar{\nu}_e$ and $\bar{B} \to D^* \mu^- \bar{\nu}_\mu$. The heavy-flavor averaging group (HFAG) provides the world averages [17]

$$\begin{aligned} \rho^2 &= 1.207 \pm 0.026, && (1.31)\\ R_1(1) &= 1.406 \pm 0.033, && (1.32)\\ R_2(1) &= 0.853 \pm 0.020. && (1.33) \end{aligned}$$

The parameter $R_0(1)$ appears in the hadronic amplitude H_s^0, which composes a product with L_s^+. Since, as shown in Eq. 1.8, L_s^+ in the $\bar{B} \to D^* e^- \bar{\nu}_e$ and $\bar{B} \to D^* \mu^- \bar{\nu}_\mu$ modes is suppressed to be $L_s^+ \sim 0$ due to the light mass of e^- and μ^-, $R_0(1)$ cannot be extracted in the same way as the other parameters. We therefore use the theoretical estimate

$$R_0(1) = 1.22, \quad (1.34)$$

based on HQET [13], with about 10% uncertainty.

Using the amplitudes and the FFs shown in this section, decay kinematics of the $\bar{B} \to D^* l^- \bar{\nu}_l$ decay is fully described.

[3] Here, equality of $\mathcal{F}(1)$ in Ref. [15] and $h_{A_1}(1)$ is used [16].

1.2.2 Observables

While the masses of e^- and μ^- are negligibly light compared to the mass of B and D^*, the τ lepton mass of 1.78 GeV/c^2 is comparable to the masses of these mesons. This difference characterizes the decay $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ differently from $\bar{B} \to D^* e^- \bar{\nu}_e$ and $\bar{B} \to D^* \mu^- \bar{\nu}_\mu$. Hereafter, we refer to the light charged leptons e^- and μ^- by ℓ^-. Using the ratio of the branching fractions,

$$R(D^*) \equiv \frac{BF(\bar{B} \to D^* \tau^- \bar{\nu}_\tau)}{BF(\bar{B} \to D^* \ell^- \bar{\nu}_\ell)}, \tag{1.35}$$

it is possible to test the lepton universality precisely. Throughout this thesis, the decay modes $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ and $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$ are referred to as the signal mode and the normalization mode, respectively. The ratio cancels some common factors such as $|V_{cb}|$ and $h_{A_1}(w)$, and thus reduces theoretical uncertainties. It also reduces experimental uncertainties such as the detector efficiencies. The SM predicts $R(D^*) = 0.252 \pm 0.003$ [13]. Since, the phase space for $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ is suppressed by the large mass of the τ lepton, the value of $R(D^*)$ is therefore significantly smaller than 1 even in the SM prediction.

The SM can be also tested using the decay kinematics. Figure 1.2 shows the full decay kinematics of $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$, where the angles θ_V, θ_τ and $\theta_{\rm hel}$ are the helicity angles of D^*, W^{*-} and τ^-, respectively. These are defined as the directions of the D momentum with respect to the direction opposite to the $\bar{B}$ momentum in the rest frame of D^*, the τ momentum with respect to the direction opposite to the $\bar{B}$ momentum in the rest frame of W^{*-}, and the τ-daughter momentum with respect to the direction opposite to the W^* momentum in the rest frame of τ, respectively. The angle between the decay plane for $W^{*-} \to \tau^- \bar{\nu}_\tau$ and that for $D^* \to D(\pi, \gamma)$ in the rest frame of $\bar{B}$ is represented by χ. A longitudinal polarization of the τ lepton,[4]

$$P_\tau(D^*) \equiv \frac{\Gamma^+(D^*) - \Gamma^-(D^*)}{\Gamma^+(D^*) + \Gamma^-(D^*)}, \tag{1.36}$$

where $\Gamma^{+(-)}(D^*)$ denotes the decay rate of $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ with the positive (negative) helicity of the τ lepton, is a sensitive probe for the NP coupling to the τ lepton. The SM predicts $P_\tau(D^*) = -0.497 \pm 0.013$ [11]. As described later in Chapter 3, $P_\tau(D^*)$ is extracted from the distribution of $\cos\theta_{\rm hel}$.

[4]The longitudinal polarization is the frame-dependent quantity. We always discuss the value of $P_\tau(D^*)$ in the rest frame of W^* in this thesis.

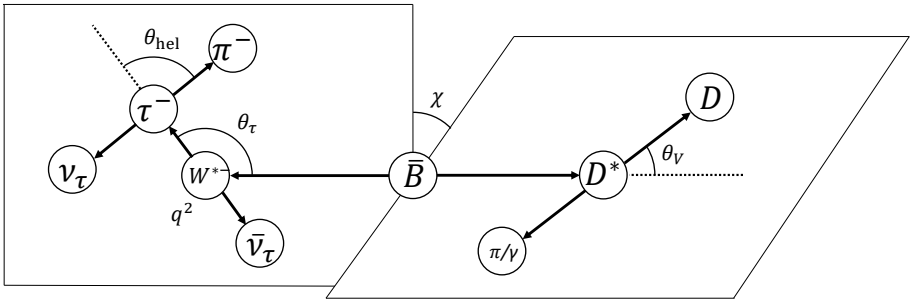


Figure 1.2: Kinematic variables characterizing the decay topology of $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$.

1.2.3 NP Effects on $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$

If there exists an unknown boson preferentially coupling to the τ lepton, it interferes with the SM and modifies the decay amplitude for $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$. A model-independent theoretical study [11] extends the the effective Lagrangian in Eq. 1.1 as

$$\mathcal{L}_{\text{eff}} = -2\sqrt{2} G_F V_{cb} \left[\mathcal{O}_{V_1} + \sum_{i=V_1,V_2,S_1,S_2,T} C_i \mathcal{O}_i \right]. \tag{1.37}$$

where

$$\mathcal{O}_{V_1} = \bar{c}\gamma^\mu(1-\gamma^5)b\bar{\tau}\gamma_\mu(1-\gamma^5)\nu_\tau, \tag{1.38}$$
$$\mathcal{O}_{V_2} = \bar{c}\gamma^\mu(1+\gamma^5)b\bar{\tau}\gamma_\mu(1-\gamma^5)\nu_\tau, \tag{1.39}$$
$$\mathcal{O}_{S_1} = \bar{c}(1-\gamma^5)b\bar{\tau}(1-\gamma^5)\nu_\tau, \tag{1.40}$$
$$\mathcal{O}_{S_2} = \bar{c}(1+\gamma^5)b\bar{\tau}(1-\gamma^5)\nu_\tau, \tag{1.41}$$
$$\mathcal{O}_T = \bar{c}\sigma^{\mu\nu}(1-\gamma^5)b\bar{\tau}\sigma_{\mu\nu}(1-\gamma^5)\nu_\tau. \tag{1.42}$$

and $\sigma^{\mu\nu} = (i/2)(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)$. The subscripts V, S and T denote the vector-, the scalar- and the tensor-type currents, respectively. For the vector- and the scalar-type operators, the index 1 and 2 indicate the difference of the quark chiralities. The right-handed neutrino is not considered, and the tensor current with the opposite combination of the quark chiralities vanishes. The complex coefficients C_i (Wilson coefficients) express the strength of the NP currents with respect to the SM.

The NP currents modify $R(D^*)$ and $P_\tau(D^*)$. Figure 1.3 illustrates the possible regions of $R(D^*)$ and $P_\tau(D^*)$, where only one of the Wilson coefficients in Eq. 1.37 has a non-zero

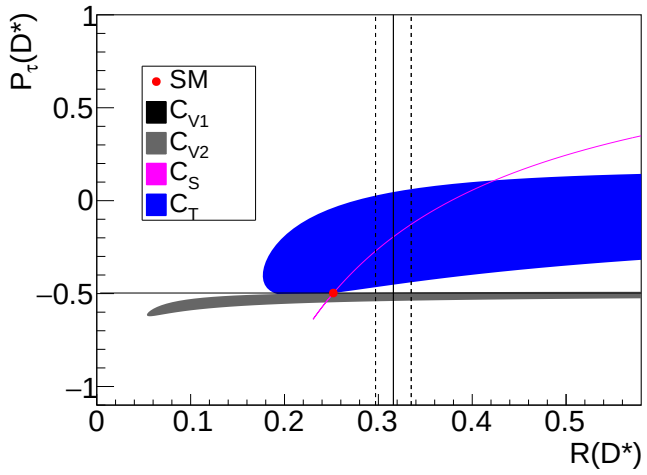


Figure 1.3: Correlation between $R(D^*)$ and $P_\tau(D^*)$ in $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ [11]. The horizontal black line, gray region, magenta line and blue region show the allowed regions with assuming that only one of the Wilson coefficients C_{V_1}, C_{V_2}, $C_{S_{1,2}}$ and C_T, respectively, has a non-zero value. The SM prediction is shown by the red dot. The experimental $R(D^*)$ average, which is described later in Sec. 1.3, is shown by the vertical black line with the $\pm 1\sigma$ region.

value [11]. The value of $P_\tau(D^*)$ is sensitive to non-vector type NP operators. A combined measurement of $R(D^*)$ and $P_\tau(D^*)$ is capable of discriminating the NP types.

In this thesis, we consider the following two NP models as benchmark models.

Two Higgs Doublet Model (2HDM)

In the SM, only one complex Higgs doublet is introduced. However, the number of doublets is not required to be exactly one. A minimal extension of the Higgs sector with one more Higgs doublet is called the Two-Higgs-Doublet model (2HDM) [18]. In the electroweak symmetry breaking, five physical Higgs bosons appear in total: h^0, H^0, A^0, H^+ and H^-, where h^0 and H^0 are the scalar Higgs bosons, A^0 is the pseudo-scalar Higgs boson and $H^\pm$ is the charged Higgs boson. Since the SM Higgs boson is electrically neutral, discovery of the charged Higgs boson is a clear evidence of NP.

In the general 2HDM, the flavor-changing neutral current (FCNC) is introduced at tree level. This is because, if two Higgs doublets Φ_1 and Φ_2 couples to the same fermions, their Yukawa matrices are not diagonal in general. According to many experimental measurements, the tree-level FCNCs are strongly suppressed. To prevent them, all the fermions in each group, up-type quark, down-type quark or charged lepton, need to couple with the same Higgs doublet. This requires the Lagrangian for the Yukawa interaction to be

$$\mathcal{L}_Y = Y^d \bar{Q}_L \Phi_i d_R + Y^u \bar{Q}_L \Phi_j u_R + Y^\ell \bar{L}_L \Phi_k e_R \quad (i,j,k=1,2), \tag{1.43}$$

where Y^d, Y^u and Y^ℓ are the Yukawa matrices for the down-type quarks, the up-type quarks and the charged leptons, respectively. The 2HDM theories are grouped as shown in Table 1.3. Among the four types of the 2HDM, the type-II 2HDM has a connection to the minimum-supersymmetric standard model (MSSM) [19], which is a prominent candidate to solve many problems in the SM, including the DM candidates [20]. To conserve the symmetry between fermions and bosons (supersymmetry) in the Lagrangian, the theory needs to have the same structure of the Higgs sector as the type-II 2HDM.

The charged Higgs may contribute to the decay $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ [11, 23, 24, 25, 26]. Two Wilson coefficients appear as [11]

$$C_{S_1} = -m_b m_\tau \frac{\tan^2 \beta}{m^2_{H^\pm}}, \tag{1.44}$$

$$C_{S_2} = -\frac{m_c m_\tau}{m^2_{H^\pm}}, \tag{1.45}$$

Table 1.3: Higgs doublet coupling with fermions in different types of the 2HDM.

Model	u-type quark	d-type quark	charged lepton
Type-I	Φ_2	Φ_2	Φ_2
Type-II	Φ_2	Φ_1	Φ_1
Type-X	Φ_2	Φ_2	Φ_1
Type-Y	Φ_2	Φ_1	Φ_2

where m_b and m_c denote the masses of the b and the c quarks, respectively. The parameter $\tan\beta$ is the ratio of the vacuum expectation values of two Higgs doublets, and the region $1 < \tan\beta < 60$ is favored in terms of the strength of the Yukawa couplings to the t and the b quarks. The charged Higgs mass is denoted by $m_{H^\pm}$. The direct measurements at the Large Electron-Positron collider using e^+e^- collisions have excluded $m_{H^\pm} < 80$ GeV/c^2 at 95% confidence level (C.L.) at any values of $\tan\beta$ [27], and hence the charged Higgs mass is much greater than m_c and m_τ. The coefficient C_{S_2} is therefore negligibly small.

Another B meson decay sensitive to the charged Higgs is the radiative decays $B \to X_s\gamma$, where X_s denotes any hadronic final state containing an s quark. In these decay modes, the charged Higgs boson appears in a loop of the decay diagram. Currently, $m_{H^\pm} < 480$ GeV/c^2 [28] or $m_{H^\pm} < 493$ GeV/c^2 [29] is excluded at 95% C.L. Compared to $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$, the $B \to X_s\gamma$ process is less sensitive to $\tan\beta$.

The charged Higgs production is also directly investigated using collisions of high energy protons at LHC. Figure 1.4 shows the latest results from the ATLAS and the CMS experiments on the plane for $m_{H^\pm}$ and $\tan\beta$ [21, 22].

Scalar Leptoquark

The scalar leptoquark is an elementary particle carrying both a quark number and a lepton number. It appears in various extensions of the SM for unification of the three SM interactions or the quark and the lepton sectors [30]. The contribution of the leptoquark to the $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decay is phenomenologically studied [11, 31, 32, 33, 34, 35, 36], although its particular connection to high energy theories has not been established yet.

There are three types of the scalar leptoquarks satisfying the following assumption.

- Interaction with the SM fermions is dimensionless and invariant under the SM gauge transformation.

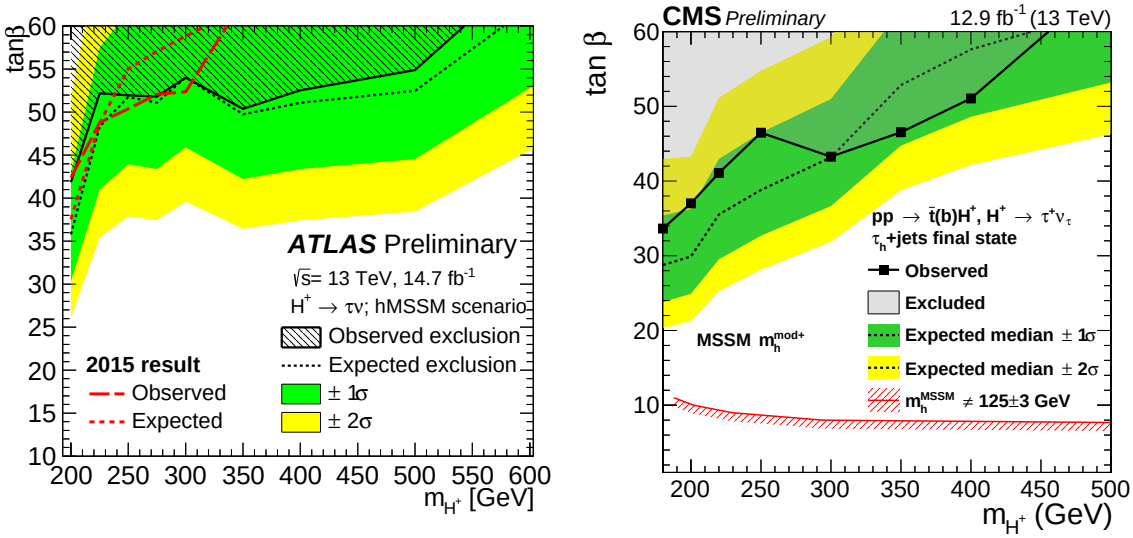


Figure 1.4: Latest results of the direct charged Higgs boson searches at ATLAS (left) [21] and CMS (right) [22] with $H^+ \to \tau^+\nu_\tau$ followed by hadronic τ decays at LHC 13-TeV Run. In both measurements, the hMSSM scenario is assumed, where the 125-GeV Higgs boson is regarded as the lighter scalar Higgs boson h^0.

- It is possible to mediate the process $b \to c\tau^-\bar{\nu}_\tau$ without introducing a right-handed neutrino.

The interaction between a scalar leptoquark, a quark and a lepton is represented by the product of the quark and the lepton fields. Using the Fierz transformation, it is reformed as a linear sum of the operators in Eqs. 1.38 to 1.42. In this thesis, we focus on the two leptoquarks which have the non-V_1 operators [33]. Their quantum numbers are summarized in Table 1.4, and corresponding Feynman diagrams are shown in Fig. 1.5. Following Ref. [33], we call them R_2 model and S_1 model, respectively, and assume $m_{LQ} = 1\ \mathrm{TeV}/c^2$. According to the direct search for the scalar leptoquark coupling to the b quark and the τ lepton at CMS, the current lower mass limit is $m_{LQ} = 900\ \mathrm{GeV}/c^2$ at 90% C.L. [37], and therefore the mass of $1\ \mathrm{TeV}/c^2$ is allowed. In these leptoquark models, two Wilson coefficients appear with the relation [33]

$$C_{S_2} = \pm 7.8 C_T, \tag{1.46}$$

where the positive (negative) sign is for the R_2 (S_1) model.

1.3 Previous Experimental Measurements

A measurement of $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ is experimentally challenging since there are always two or more neutrinos in the final state: one directly from the B meson decay and the remaining from the τ lepton decay.

At the B-factory experiments such as Belle [38] and BaBar [39], an electron and a positron are collided at the center-of-mass (CM) energy of 10.58 GeV. This energy is

Table 1.4: Quantum numbers of the leptoquarks in the R_2 and the S_2 models. The field with a superscript c denotes the charge-conjugated fermion field.

	R_2 leptoquark	S_1 leptoquark
Spin	0	0
$3B + L$	0	-2
Y	7/6	1/3
T^3	$\pm 1/2$	0
Coupling	$\bar{b}_L\tau_R$, $\bar{c}_R\nu_{\tau L}$	$\bar{b}^c_L\nu_{\tau L}$, $\bar{c}^c_L\tau_L$

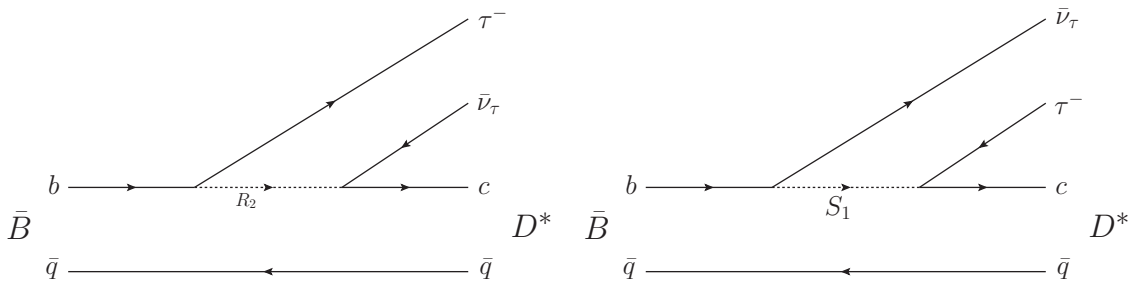


Figure 1.5: Diagrams of the $\bar{B} \to D^*\tau^-\bar{\nu}$ decay mediated by the R_2 leptoquark (left) or the S_1 leptoquark (right).

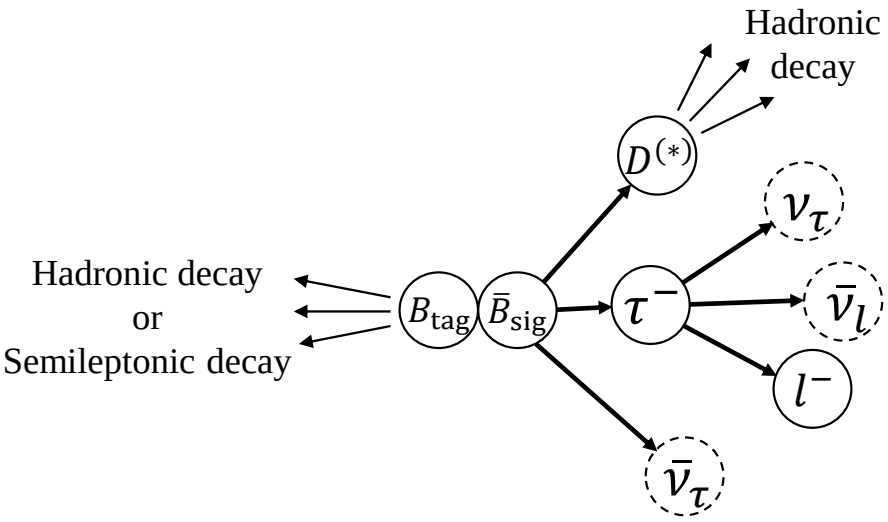


Figure 1.6: Illustration of the two B meson decays at an e^+e^- collision.

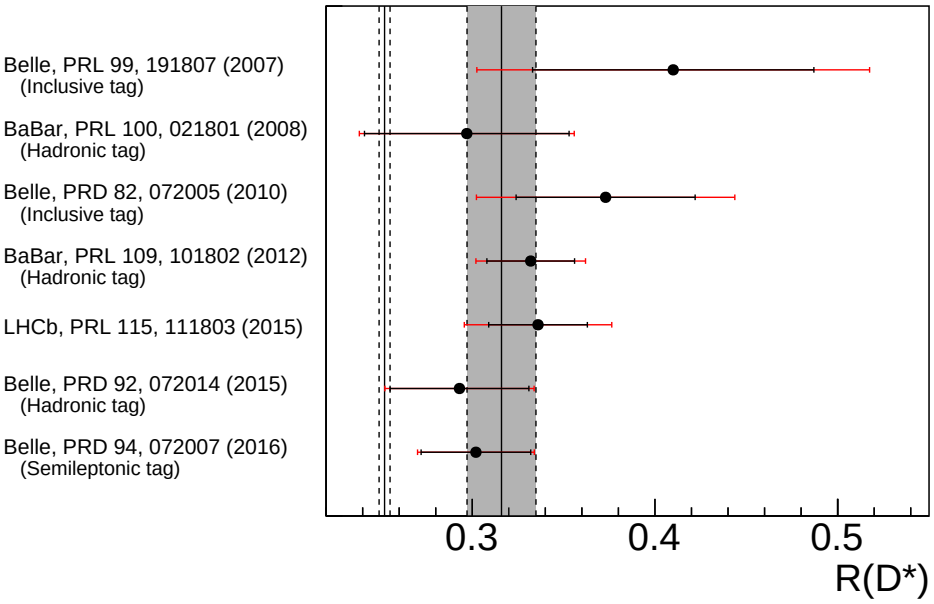


Figure 1.7: Summary of the previous $R(D^*)$ measurements. The red error bars show the total uncertainties while the black error bars are statistical. The shaded and the blank vertical lines show the world-average $R(D^*)$ [17] and the SM prediction [13], respectively.

consistent with the mass of $\Upsilon(4S)$, which is a resonance of a pair of b and $\bar{b}$. The $\Upsilon(4S)$ decays into a $B\bar{B}$ pair with more than 96% branching fraction [10]. In the $\Upsilon(4S)$ decay, no particle except for the two B mesons is produced as illustrated in Fig. 1.6. This feature provides an advantage for $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ measurements. By identifying one of the two B mesons (denoted as the tag-side B meson or $B_{\rm tag}$), information of the signal-side B meson ($B_{\rm sig}$) such as a four-momentum and associated particles is indirectly obtained without explicitly reconstructing $B_{\rm sig}$.

Figure 1.7 is a summary of the previous experimental studies of $R(D^*)$. The decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ was firstly observed by Belle in 2007 [40]. Since then, this decay has been studied at Belle [41, 42, 43] and BaBar [44, 45]. In 2012, BaBar reported a 2.7σ discrepancy in their measurement of $R(D^*)$ from the SM prediction. The report also included a measurement of $R(D) \equiv BF(\bar{B} \to D\tau^-\bar{\nu}_\tau)/BF(\bar{B} \to D\ell^-\bar{\nu}_\ell)$ for the similar decay mode $\bar{B} \to D\tau^-\bar{\nu}_\tau$. This result was by 2.0σ away from the SM prediction $R(D) = 0.300 \pm 0.008$ [46] or $R(D) = 0.299 \pm 0.011$ [47]. The overall discrepancy for $R(D^*)$ and $R(D)$ reached 3.4σ. While these measurements at Belle and BaBar reconstructed $B_{\rm tag}$ from its hadronic decay, in the latest $R(D^*)$ measurement at Belle, the semileptonic decay was adopted to the $B_{\rm tag}$ reconstruction. Despite the disadvantage that $B_{\rm tag}$ is not fully reconstructed due to one neutrino in the $B_{\rm tag}$ decay, this measurement achieved a comparable precision to the past $R(D^*)$ measurements. LHCb [48] also measured $R(D^*)$ and demonstrated a capability of the $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ study at hadron colliders [49].

The world-average $R(D^{(*)})$ with measurements in Refs. [42, 43, 45, 49] are estimated by HFAG as [17]

$$R(D) = 0.397 \pm 0.040(\text{stat.}) \pm 0.028(\text{syst.}), \tag{1.47}$$

$$R(D^*) = 0.316 \pm 0.016(\text{stat.}) \pm 0.010(\text{syst.}). \tag{1.48}$$

Here, two $R(D^*)$ measurements at Belle [42, 43] are used since they exploited different B decay modes for the $B_{\rm tag}$ reconstruction and therefore statistically independent. The other measurements with a partial data set of Belle [40, 41] and that of BaBar [44] are excluded to avoid statistical overlap. These values are 1.9σ and 3.3σ away from the SM prediction, respectively. The overall discrepancy is about 4σ.

In the past measurements, the leptonic τ decays $\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau$ are mostly used. The existence of one charged lepton in the final state of the $B_{\rm sig}$ decay is one of the advantages for the event selection. On the other hand, it requires a difficult estimation for the $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ background, where D^{**} denotes excited states of the charmed meson heavier than D^*. If π^0 from the $D^{**} \to D^*\pi^0$ decay is not detected, $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ events mimic the signal $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$. Since some of the D^{**} states have wide mass widths around 300 MeV/c^2 [10], experimental measurements for $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ are not trivial. Besides, there are several theoretically-predicted D^{**} states, which have not been measured yet [50]. This background mode has caused a major systematic uncertainty in the past measurements.

1.4 Objective of This Thesis

In this thesis, using the full data sample containing 7.72×10^8 $B\bar{B}$ pairs accumulated at the Belle experiment, I aim to

- investigate the $R(D^*)$ discrepancy with an independent measurement, and
- add a new observable

to investigate the SM and possibilities of the NP. I reconstruct signal events using hadronic τ decays $\tau^- \to \pi^-\nu_\tau$ and $\rho^-\nu_\tau$. Since the previous measurement at Belle [42] has used only the leptonic τ decays, an independent data sample is obtained. Besides, the main background in this analysis is predicted to originate from hadronic B decays, which is

different from $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ in the previous analyses. The $R(D^*)$ discrepancy is investigated with the different background-induced systematic uncertainty. On the other hand, estimation of the hadronic B decay background is a challenging point of my analysis since hadronic B decays are too complicated to be quantitatively predicted theoretically, and experimental measurements do not fully cover all the possible hadronic B decay modes.

The polarization $P_\tau(D^*)$ has a sensitivity to NP but has not been measured yet. It was difficult to measure $P_\tau(D^*)$ using the leptonic τ decays, where the two neutrinos from the τ decay weaken information of the polarization. My analysis using the two-body τ decays enables to measure $P_\tau(D^*)$ by using the feature of the two-body τ decays. Nevertheless, the $P_\tau(D^*)$ measurement is challenging as the τ lepton is not fully reconstructed due to a neutrino. I have developed a method to extract $P_\tau(D^*)$ from data using kinematic constraints of the $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decay.

Summarizing above, the goal of this thesis is to conduct

- a new $R(D^*)$ measurement using $\tau^- \to \pi^-\nu_\tau$ and $\rho^-\nu_\tau$, and
- the first $P_\tau(D^*)$ measurement

using the full data sample at Belle.

Chapter 2

Experimental Apparatus

In this study, we use the data sample of e^+e^- collisions at the KEKB accelerator [51]. The data are recorded with the Belle detector [38]. First, the KEKB accelerator and the Belle detector are overviewed. The particle reconstruction and identification methods are next discussed.

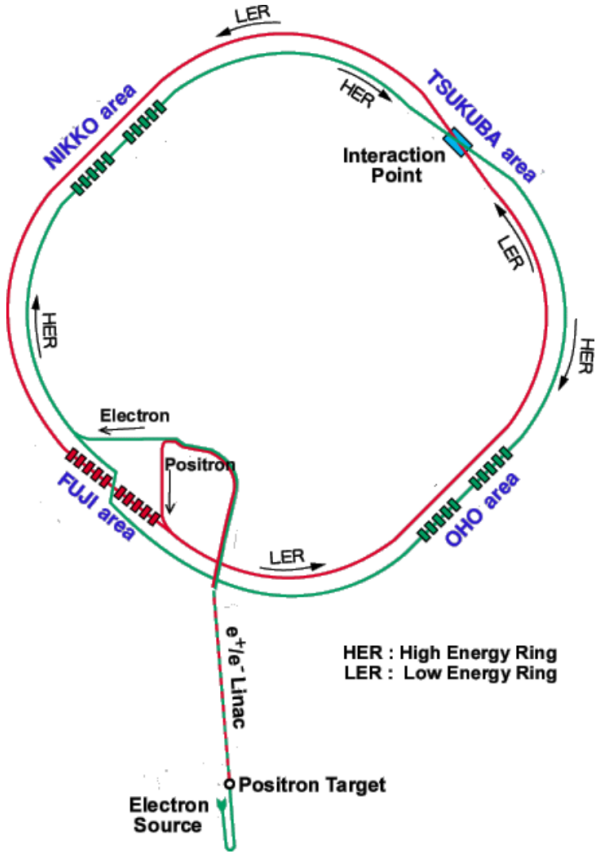


Figure 2.1: Drawing of the KEKB accelerator. The HER and the LER are the e^- and the e^+ beams, respectively. Four experimental halls, FUJI, NIKKO, OHO and TSUKUBA, are shown.

2.1 KEKB Accelerator

The KEKB accelerator is an asymmetric-energy e^+e^- collider located at KEK, Japan. As shown in Fig. 2.1, the accelerator consists of a linear accelerator (Linac) and a 3-km-circumference main ring (MR). Electrons produced by a thermal electron gun are accelerated to 8 GeV in the Linac and injected to the MR. A part of the electrons are collided with the tungsten target in the middle of the Linac to produce positrons. The positrons are collected, accelerated to 3.5 GeV and injected to the MR. In the MR, the e^+ and the e^- beams are stored with 508.88 MHz RF cavities. They are collided at the interaction point (IP) in the TSUKUBA experimental hall, where the Belle detector is located, with a crossing angle of 22 mrad. Several different states of the e^+e^- center-of-mass (CM) energy $\sqrt{s}$ from 9.4 GeV to 11.2 GeV are selected for the collision.

The data used for our study is taken at $\sqrt{s} = 10.58$ GeV, corresponding to the mass of the $\Upsilon(4S)$ resonance. A pair of B mesons is produced in the decay of $\Upsilon(4S)$. Decay particles from the B mesons are detected by the Belle detector, surrounding the IP. As mentioned in Sec. 1.3, there is no additional particle produced in an e^+e^- collision. This feature is essential for physics analyses with a large missing momentum. Other physical processes at this energy, such as $e^+e^- \to \tau^+\tau^-$, $q\bar{q}$ ($q = u, d, s, c$) and the processes by the quantum electrodynamics (QED) are summarized in Table 2.1.

KEKB has been operated from 1999 to 2010. It achieved the world-record instantaneous luminosity of 2.11×10^{34} cm^{-2}s^{-1} and the total integrated luminosity of 1041 fb^{-1}. Out of 1041 fb^{-1}, 711 fb^{-1} data were taken at $\sqrt{s} = 10.58$ GeV. The data set contains 7.71×10^8 $B\bar{B}$ pairs.

2.2 Belle Detector

As shown in Fig. 2.2, the Belle detector is a complex of six different sub-detector systems. The silicon vertex detector (SVD) and the central drift chamber (CDC) are the tracking devices for charged particles. The time-of-flight (TOF) counter and the aerogel Cherenkov counter (ACC) provide particle identification (PID) information of charged hadrons and electrons. The electromagnetic calorimeter (ECL) measures energy of electrons and photons. The K_L^0 Muon detector (KLM) is dedicated to detecting muons and neutral hadrons. The extreme forward calorimeter (EFC) is used as the instantaneous luminosity monitor and the triggering device for Bhabha scattering events (Since the EFC

Table 2.1: Cross sections of the main physics processes in e^+e^- collisions at $\sqrt{s} = 10.58$ GeV [38].

$e^+e^- \to b\bar{b}$	1.2 nb
$e^+e^- \to q\bar{q}$	2.8 nb
$e^+e^- \to \mu^+\mu^-, \tau^+\tau^-$	1.6 nb
Bhabha scattering (within detector acceptance)	44 nb
Other QED process (within detector acceptance)	$\sim$ 17 nb
Total	$\sim$67 nb

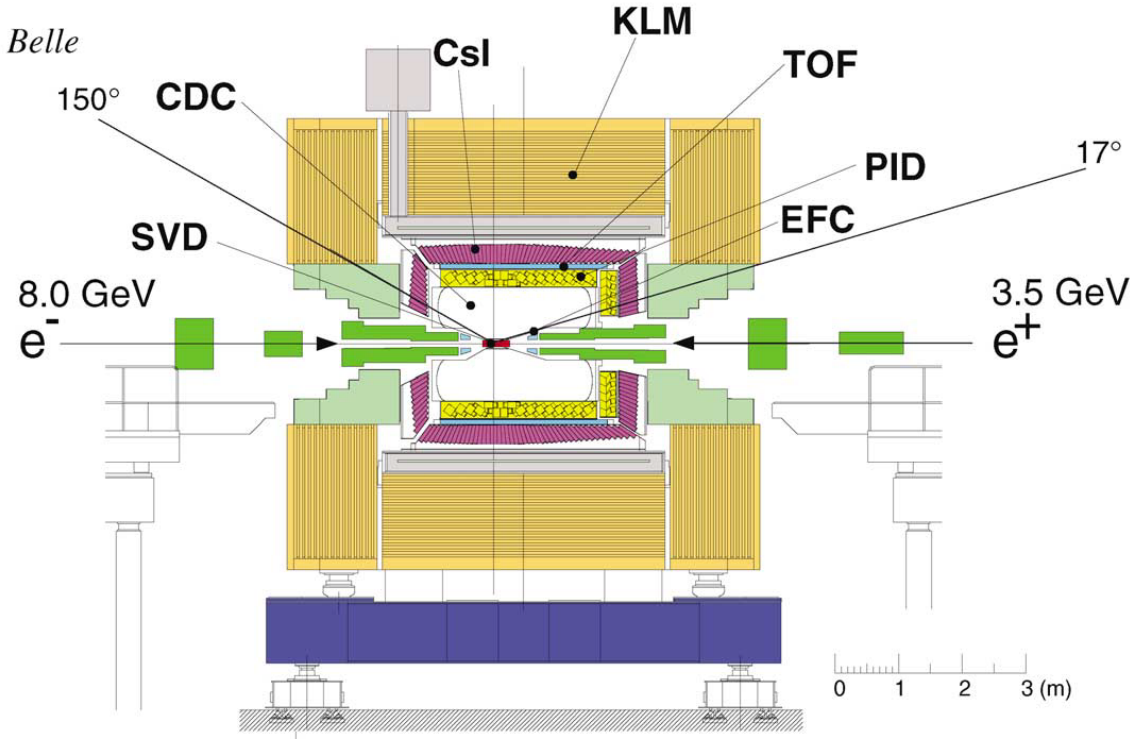


Figure 2.2: Cross section of the Belle detector [38]. CsI and PID indicate the positions of the ECL and the ACC, respectively.

is not directly used in our analysis, we omit description for the EFC; see, for example, Section 3 in Ref. [38] for details). To determine the momentum of charged particles, the inner sub-detectors (SVD, CDC, TOF, ACC and ECL) and EFC are immersed in the 1.5 T magnetic field supplied by the superconducting solenoid.

The coordinates of the Belle detector are the left-handed system with the origin at the IP and the z axis corresponding to the direction opposite to the e^+ beam; the polar and cylindrical angles are represented by θ and ϕ, respectively; the radial distance by r.

In our analysis, key features of the Belle detector are as follows.

- Good particle detection efficiency for 4-π region to collect all the particles in an e^+e^- collision. This is important not only for the good signal reconstruction efficiency, but also for background reduction since un-detected particles from various B decay modes are the main source of the background.

- Good reconstruction efficiency for low-momentum particles. Since decay particles from D^* mesons and τ leptons are produced in a chain of multi-body decays, their momenta are about 0.7 GeV/c on average and lower than 1.5 GeV/c.

- Good momentum resolution in the tracking devices and energy resolution in the ECL.

- Good PID performance for charged particles.

2.2.1 Silicon Vertex Detector (SVD)

For the SVD, which is the innermost sub-detector system, about 100 μm position resolution is required to determine vertices of the two B mesons in the measurement of the time-dependent CP-violation [52]. To achieve such an excellent position resolution, double-sided silicon detector (DSSD) was developed.

In the beginning of the experiment, the three-layer SVD (SVD1) was installed, and the data containing 1.52×10^8 $B\bar{B}$ pairs were recorded. The SVD1 was upgraded to the four-layer SVD system (SVD2), and 6.20×10^8 $B\bar{B}$ pairs were accumulated. Figure 2.3 shows the structure of the SVD2. The DSSD has a size of about $8 \times 3(5 \times 3)$ cm^2, where the value without (with) the parentheses is the specification of the SVD2 (SVD1). Many sensor strips are placed on the both sides of the plate: one side with 50–55(42) μm width for the z direction and the other side with 10–12(25) μm width for the ϕ direction. Two to six (two to four) DSSD plates compose a "ladder", covering the polar angle $17(23)^\circ < \theta < 150(140)^\circ$. Cylindrical four (three) layers with 8 to 18 (8 to 14) ladders surround the beam pipe with the inner radius of 20(30) mm. Passing through the DSSD by about 300 μm, a minimum-ionization particle creates about 2×10^5 electrons. The output charge is shaped, amplified and digitized by a flash analog-to-digital converter.

For the signal processing in the SVD1, a VA1 chip [54] was used. However, a gain of the VA1 chip was degraded due to damage by γ rays. As shown in Fig. 2.4 (left), it caused a slight efficiency drop by 1%. In the SVD2, the VA1 chip was replaced by the VA1TA chip with a 0.35 μm process [53]. The thinner process than 1.2 μm for VA1 achieved a better tolerance to γ rays.

Figure 2.4 (right) shows the impact parameter resolution evaluated using cosmic-ray tracks. The impact parameter is defined as the closest approach of a charged particle to the IP. In our study, its resolution is important to select good-quality charged tracks originating from the e^+e^- collision. Both for the r direction (dr) and the z direction (dz), the resolution better than 150 μm is achieved for more than 500 MeV/c particles.

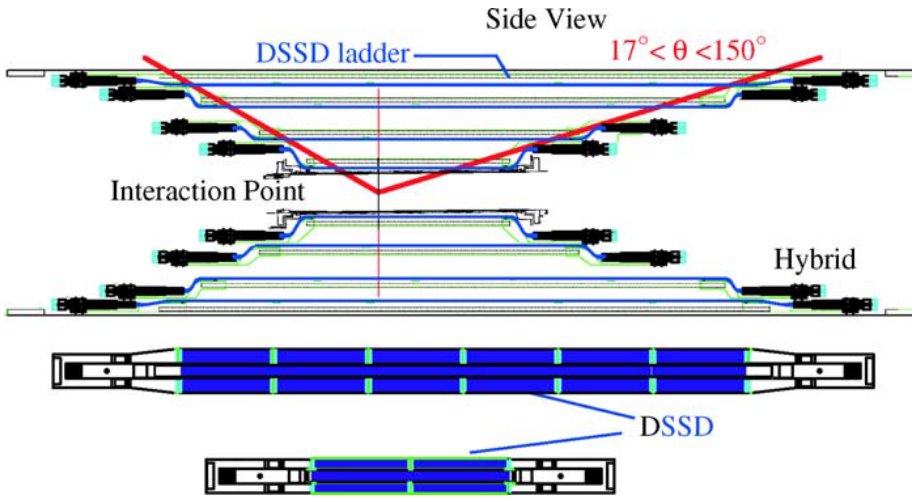


Figure 2.3: Structure of the SVD2 [53]. The top drawing is the cross section on the $y-z$ plane, and the bottom shows the structure of ladders for the first and the forth layers.

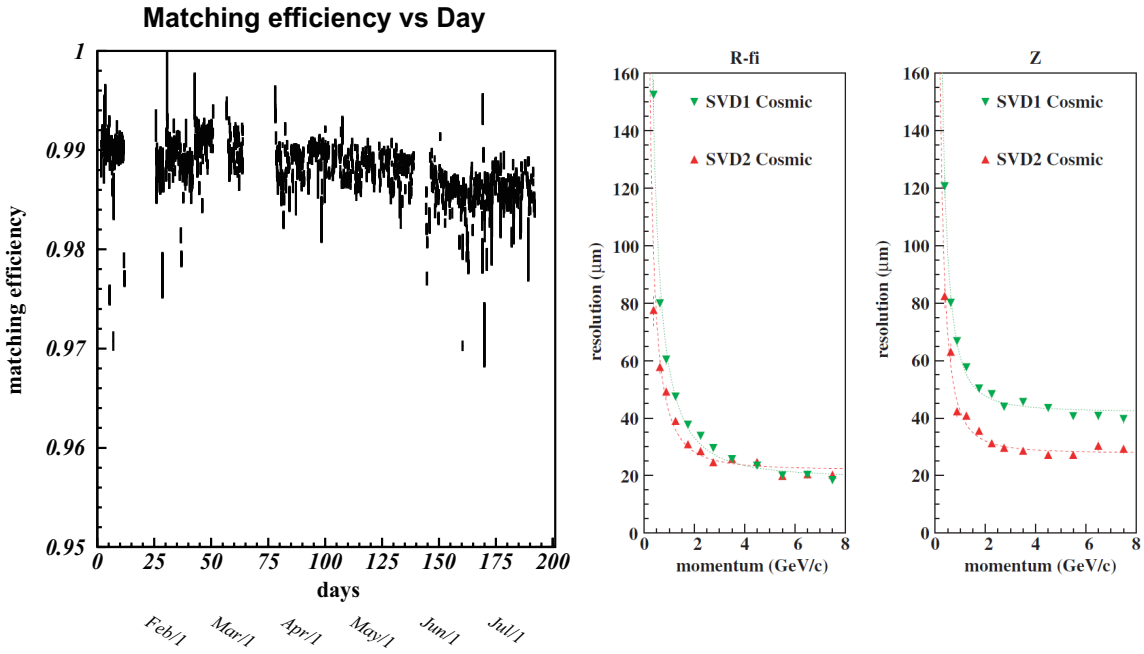


Figure 2.4: Performance of the SVD. (left) Tracking efficiency [38]. (right) Comparison of the impact parameter resolution between the SVD1 and the SVD2 [53].

2.2.2 Central Drift Chamber (CDC)

The CDC is a large-volume tracking device located at the central part of the Belle detector. It plays three important roles: measurements of a track, momentum and energy loss (dE/dx) for a charged particle. It has a cylindrical structure with an 88 cm radius, a 2.4 m length and an acceptance $17° < \theta < 150°$. Gold-plated tungsten wires with a 30 μm diameter and aluminum wires with a 126 μm diameter are used for the sense wires and the field wires, respectively. The field wires are connected to the ground, and a high voltage (HV) of typically 2.35 kV is applied to the sense wires. These wires compose a square-shape cell as shown in Fig. 2.5. There are 50 cylindrical sense-wire layers in total, and three to five layers compose a superlayer. The even-number superlayers are inclined with respect to the z-axis by 40–70 mrad, that provides about 600 μm resolution for the z direction. The z information is used to distinguish tracks originating from the IP from background-induced tracks. On the innermost surface of the CDC cylinder and between the second and the third layers, 7.4 mm wide cathode strips are equipped perpendicularly to the sense wires to measure the z-direction.

The cylinder is filled with a mixture gas of 50%-helium and 50%-ethane. This gas has a small cross section of the photoelectric effect. This is important to reduce background electrons induced by the synchrotron radiation from the beam. It has a radiation length of 640 m, which is sufficiently long to reduce the Coulomb scattering of charged particles.

The output from the sense wire is amplified and sent to a charge-to-time conversion (QTC) module. The QTC module outputs a digital pulse, the leading-edge timing and the width of which correspond to the drift time and the output charge from the CDC, respectively. The QTC output is measured by the FASTBUS multi-hit TDC module.

The momentum of a charged particle is measured using a curvature of the charged

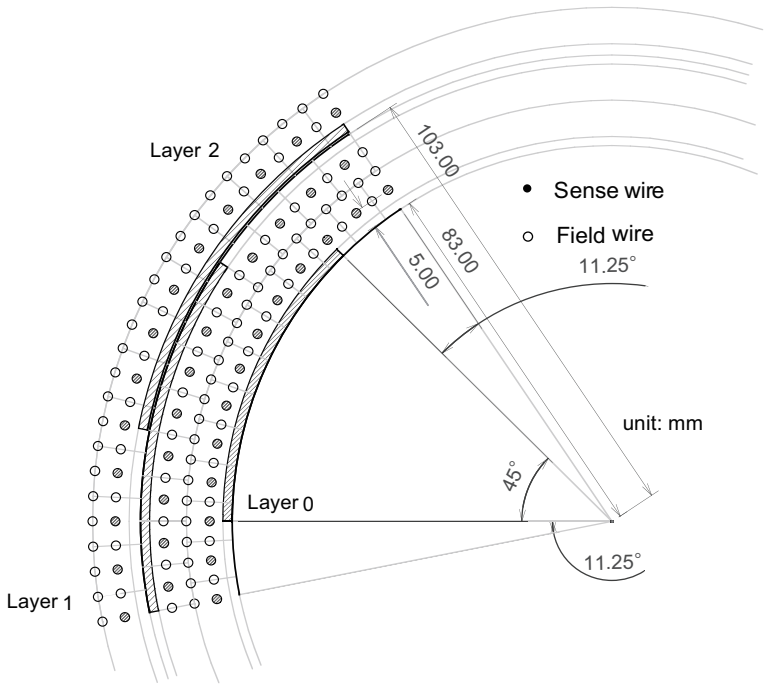


Figure 2.5: Wire configuration in the inner two sense-wire layers of the CDC [38]. The shaded and the open circles indicate positions of the sense wires and the field wires, respectively.

track with the relation

$$p = eBr, \tag{2.1}$$

where p, e, B and r are the momentum, the elementary charge, the magnetic field and the radius of the track curvature, respectively. The momentum resolution is evaluated using cosmic-ray muons passing through the IP. The track is divided into two segments at the IP, and the momentum of each segment is independently measured. The momentum resolution is evaluated from the difference of the obtained momenta. Figure 2.6 (left) shows the momentum resolution as a function of the transverse momentum p_t.The momentum resolution in the high p_t region is worse due to the small track curvature. The typical momentum resolution for a 1 GeV/c particle is about 0.4%.

The dE/dx measurement is performed by collecting output charges from the sense wires on a track. The output charge distributes as the Landau distribution with a long tail. To improve the dE/dx resolution, higher 20% of the sense wire outputs are discarded, and the mean of the output charges (truncated mean) is calculated. The dE/dx resolution of about 5% is obtained. Figure 2.6 (right) shows the dE/dx as a function of the momentum for each charged-particle type: $e^{\pm}$, p, $K^{\pm}$ and $\pi^{\pm}$. The dE/dx measurement provides a good discrimination for $K^{\pm}$ and $\pi^{\pm}$ in the momentum range below 0.5 GeV/c, corresponding to $\log_{10}(p) \sim -0.3$ in Fig. 2.6 (right). Since charged particles in this momentum region do not reach the outer detectors, the dE/dx measurement provides the PID information complementary to the measurements in the TOF and the ACC.

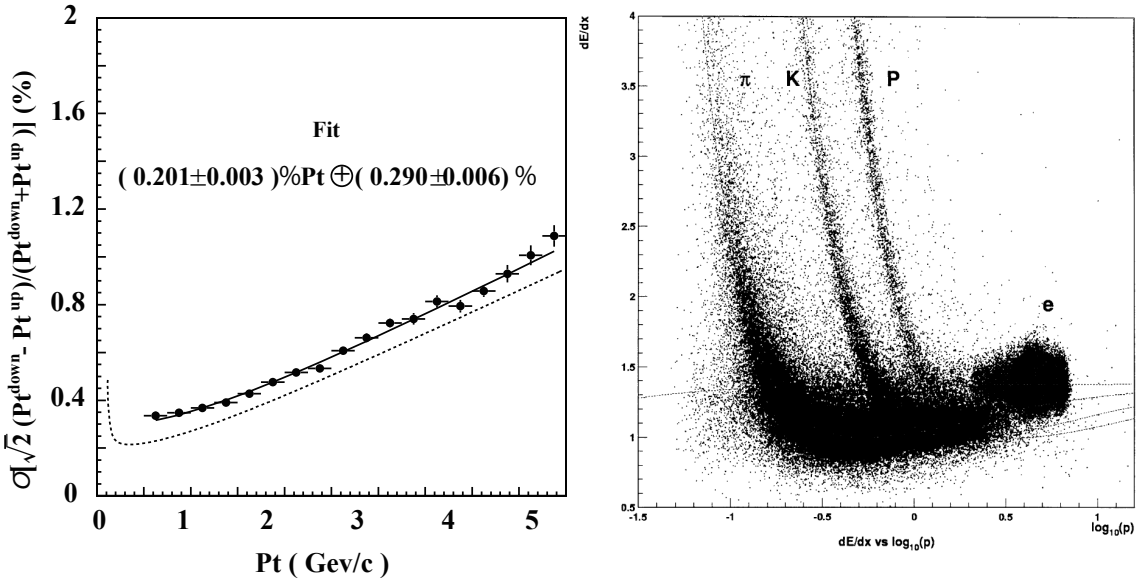


Figure 2.6: CDC performance reported in Ref. [38]. (left) Momentum resolution as a function of the transverse momentum. The dashed line indicates the prediction using the MC simulation while the black dots are obtained from data. (right) Truncated mean of dE/dx as a function of the momentum.

2.2.3 Time Of Flight Counter (TOF)

The TOF performs PID, especially for $K^{\pm}$ and $\pi^{\pm}$, on the barrel region ($33^{\circ} < \theta < 121^{\circ}$) by measuring the time of flight of a charged particle between the IP and the TOF. Using the time of flight t, the velocity is obtained by

$$\beta = \frac{L}{ct} \tag{2.2}$$

where L and c denote the flight distance of the charged particle and the speed of light, respectively. Depending on the polar angle, L takes a value from 122 to 224 cm. Combining the obtained β with the momentum p measured in the CDC, the particle type is identified.

One TOF module consists of a 4 cm thick × 6 cm wide × 255 cm long plastic scintillator bar and two fine-mesh PMTs at the two ends of the bar. We use the polyvinyltoluene-based plastic scintillator with a decay constant of 2.4 ns and the light propagation velocity of 14.4 cm/s, and the fine-mesh PMT Hamamatsu R6680 with a time resolution of about 320 ps for single-photon detection. The fine-mesh PMT has a good tolerance to the magnetic field. In total, 128 TOF modules cover the barrel region. A trigger scintillation counter (TSC) is attached to a pair of the TOF modules. It consists of a thinner scintillator bar with dimensions of 0.5 cm thick × 12 cm wide × 263 cm long and a fine-mesh PMT. By requiring coincidence hits in the TOF and the TSC, background hit rates are suppressed in a high beam-background environment.

An output from the fine-mesh PMT is divided into two paths; one is used for the charge measurement and another for the time-of-flight measurement. The pulse-height dependence of the signal timing (time walk) is corrected using the output charge. To achieve the good discrimination between $K^{\pm}$ and $\pi^{\pm}$, a time-of-flight resolution less than 100 ps

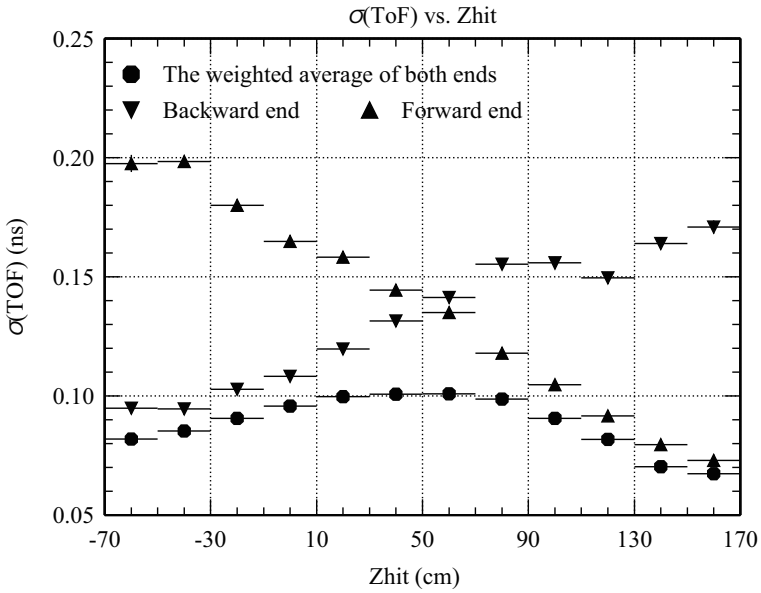


Figure 2.7: Time resolution of a single TOF module [55]. Zhit is the hit position of charged particles on the TOF counter.

is required. As shown in Fig. 2.7, using the weighted average of the PMT signal timing on both ends, the time resolution better than 100 ps is achieved. In the data analysis, the PID for charged particles is performed based on a likelihood ratio constructed from detector responses in the CDC and the ACC as well as the TOF. The PID performance is discussed in Sec. 2.4.

2.2.4 Aerogel Cherenkov Counter (ACC)

The time-of-flight difference between $K^{\pm}$ and $\pi^{\pm}$ at the same momentum is closer to zero in the higher-momentum region where β is almost equal to 1. The TOF is not capable of performing a good PID above 1 GeV/c. To perform PID for the high-momentum particles, the ACC is introduced on the barrel ($34^\circ < \theta < 127^\circ$) and the forward endcap regions ($17^\circ < \theta < 34^\circ$).

The ACC is the threshold-type Cherenkov counter using silica-aerogel as a radiator. The condition for a charged particle to emit the Cherenkov light is

$$\beta > \frac{1}{n}, \tag{2.3}$$

where n denotes the refractive index of the aerogel. Let us assume $n = 1.01$ and consider the momentum range between 0.7 GeV/c and 2.5 GeV/c, for example. In this condition, $\pi^{\pm}$ ($m_\pi = 140$ MeV/c^2) radiates the Cherenkov light while $K^{\pm}$ ($m_K = 494$ MeV/c^2) does not. The particle type is thus identified from the Cherenkov photon yield.

Figure 2.8 shows the drawings of the barrel and the endcap ACC modules [38]. Each module contains five aerogel tiles in the box with an approximate size of $12 \times 12 \times 12$ cm^3, and one or two fine-mesh PMTs are attached. On the barrel region, 60 ACC modules are annually arrayed for the ϕ-direction, composing 16 rings for the z-direction. In total, 960 modules are installed. On the endcap region, 228 modules compose the five-layer ACC array.

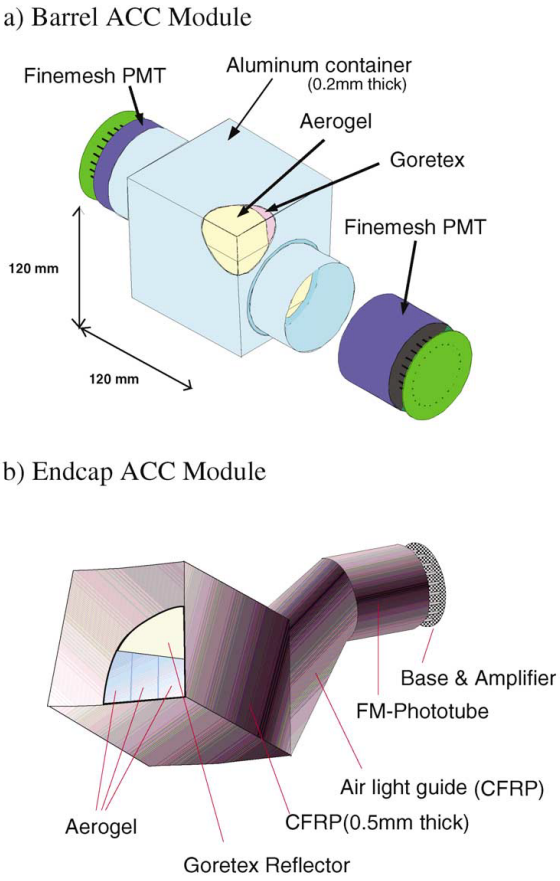


Figure 2.8: Drawing of the ACC modules [38]: (a) barrel module and (b) endcap module.

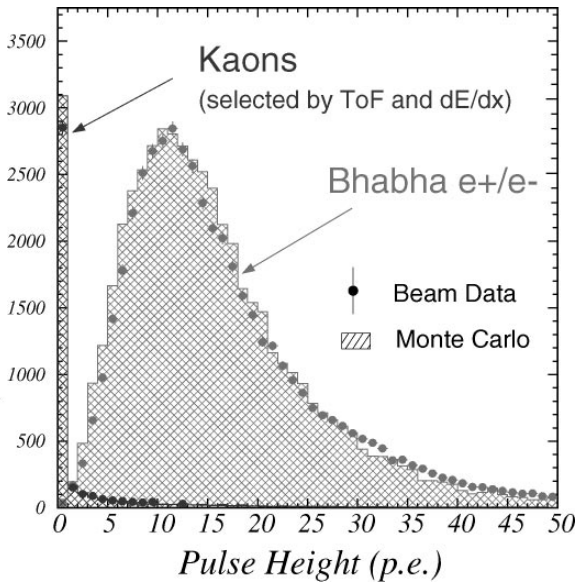


Figure 2.9: Demonstration of the separation of $K^{\pm}$ from $e^{\pm}$ [56].

Since the typical momentum of charged particles is different depending on the polar angle, six different aerogel indices are adopted: $n = 1.030$ for the endcap region and $n = 1.010$–1.028 for the barrel region. The refractive indices are controlled within 3% precision. Photons from the aerogel tiles are collected by the fine-mesh PMTs [56]. An output signal is amplified, processed by the QTC module and digitized by the TDC module. Figure 2.9 shows the pulse-height distribution with the $K^{\pm}$ and the $e^{\pm}$ samples. The $K^{\pm}$ candidates are selected based on the dE/dx and the time of flight, and the $e^{\pm}$ candidates are obtained from the Bhabha scattering $e^+e^- \to e^+e^-$. These two pulse-height distributions demonstrate a clear separation between $K^{\pm}$ and $e^{\pm}$. Details of the PID performance are discussed in Sec. 2.4.

2.2.5 Electromagnetic Calorimeter (ECL)

The ECL is mainly used for the photon reconstruction and the electron identification. In the photon reconstruction, a good efficiency, position and energy resolutions are required over the wide energy range; below 500 MeV for photons from cascade-decays of B mesons and up to a few GeV for two-body radiative decays such as $\bar{B} \to K^*\gamma$. Two nearby photon separation is crucial since some neutral mesons are reconstructed from their di-photon decay. From these points of view, Tl-doped CsI crystals are selected.

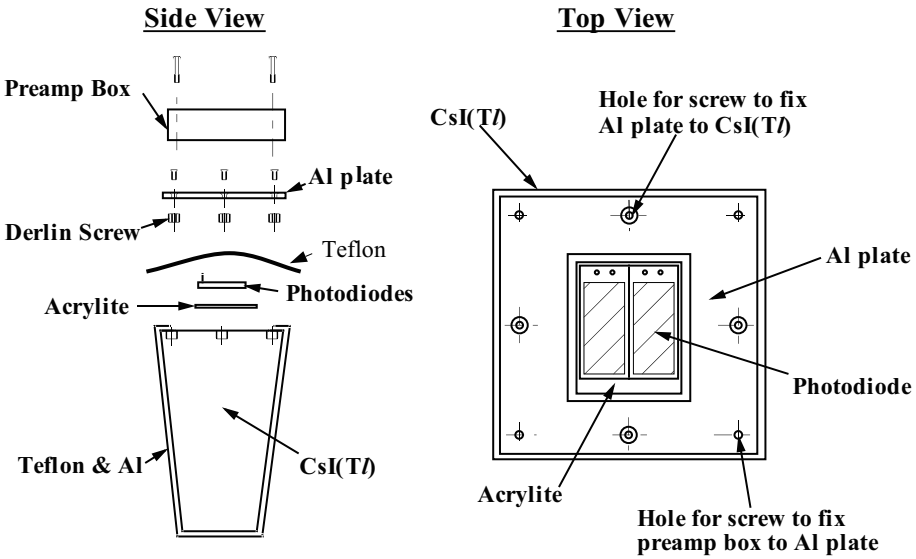


Figure 2.10: Drawing of the ECL module [38].

The ECL module is illustrated in Fig. 2.10. The modules on the barrel region ($32.2° < \theta < 128.7°$) have the crystal cross section of 55×55 mm^2 at the inner side and 65×65 mm^2 at the outer side. The cross section of the endcap modules ($12.4° < \theta < 31.4°$, $130.7° < \theta < 155.1°$) varies from 44.5 to 70.8 mm for the inner side and 54 to 82 mm for the outer side. The tower-like shape of the crystal makes an array so that every module points to the IP. The crystal length is 30 cm, corresponding to 16.2 radiation length. To increase the photon yield, the crystal is covered by 200-μm-thick Goretex teflon, which is an efficient diffuse reflector. At the outer side of the crystal, two PIN photodiodes (Hamamatsu S2744-08) are attached to read the scintillation light. A set of the crystal and the photodiodes are contained in the aluminum assembly.

Two output signals from the photodiodes are summed and then divided into two paths; one is used for the energy measurement, and the other for the trigger. The main signal is shaped with a time constant of 1 μs. The charge information is then converted to the timing information by the QCT module and measured by the multi-hit TDC.

The performance of the ECL is evaluated with Bhabha-scattering events. The energy resolution is about 4% at 100 MeV to 1.6% at 8 GeV [38].

2.2.6 K_L^0 and Muon Detector (KLM)

Using the iron return yoke for the magnetic field, the KLM identifies high penetration particles such as K_L^0 and $\mu^\pm$. While K_L^0 or other types of hadrons interacts strongly with the return yoke, $\mu^\pm$ is free from the hadronic interaction. Using the difference of the interaction with the KLM, these particles are distinguished.

The KLM consists of alternative layers of 4.7 cm thick iron plates and 3.7 cm thick resistive plate chamber (RPC) modules. There are 15(14) RPC layers on the barrel (endcap) region, covering the polar angle $20° < \theta < 155°$. Figure 2.11 shows the cross section of the RPC layer. It has a pair of the RPCs. Each RPC module consists of two 2.4 mm thick float-glass plates with a 2 mm gap. The glass plates are coated with highly resistive ink to provide HV electrodes. Signal pick-up strips are placed outside the two

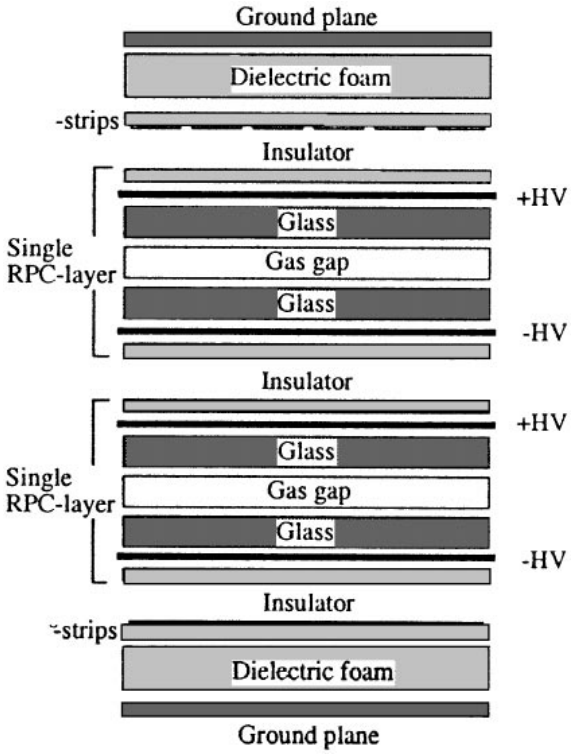


Figure 2.11: Superlayer of the KLM module [57].

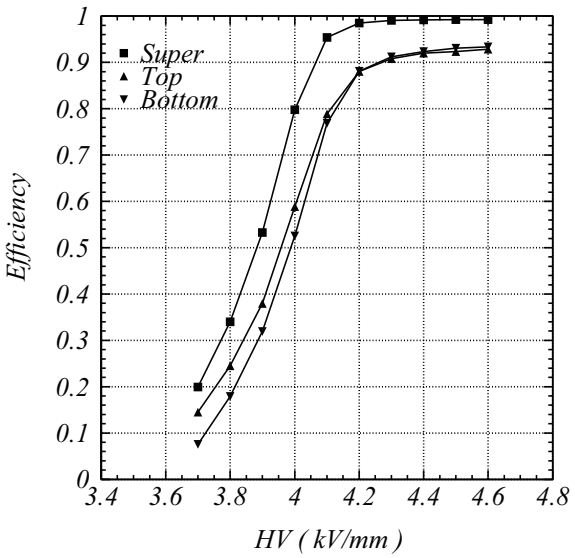


Figure 2.12: Efficiency as a function of the HV applied to the KLM module [38]. The graphs for Super, Top and Bottom show the efficiencies of the combined two RPCs in a superlayer, the top-layer RPC and the bottom-layer RPC, respectively.

RPC plates. The strips are 36 to 50 mm wide depending on the module position, and the upper and the lower strips are mutually orthogonal.

The gap is filled with a mixture gas of 62% HFC-134a (CH_2FCF_3), 30% argon and 8% butane-silver (C_4H_{10}). During the operation, HVs of -3.5 kV and $+4.5$ kV ($+4.7$ kV for the endcap modules) are applied to each HV electrode. Charged particles penetrate through the gap with ionizing the gas, that triggers a discharge. Since the resistivity of the glass surface is high, the discharge occurs locally. The induced charge is detected by both sides of the strips. A typical output is about 100 mV high and 50 ns wide. The output is digitized by a comparator. The hit information from 12 strips are multiplexed and measured by the multi-hit TDC module. Figure 2.12 shows the efficiency for a single cosmic-ray muon track as a function of the applied HV. At the plateau, the efficiency of about 99% is obtained.

In our study, separation of $\mu^\pm$ from other charged hadrons is important to reconstruct $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$ events. It is performed using a difference of the cluster shape; a $\mu^\pm$ track passes through a long range of the KLM, while a hadron track stops quickly. Details of the muon identification are discussed in Sec. 2.4.3.

2.3 Trigger and Data Acquisition

As mentioned in each sub-detector section, most of the sub-detector systems use the QTC module to digitize detector signals and the FASTBUS TDC module to record the data. As shown in Fig. 2.13, data from each detector are collected and combined by the event

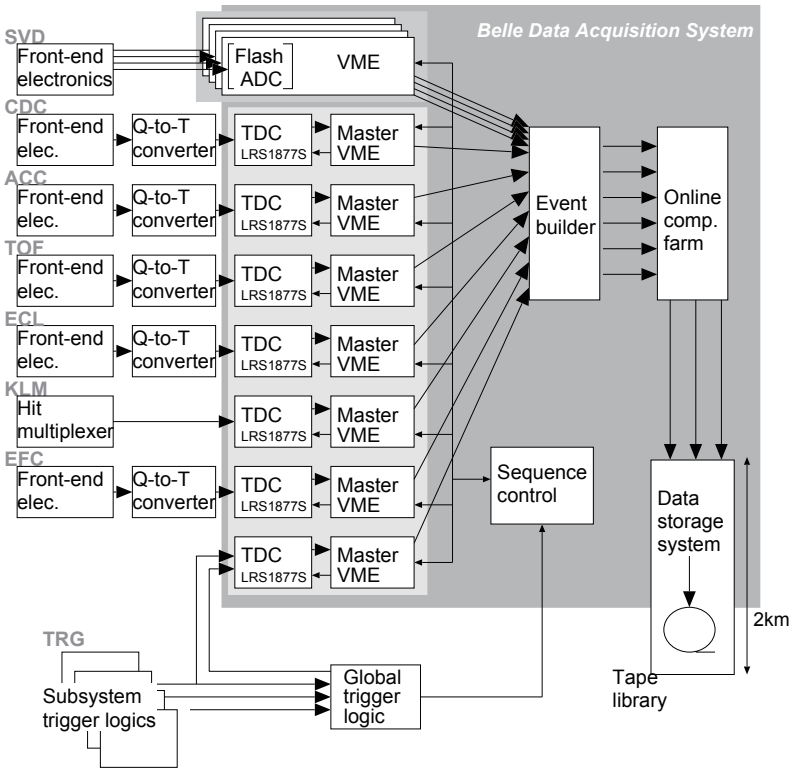


Figure 2.13: Schematic diagram of the Belle DAQ system [38].

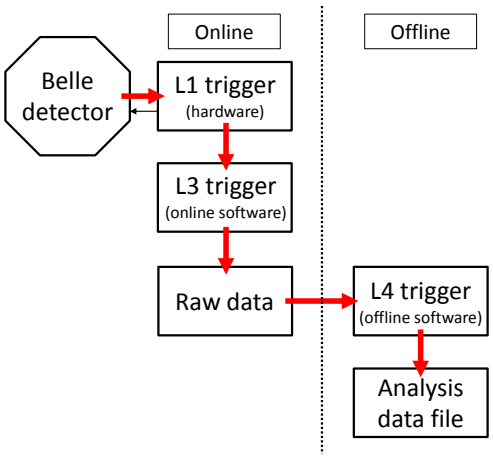


Figure 2.14: Schematic diagram of the trigger system.

builder and sent to the computer.

As shown in Fig. 2.14, Belle has the three-stage trigger logic. The first step is the Level-1 (L1) hardware trigger. Sub-trigger signals issued by the CDC, TOF, ECL, KLM and EFC are combined and judged by the global decision logic. Hadronic events such as $e^+e^- \to B\bar{B}$ and $q\bar{q}$ are selected by one of the three trigger conditions: (i) three or more tracks in the CDC, (ii) large energy deposit in the ECL (Bhabha scattering and cosmic ray events are vetoed) and (iii) four or more energy clusters in the ECL. Other trigger conditions are also prepared to take different types of data such as the Bhabha-scattering,

$e^+e^- \to \mu^+\mu^-$ and $e^+e^- \to \tau^+\tau^-$. The total L1-trigger rate is typically 200 Hz, which is dominated by the beam background. The dead time is less than 3%.

Events selected by the L1 trigger are reconstructed and sent to the computer. The next stage, the Level-3 (L3) trigger, is the online software trigger. It selects events with at least one track originating around the IP ($|dr| < 5.0$ cm) and reduces the beam background events by a factor of about 2 while keeping the efficiency of hadronic events more than 99%.

After passing the L3 trigger, the events are recorded on tapes. However, these data still contain many events from the beam background. To reduce the background events, the events are required to pass the Level-4 (L4) offline software trigger. The L4 trigger uses the fast tracking algorithm with coarse resolutions of about 1 mm for $|dr|$ and 1 cm for $|dz|$. The events with at least one good quality track with $|dr| < 1$ cm, $|dz| < 4$ cm and a transverse momentum greater than 300 MeV/c are retained and fully reconstructed for data analyses. The L4 trigger reduces the data size by a factor of about 5 while keeping the efficiency of hadronic events more than 99%.

For the $B\bar{B}$ analysis, the dedicated data skimming method is applied. The skimming is performed based on the conditions (i) at least three charged tracks with $|dr| < 2$ cm, $|dz| < 4$ cm and $p_t > 0.1$ GeV/c, (ii) the sum of the reconstructed energy ($E^*_{\rm vis}$; the variable with an asterisk is defined in the CM frame) to be $E^*_{\rm vis} > 0.2\sqrt{s}$. Furthermore, requirements to remove the radiative Bhabha scattering and high-multiplicity QED processes are employed: (i) at least two energy clusters in the ECL region $-0.7 < \cos\theta^* < 0.9$, (ii) the average energy of the clusters less than 1 GeV, (iii) the sum of the cluster energy ($E^*_{\rm sum}$) in $0.18\sqrt{s} < E^*_{\rm sum} < 0.8\sqrt{s}$ GeV. A rejected event based on the above condition is added to the analysis data if at least one J/ψ or $\psi(2S)$ candidate is found.

2.4 Particle Reconstruction

From the detector responses, $K^\pm$, $\pi^\pm$, $e^\pm$, $\mu^\pm$ and photons are reconstructed. Neutral pions and K^0_S are reconstructed from the decay modes $\pi^0 \to \gamma\gamma$ and $K^0_S \to \pi^+\pi^-$, respectively. In this section, the particle reconstruction and charged particle identification are overviewed.

2.4.1 Tracking

A charged track is reconstructed by the algorithm which connects adjacent "segments", or clusters of the CDC hit wires. For low-p_t tracks, the CDC information is combined with the SVD hits.

The track finding efficiency is evaluated using partially-reconstructed $D^{*+} \to D^0(\to K^0_S\pi^+\pi^-, K^0_S \to \pi^+\pi^-)\pi^+$ sample. The D^{*+} mesons can be reconstructed without one of the charged pions. The track finding efficiency is evaluated by checking whether or not all the charged pions are found. Figure 2.15 shows the efficiency as a function of p_t. At the p_t range from 0.4 to 1 GeV/c, where most of the decay particles in $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ distribute, the efficiency is more than 90%. The efficiency is almost determined by the limited acceptance of the SVD and the CDC. In the region $p_t < 0.1$ GeV/c, the efficiency drops significantly since p_t is close to the threshold for a charged track to penetrate a few layers of the SVD.

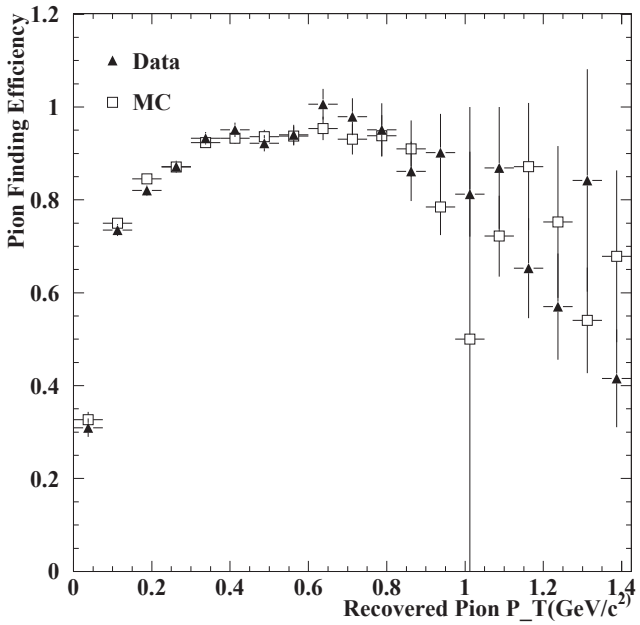


Figure 2.15: Tracking efficiency as a function of p_t [58].

When the beam background increases, the segments become unclear. The conventional track finding algorithm often fails to reconstruct a track. To recover the efficiency, the more robust track-finding algorithm based on the Hough transform is developed. Using the new algorithm, the efficiency has significantly improved by 10 to 20% per track relative to the efficiency of the conventional algorithm under the high beam-background environment. The new algorithm is applied to the SVD2 data.

2.4.2 Electron Identification

In the ECL detector, electrons and positrons cause electromagnetic showers and are thus distinguished from other heavier charged particles such as $K^{\pm}$ and $\pi^{\pm}$. In the data analysis, following five variables are used.

Matching of a charged track to an energy cluster (track matching)
Compared to other charged particles, an electron easily creates a clear energy cluster in the ECL. The electron has better track-matching efficiency to the ECL cluster.

(E_9/E_{25})
The variable E_9/E_{25} is defined as a ratio of the total energy measured in 3×3 CsI crystals to that in 5×5 cells. The cell center is determined as the center of gravity of the total energy deposit in an ECL cluster. Due to faster evolution of the electromagnetic shower, an electron has the wider shower shape compared to other charged particles. Therefore distant crystals from the shower center also detect energy from the shower, and E_9/E_{25} tends to be close to 1.

(E/p)
: Since the electron is light, it has $\beta = 1$ at any momentum region. Besides, it loses all the energy in the ECL. The value of E/p is therefore almost equal to 1. On the other hand, massive charged particles tend to have a smaller value of E/p since they lose a small fraction of the energy in the ECL. This variable is useful in the high momentum region $p > 0.5$ GeV/c. Due to the Coulomb scattering with the detector materials, the low-momentum electron has a value of E/p significantly smaller than 1 and overlaps with that of the other charged particles.

Light yield in ACC
: Assuming that a typical refractive index of the ACC is 1.015, the threshold momentum for electrons to emit Cherenkov light in the ACC is 4 MeV/c. In the Belle detector, at least 50 MeV/c is required for charged particles to be reconstructed. Reconstructed electrons therefore always emit the Cherenkov light in the ACC. This is useful to identify low-energy electrons.

dE/dx **in CDC**
: The variable dE/dx provides a good electron identification in the momentum region $p < 0.5$ GeV/c as shown in Fig. 2.6 (right). As the electrons in this momentum region do not reach the ECL, dE/dx is complementary to the track matching, E_9/E_{25} and E/p.

Using the MC sample, the probability for a charged track to be an electron (P_e^i) or a non-electron particle ($P_{\rm not}^i$) is calculated. Here, the index i denotes one of the five variables described above. We calculate the electron identification (e ID) probability $\mathcal{P}_e$ as

$$\mathcal{P}_e = \frac{\prod_{i=1}^{5} P_e^i}{\prod_{i=1}^{5} P_e^i + \prod_{i=1}^{5} P_{\rm not}^i}. \tag{2.4}$$

The e ID efficiency is measured using the QED process such as the radiative Bhabha scattering [59]. By tagging one of the tracks to be an electron, the accompanying charged track is implied to be an election without imposing any PID requirement. As shown in Fig. 2.16, about 95% efficiency is obtained for the momentum larger than 1 GeV/c. Since e ID largely relies on the ECL information, the lower-momentum region shows the worse efficiency. A slight discrepancy between the data and the Monte Carlo (MC) simulation is observed, which is corrected from this comparison.

To take into account high-multiplicity effects in $B\bar{B}$ events, the e ID efficiency is also measured using $J/\psi \to e^+e^-$ events originating from B meson decays. The J/ψ candidates with an invariant mass between 2.5 and 3.5 GeV/c^2 are selected. The e ID efficiency is evaluated using the same method as that applied to the QED process sample. The difference of the efficiency between two samples is taken as a systematic uncertainty.

2.4.3 Muon Identification

The most probable particle that can be mis-identified as $\mu^\pm$ is the charged hadron such as $\pi^\pm$. When a series of the KLM hits matching to a track extrapolated from the CDC is found, following two variables are used to separate $\mu^\pm$ from charged hadrons.

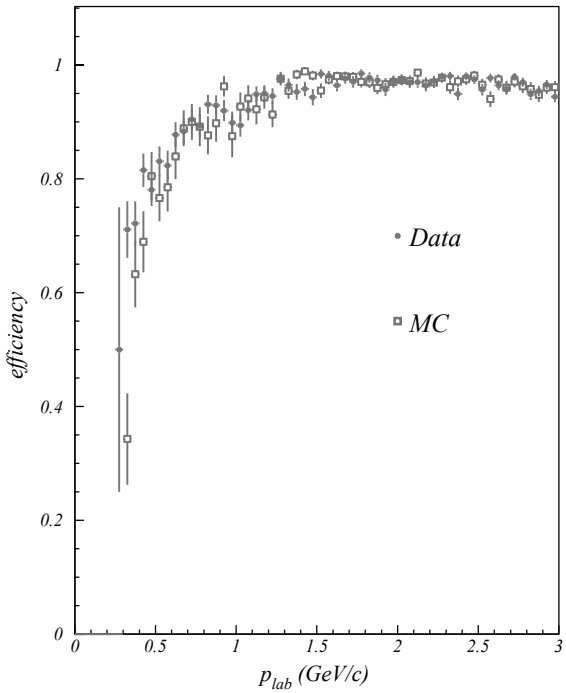


Figure 2.16: Efficiency of e ID as a function of the momentum in the laboratory frame [59]. The requirement $\mathcal{P}_e > 0.5$ is applied.

ΔR

We firstly find the layer number of the outermost RPC having a hit associated with an extrapolated track ($L_{\rm hit}$). It is compared with the layer number of the outermost RPC which the extrapolated track ($L_{\rm ext}$) reach with the hypothesis that the track is not scattered in the KLM. The variable ΔR is then defined by $\Delta R \equiv L_{\rm ext} - L_{\rm hit}$. Since $\mu^{\pm}$ has a good penetrating power over the KLM, it tends to reach the RPC layer $L_{\rm ext}$ and hence ΔR becomes close to zero.

χ^2

The χ^2 is calculated from the deviation of each KLM hit position from the extrapolated track. Since hadrons are scattered more strongly in the KLM than $\mu^{\pm}$, they have a larger χ^2 value. If a charged track does not reach the KLM, $\chi^2 = 0$ is assigned to the track.

We then define the muon identification (μ ID) probability by

$$\mathcal{P}_\mu \;=\; p(\Delta R)p(\chi^2), \tag{2.5}$$

where $p(\Delta R)$ and $p(\chi^2)$ are the probability density functions for $\mu^{\pm}$ to take values of ΔR and χ^2. Since we assign $p(\chi^2 = 0) = 0$, $\mathcal{P}_\mu$ is equal to 0 for a track which does not reach the KLM.

Using the similar method applied to the e ID efficiency determination, the μ ID efficiency is evaluated [60]. As shown in Fig. 2.17, the efficiency of about 90% is obtained with the $\mu^{\pm}$ momentum greater than 1 GeV/c. Since at least about 0.6 GeV/c is required for muons to reach the KLM, the efficiency significantly drops in the lower momentum region.

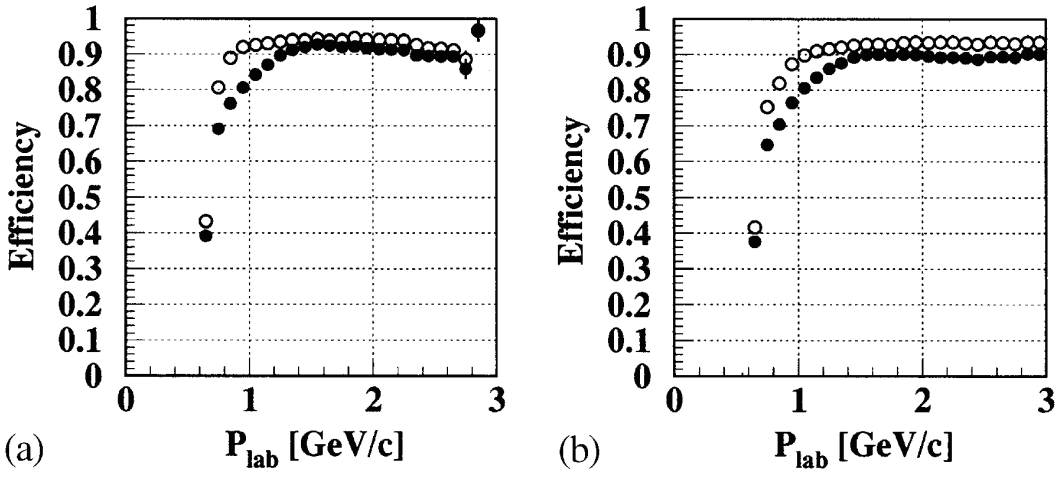


Figure 2.17: Efficiency of μ ID as a function of the momentum in the laboratory frame with $\mathcal{P}_\mu > 0.1$ (open circle) and $\mathcal{P}_\mu > 0.9$ (closed circle) [60]. (a) Barrel region ($51° < \theta < 117°$). (b) Entire region ($25° < \theta < 145°$).

2.4.4 Charged Hadron Identification

Since multiple $K^\pm$ and $\pi^\pm$ are often found in the final state of B meson decay, their identification is essential. To perform the $K^\pm$ and $\pi^\pm$ identification, we use three inner sub-detectors: CDC, TOF and ACC. These detectors provide measurements related to the velocity β. The particle mass is determined once the momentum is given. On each variable, a likelihood for the $K^\pm(\pi^\pm)$ hypothesis is calculated as $L^i_{K(\pi)}$, where $L^1_{K(\pi)}$ is the likelihood based on dE/dX at the CDC, $L^2_{K(\pi)}$ based on the time of flight at the TOF and $L^3_{K(\pi)}$ based on the photon yield at the ACC. These likelihoods are combined into a $K^\pm$ identification (K ID) or a $\pi^\pm$ identification (π ID), which is defined as

$$\mathcal{P}_{K(\pi)} = \frac{\prod_{i=1}^{3} L^i_{K(\pi)}}{\prod_{i=1}^{3} L^i_{K(\pi)} + \prod_{i=1}^{3} L^i_{\pi(K)}}. \tag{2.6}$$

Similarly, a proton identification (p ID) is defined as

$$\mathcal{P}_p = \frac{\prod_{i=1}^{3} L^i_{p}}{\prod_{i=1}^{3} L^i_{p} + \prod_{i=1}^{3} L^i_{\pi}}. \tag{2.7}$$

In our analysis, the p ID is used to reject background events containing protons.

To measure the K ID and π ID efficiencies, the $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$ sample is used. Owing to its final state containing only three charged particles, the D^{*+} mesons are cleanly reconstructed without imposing any PID requirement. The particle types are determined from their charges; two particles with the same charge are $\pi^\pm$, and the remaining is $K^\pm$. As shown in Fig. 2.18, about 85% of the PID efficiency and less than 10% fake rate is typically obtained. The visible discrepancy between the data and the MC simulation is corrected based on this comparison.

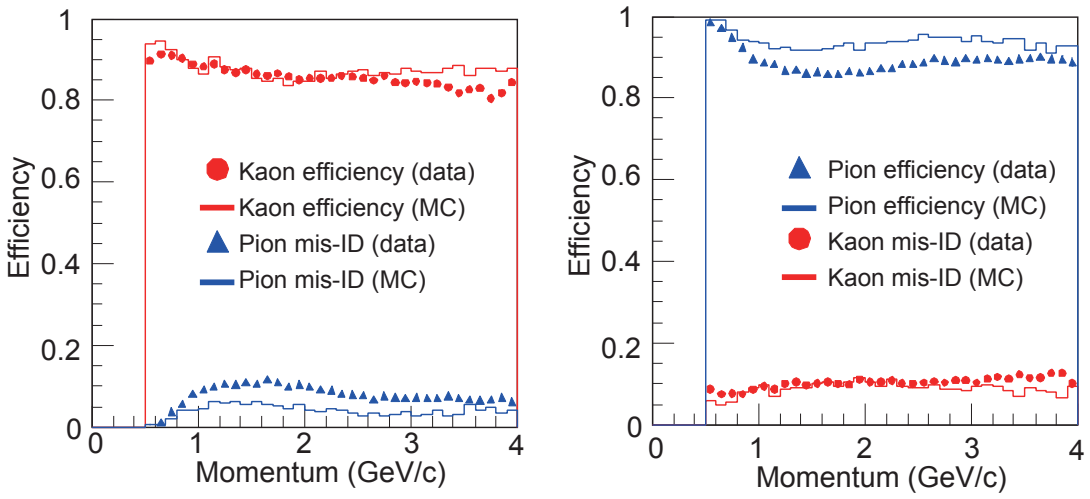


Figure 2.18: Efficiency of K ID (left) and π ID (right) as a function of the momentum in the laboratory frame [61].

2.4.5 Photon Reconstruction

Photon candidates are reconstructed using the ECL. First, segments are formed using the CsI crystals with measured energy greater than 0.5 MeV, and a crystal with the maximum energy (seed) is found for each segment. The energy cluster is determined as the 5×5 crystals around the seed. If the 5×5 region is overlapped with a neighbor cluster, the energy on the overlapped region is assigned to each cluster by weighting based on the total energy in the non-overlapped region. Once the cluster is formed, the center of gravity of the energy in the 5×5 region is found as the cluster center. The energy cluster is then compared to the extrapolated charged tracks in the CDC. If there is no track matching to the cluster, it is recorded as a photon candidate.

Due to the limited length of the CsI crystals, there is energy leakage of the electromagnetic shower from the crystal. The measured energy of the cluster is hence significantly underestimated by 1 to 10%, depending on the incident photon energy. Besides, since a part of the energy is lost in the gap between the crystals, the finite crystal cross section makes the center of gravity closer to the crystal center than the actual photon incident position. These effects are corrected based on the MC simulation, validated by reconstructed π^0 candidates from pairs of the photons.

2.4.6 π^0 Reconstruction

Neutral pions decay into a pair of the photons with a 99% branching fraction [10]. We therefore reconstruct π^0 candidates from a pair of photons and select them based on the invariant mass $M_{\gamma\gamma}$, which must be around the nominal π^0 mass of 135 GeV/c^2 [10]. Figure 2.19 is one example of the $M_{\gamma\gamma}$ distribution. A typical mass resolution is 5 MeV/c^2.

The π^0 reconstruction efficiency is evaluated using $\tau^- \to \pi^-\pi^0\nu_\tau$ events in the $e^+e^- \to \tau^+\tau^-$ sample. First, the number of $(\tau^+ \to \ell^+\nu_\ell\bar{\nu}_\tau, \tau^- \to \pi^-\pi^0\nu_\tau)$ pairs (single-signal events) and $(\tau^+ \to \pi^+\pi^0\bar{\nu}_\tau, \tau^- \to \pi^-\pi^0\nu_\tau)$ pairs (double-signal events) are counted. Since the branching fraction of $\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau$ is well determined as $(17.82 \pm 0.04)\%$ for the

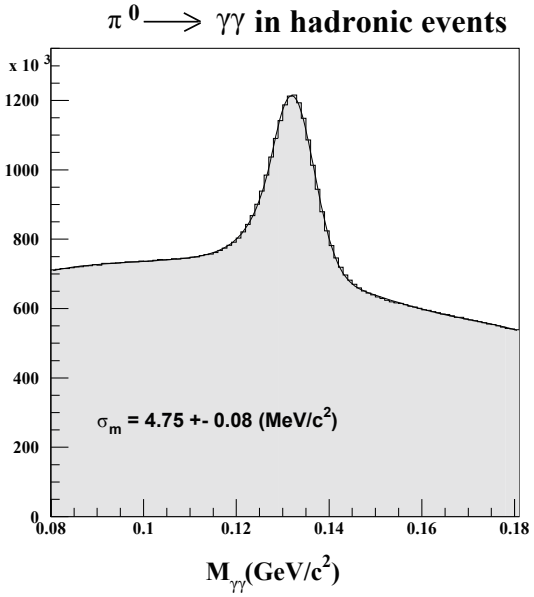


Figure 2.19: Invariant mass distribution for the π^0 candidates [38].

$\ell^- = e^-$ mode and $(17.39 \pm 0.04)\%$ for the $\ell^- = \mu^-$ mode, the efficiency is obtained from the comparison of the number of single- and double-signal events. The π^0 reconstruction efficiency is found to be typically 40 to 50%.

2.4.7 K^0_S Reconstruction

Due to the short lifetime of 895 ps [10], K^0_S mostly decays in the detector. They are reconstructed from the decay $K^0_S \to \pi^+\pi^-$ with the branching fraction of 69% [10]. At first, two oppositely-charged tracks are combined with a pion mass hypothesis. Since most of K^0_S decay in the SVD, their vertices are determined precisely. Using the feature that K^0_S has a vertex detached from the IP, following five variables are used to select K^0_S candidates.

$dr_{\rm min}$: Smaller value of $|dr|$ between two $\pi^\pm$ tracks. True K^0_S candidates must have a larger $dr_{\rm min}$ since the decay vertex of K^0_S is inconsistent with the IP.

$d\phi$: Angle between the momentum vector of the K^0_S candidate and the K^0_S vertex direction with respect to the origin of the Belle coordinates. These are aligned if the K^0_S candidate is true.

$z_{\rm dist}$: Closest distance between two $\pi^\pm$ tracks. True K^0_S candidates tends to have $z_{\rm dist} \sim 0$ as they are produced at the same point.

fl : Flight distance of a K^0_S candidate in the $x-y$ plane. While mis-combination of two prompt $\pi^\pm$ have $fl \sim 0$, true K^0_S candidates have non-zero values of fl.

$M_{\pi\pi}$: Invariant mass of the K^0_S candidate.

The basic requirement on each variable is summarized in Table 2.2.

Table 2.2: Requirements for the K^0_S selection. $m^{\rm PDG}_{K^0_S} = 497.6$ GeV/c^2 is the nominal K^0_S mass [10].

Momentum (GeV/c)	dr (cm)	$d\phi$ (rad)	$z_{\rm dist}$ (cm)	fl (cm)	$\lvert M_{\pi\pi} - m^{\rm PDG}_{K^0_S}\rvert$ (MeV/c^2)
< 0.5	> 0.05	< 0.3	< 0.8	–	< 30
0.5–1.5	> 0.03	< 0.1	< 1.8	> 0.08	< 30
> 1.5	> 0.02	< 0.03	< 2.4	> 0.22	< 30

The K^0_S reconstruction efficiency is estimated using the $D^{*+} \to D^0(\to K^0_S\pi^+\pi^-)\pi^-$ sample. Since D^{*+} candidates are cleanly reconstructed, the K^0_S sample with loose selection, where only the $M_{\pi\pi}$ requirement to be $350 < M_{\pi\pi} < 650$ MeV/c^2 is imposed, is expected to have almost 100% efficiency with a good purity. By comparing the numbers of K^0_S candidates with the loose selection and with the basic requirements, the efficiency is estimated to be 85 to 90%.

Chapter 3

Event Reconstruction

In this chapter, we describe the data analysis procedure. First, the method of the $R(D^*)$ and $P_\tau(D^*)$ measurement are overviewed in the first section. Next, the data samples used for our study are introduced in the second section. Based on the described method and the data set, the signal selection criteria are described in the third and the forth sections. In the last section, calibration and validation of the MC simulation samples using the calibration data samples are discussed.

To avoid subjective bias of the analyzer, we conduct a blind data analysis, where the signal sample of the real data is not used until the analysis procedure is completely validated.

3.1 Measurement Method

3.1.1 $R(D^*)$ Measurement

For the $R(D^*)$ measurement, $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ and $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ events need to be reconstructed to determine the numerator and the denominator, respectively, of $R(D^*)$. We call these modes "signal mode" and "normalization mode", respectively.

Since there are one and two neutrinos in the normalization mode and the signal mode, respectively, these modes cannot be fully reconstructed. We employ the method "hadronic tag", which exploits the feature that exactly two B mesons are produced in a e^+e^- collision without any other extra particles, as described in Sec. 1.3. In this method, $B_{\rm tag}$ is reconstructed from its hadronic decay. Using the beam momentum and the fully-reconstructed $B_{\rm tag}$ momentum, the complete momentum of $B_{\rm sig}$ is obtained.

To extract the yields of the normalization mode, we use the missing mass squared

$$M^2_{\rm miss} \equiv (E_{e^+e^-} - E_{\rm tag} - E_{D^*} - E_\ell)^2/c^4 - (\vec{p}_{e^+e^-} - \vec{p}_{\rm tag} - \vec{p}_{D^*} - \vec{p}_\ell)^2/c^2, \quad (3.1)$$

where E and $\vec{p}$ denote the energy and the three-momentum, respectively, for the e^+e^- beam, $B_{\rm tag}$, D^* and ℓ^-. The $M^2_{\rm miss}$ is equivalent to the square of the invariant mass of the missing four-momentum. Since only one neutrino is a missing particle in the $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decay, the normalization events are expected to distribute around $M^2_{\rm miss} = 0$ GeV$^2/c^4$.

It is difficult to use $M^2_{\rm miss}$ for the extraction of the signal yield since it has a strong correlation to $P_\tau(D^*)$. As shown in Fig. 3.1 (left), the $M^2_{\rm miss}$ shape strongly depends on the value of $P_\tau(D^*)$. For the signal mode, instead of $M^2_{\rm miss}$, we use $E_{\rm ECL}$, which is a

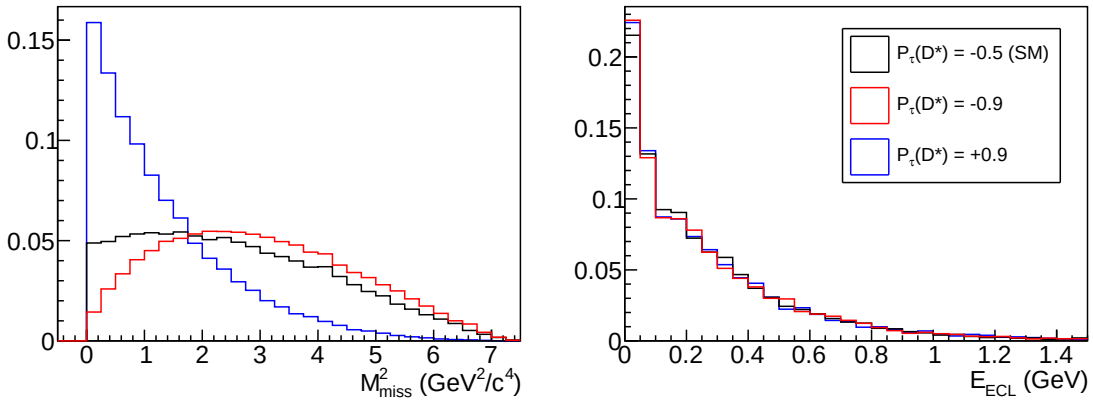


Figure 3.1: Comparison of the $M^2_{\rm miss}$ (left) and the $E_{\rm ECL}$ (right) shapes at $P_\tau(D^*) = -0.5$ (SM), -0.9 and $+0.9$.

linear sum of the energy of ECL clusters not used for the $B_{\rm tag}$ and $B_{\rm sig}$ reconstruction. Signal events ideally have $E_{\rm ECL}$ equal to zero and have a long tail in the $E_{\rm ECL}$ distribution due to the beam background and split-off showers, separated from the main ECL cluster and reconstructed as photon candidates. As shown in Fig. 3.1 (right), the $E_{\rm ECL}$ shape is independent of $P_\tau(D^*)$. According to the previous studies, for example, Ref. [43], it also provides a clear discrimination between the signal and the other background modes.

3.1.2 $P_\tau(D^*)$ Measurement

The value of $P_\tau(D^*)$ is determined using the differential decay rate

$$\frac{1}{\Gamma}\frac{d\Gamma}{d(\cos\theta_{\rm hel})} = \frac{1}{2}\left[1 + \alpha P_\tau(D^*)\cos\theta_{\rm hel}\right], \tag{3.2}$$

$$\alpha = \begin{cases} 1, & \text{for } \tau^- \to \pi^-\nu_\tau, \\ \frac{m_\tau^2 - 2m_\rho^2}{m_\tau^2 + 2m_\rho^2}, & \text{for } \tau^- \to \rho^-\nu_\tau, \end{cases} \tag{3.3}$$

where the coefficient α represents a sensitivity to $P_\tau(D^*)$ depending on each τ decay mode [62]. Using $m_\rho = 775$ MeV/c^2 [10], α is found to be 0.45. Equation 3.2 is calculated based on the combination of the spin states of the τ lepton and the $\pi(\rho) - \bar{\nu}_\tau$ system, requiring for the (anti-)neutrino to be left(right)-handed. Details of the calculation are described in Appendix A.

In the previous measurement using the $\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau$ decay [42], about 500 signal events were obtained. There are a few major differences between this measurement and our study.

- In the previous measurement, about half of the $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ events were extracted as feed-down events reconstructed as $\bar{B} \to D\tau^-\bar{\nu}_\tau$. Since we do not reconstruct $\bar{B} \to D\tau^-\bar{\nu}_\tau$, such events cannot be used in our analysis.

- The sum of the branching fractions of $\tau^- \to \pi^-\nu_\tau$ and $\tau^- \to \rho^-\nu_\tau$ is 36.0% [10]. This is almost equal to $BF(\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau) = 35.2\%$.

Taking the above differences into account, the expected number of signal events in our study is roughly estimated to be 250. This is not sufficient to measure the angular distribution in Eq. 3.2. We therefore extract $P_\tau(D^*)$ as an asymmetry of the signal yields in two $\cos\theta_{\rm hel}$ regions: $\cos\theta_{\rm hel} > 0$ (forward) and $\cos\theta_{\rm hel} < 0$ (backward). By integrating Eq. 3.2 in $[0, 1]$ and $[-1, 0]$, the number of signal events in the forward region ($N_{\rm F}$) and in the backward region ($N_{\rm B}$) are, respectively, obtained as

$$N_{\rm F} = \frac{N_{\rm F} + N_{\rm B}}{2}\left[1 + \frac{\alpha}{2}P_\tau(D^*)\right], \tag{3.4}$$

$$N_{\rm B} = \frac{N_{\rm F} + N_{\rm B}}{2}\left[1 - \frac{\alpha}{2}P_\tau(D^*)\right]. \tag{3.5}$$

The value of $P_\tau(D^*)$ is then obtained by

$$P_\tau(D^*) = \frac{2}{\alpha}\frac{N_{\rm F} - N_{\rm B}}{N_{\rm F} + N_{\rm B}}. \tag{3.6}$$

Although $\cos\theta_{\rm hel}$ needs to be measured in the rest frame of τ, a complete four-momentum of the τ lepton cannot be determined due to one unconstrained degree of freedom (Appendix C for further discussion). We have established a method to retrieve a correct value of $\cos\theta_{\rm hel}$ as follows.

First, we take the rest frame of W^* as shown in Fig. 3.2 (a). This frame is obtained by boosting the laboratory frame along

$$\vec{q} = \vec{p}_{\rm sig} - \vec{p}_{D^*}, \tag{3.7}$$

$$= \vec{p}_{e^+e^-} - \vec{p}_{\rm tag} - \vec{p}_{D^*}, \tag{3.8}$$

where $\vec{q}$ represents the three momentum vector of W^* and the other variables are defined in Eq. 3.1. In this frame, a cosine of the angle between the τ momentum and its daughter meson ($\cos\theta_{\tau d}$) is calculated as

$$\cos\theta_{\tau d} = \frac{2E_\tau E_d - m_\tau^2 - m_d^2}{2|\vec{p}_\tau||\vec{p}_d|}, \tag{3.9}$$

where E, $\vec{p}$ and m are the energy, the momentum and the mass, respectively, of τ and its daughter $d = \pi$ or ρ. Since $W^{*-} \to \tau^-\bar{\nu}_\tau$ is a two-body decay of the stationary W^* boson, $|\vec{p}_\tau|$ depends only on q^2:

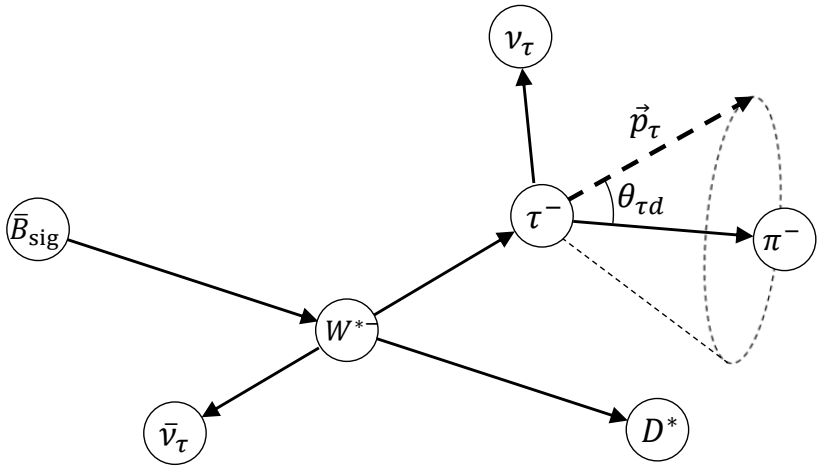
$$|\vec{p}_\tau| = \frac{q^2 - m_\tau^2/c^2}{2\sqrt{q^2}}, \tag{3.10}$$

where q^2 is experimentally determined by

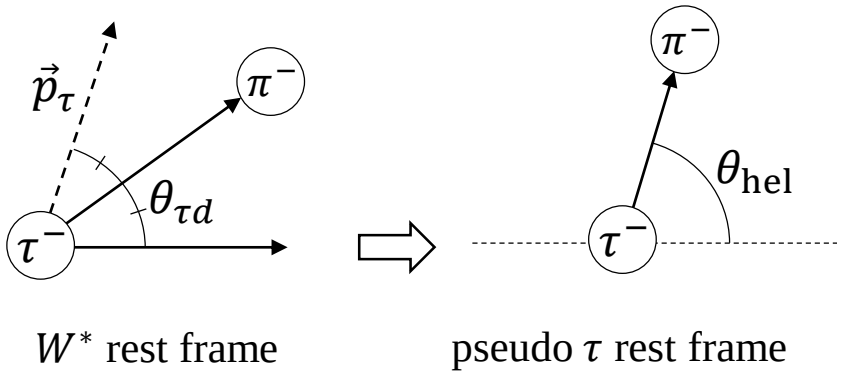
$$q^2 = (E_{e^+e^-} - E_{\rm tag} - E_{D^*})^2/c^2 - (\vec{p}_{e^+e^-} - \vec{p}_{\rm tag} - \vec{p}_{D^*})^2. \tag{3.11}$$

The energy of the τ lepton is also determined by $E_\tau = \sqrt{m_\tau^2c^4 + |\vec{p}_\tau|^2c^2}$. The values of $|\vec{p}_d|$ and E_d are measurable, and we use $m_\tau = 1.78$ GeV/c^2 and $m_d = 140$ MeV/c^2 ($d = \pi^\pm$) or 775 MeV/c^2 ($d = \rho^\pm$) reported in Ref. [10]. The value of $\cos\theta_{\tau d}$ is hence completely determined.

Now, as shown in Fig. 3.2 (a), $\vec{p}_\tau$ is constrained on the cone around $\vec{p}_d$ with the opening angle $\theta_{\tau d}$. Even though the direction of $\vec{p}_\tau$ is not completely fixed, the remaining one degree of freedom is confined to the rotation symmetry of the the cone. Namely, we can take any arbitrary vector on the cone as a boost vector. Choosing one particular direction as shown in Fig. 3.2 (b), the frame is boosted to the "pseudo" τ-rest frame, where a correct value of $\cos\theta_{\rm hel}$ is uniquely obtained.



(a)



(b)

Figure 3.2: Illustration of the decay topology and the boost of the rest frame of W^*. (a) Decay topology of $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ in the rest frame of W^*. The arrow indicates the direction of the momentum vector of each particle (length is not to scale). Every particle except for $\bar{B}_{\rm sig}$ is indicated at the end of the vector. The dashed arrow expresses the τ momentum, which is on the cone with the opening angle $\theta_{\tau d}$ around the π^- momentum. (b) Transformation from the rest frame of W (left) to the pseudo τ rest frame (right). The boost axis is taken as the horizontal solid arrow in the left panel, the magnitude of which is equal to $|\vec{p}_\tau|$. The direction of the boost vector is indicated as the horizontal dashed line in the right side. As the frame obtained by this boost is not necessarily consistent with the rest frame of τ, we name this frame "pseudo" τ rest frame.

3.2 Data Sample

In the data analysis, we use two types of the data samples. The real data sample, or simply referred to as data, are the sample accumulated in the actual experiment with the Belle detector. The second one is the MC simulation sample produced on computers, which is used to estimate the signal reconstruction efficiency and study the background.

3.2.1 Real Data Sample

In this study, we use the full data sample recorded at $\sqrt{s} = 10.58$ GeV, containing $(772 \pm 11) \times 10^6$ $B\bar{B}$ pairs. The number of $B\bar{B}$ pairs ($N_{B\bar{B}}$) is estimated using the formula

$$N_{B\bar{B}} = \frac{N^{\rm on} - r(\epsilon_{q\bar{q}})\alpha N^{\rm off}_{q\bar{q}}}{\epsilon_{B\bar{B}}}. \tag{3.12}$$

The variable $N^{\rm on}$ is the number of events satisfying the hadronic event conditions described in Sec. 2.3 at the $\Upsilon(4S)$ resonance (on-resonance). This is the sum of the number of $B\bar{B}$ events and $q\bar{q}$ events. The number of $q\bar{q}$ events $N^{\rm off}_{q\bar{q}}$ is estimated at the off-resonance energy, by 60 MeV below the $\Upsilon(4S)$ resonance. Since no B meson is produced at off-resonance, all of hadronic events originate from the $e^+e^- \to q\bar{q}$ process. The total number of $B\bar{B}$ events recorded with the Belle detector is obtained by subtracting $N^{\rm off}_{q\bar{q}}$ from $N^{\rm on}$. The coefficients $r(\epsilon_{q\bar{q}})$ and α are the relative $q\bar{q}$ efficiency and cross-section, respectively, at on-resonance with respect to the off-resonance. The denominator $\epsilon_{B\bar{B}}$ is the efficiency of the $B\bar{B}$ events, which is around 99% over the operation period.

3.2.2 Monte Carlo Sample

To simulate particle decays and calculate a four-momentum of each particle, we use the following event generators and simulation packages. Events with a $B\bar{B}$ pair are generated with EvtGen [63], and the B meson decays are reproduced based on branching fractions reported in Ref. [10]. The hadronization process of the B meson decay with no experimentally-measured branching fractions are inclusively reproduced by PYTHIA [64]. For the $q\bar{q}$ events, the initial quark pair is hadronized by PYTHIA, and decays of the produced hadrons are modeled by EvtGen. The final-state radiation from charged particles is added by PHOTOS [65]. The generated particles are processed by the Belle detector simulator based on GEANT3 [66] to reproduce detector responses. Types of the produced MC samples are described below.

$\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ **(signal) MC**
: The signal MC sample is used to study the signal distributions and the reconstruction efficiency. The MC sample is produced with the FFs in Eqs. 1.31 to 1.34. The statistics of the signal MC sample is equivalent to the 40 times larger luminosity than the Belle experiment. Branching fractions of the decays $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ are set to be $BF(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau) = 1.25\%$ and $BF(B^- \to D^{*0}\tau^-\bar{\nu}_\tau) = 1.43\%$. These are determined by $BF(\bar{B} \to D^*\tau^-\bar{\nu}_\tau) = R(D^*) \times BF(\bar{B} \to D^*\ell^-\bar{\nu}_\ell)$ with $R(D^*) = 0.252$ based on the SM prediction [13] and $BF(B^- \to D^*\ell^-\bar{\nu}_\ell) = 5.69\%$ for the B^- mode or $BF(\bar{B}^0 \to D^*\ell^-\bar{\nu}_\ell) = 4.95\%$ for the $\bar{B}^0$ mode [10].

$\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ **(normalization) MC**
: The $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ MC sample equivalent to 10 times larger luminosity than the Belle experiment is produced with the HQET FFs. The older values of the FF parameters were used in the MC production, and the decay kinematics is corrected to match to that generated using the values in Eqs. 1.31 to 1.34. Details of the correction are further discussed in Appendix B.

$\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ MC

The decay $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$, where D^{**} denotes excited D meson states heavier than D^*, is one of the important background modes in our analysis. To study the systematic uncertainty arising from this mode, we generate the $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ MC sample with the statistics equivalent to 40 times larger luminosity than the Belle experiment. The branching fractions for the orbitally-excited states D_0^*, D_1, D_1' and D_2^* are taken from the world averages [17], and the isospin symmetry is imposed on the D^{**} decays. Other possible modes such as $D_0^* \to D^{(*)}\eta$ are also included based on Ref. [50]. The radially-excited $D^{(*)}(2S)$ modes are included to fill the observed gap

Table 3.1: Branching fractions of the decays $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ and D^{**} in the MC simulation.

Decay mode	$BF(\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell)$	$BF(D^{**})$	
$\bar{B} \to D_0^*\ell^-\bar{\nu}_\ell$	0.67%	$D_0^* \to D\pi$	64.93%
		$D_0^* \to D\pi\pi$	11.69%
		$D_0^* \to D^*\pi\pi$	11.69%
		$D_0^* \to D\eta$	5.85%
		$D_0^* \to D^*\eta$	5.85%
$\bar{B} \to D_1\ell^-\bar{\nu}_\ell$	0.68%	$D_1 \to D\pi\pi$	37.13%
		$D_1 \to D^*\pi$	62.86%
$\bar{B} \to D_1'\ell^-\bar{\nu}_\ell$	0.30%	$D_1' \to D\pi\pi$	14.29%
		$D_1' \to D^*\pi$	75.01%
		$D_1' \to D\eta$	7.14%
		$D_1' \to D\rho$	2.38%
		$D_1' \to D^*\rho$	1.19%
$\bar{B} \to D_2^*\ell^-\bar{\nu}_\ell$	0.60%	$D_2^* \to D\pi$	35.01%
		$D_2^* \to D^*\pi$	17.5%
		$D_2^* \to D\pi\pi$	31.67%
		$D_2^* \to D\eta$	15.83%
$\bar{B} \to D(2S)\ell^-\bar{\nu}_\ell$	0.50%	$D(2S) \to D\pi\pi$	33.34%
		$D(2S) \to D_0^*\pi$	33.33%
		$D(2S) \to D^*\pi$	16.67%
		$D(2S) \to D_2^*\pi$	16.67%
$\bar{B} \to D^*(2S)\ell^-\bar{\nu}_\ell$	0.50%	$D(2S)^* \to D^*\pi\pi$	33.34%
		$D^*(2S) \to D_1'\pi$	33.33%
		$D^*(2S) \to D(2S)\gamma$	16.67%
		$D^*(2S) \to D(2S)\pi$	16.67%
Total	2.99%		

between the inclusive and the sum of exclusive measurements of the $\bar{B} \to X_c \ell^- \bar{\nu}_\ell$ branching fractions [50], where X_c denotes all the possible charmed-meson states. Values of the branching fractions are summarized in Table 3.1. The $\bar{B} \to D^{**} \ell^- \bar{\nu}_\ell$ MC sample is produced using the FFs based on the ISGW model [67], and the kinematics is corrected to match the LLSW model [68]. The kinematics correction is described further in Appendix B.

$\bar{B} \to D^{**} \tau^- \bar{\nu}_\tau$ MC
The decays $\bar{B} \to D^{**} \tau^- \bar{\nu}_\tau$ are the same types of the background as $\bar{B} \to D^{**} \ell^- \bar{\nu}_\ell$. To generate the $\bar{B} \to D^{**} \tau^- \bar{\nu}_\tau$ MC sample, we use the branching fractions in Table 3.2, which are calculated from the theoretical estimates of $R(D^{**}) \equiv BF(\bar{B} \to D^{**} \tau^- \bar{\nu}_\tau)/BF(\bar{B} \to D^{**} \ell^- \bar{\nu}_\ell)$ for each D^{**} state [69] and $BF(\bar{B} \to D^{**} \ell^- \bar{\nu}_\ell)$ in Table 3.1. The branching fraction is by one order of magnitude smaller than that of the $\bar{B} \to D^{**} \ell^- \bar{\nu}_\ell$ decays because the phase space is suppressed due to the mass of the τ lepton. We produce the MC sample with the statistics equivalent to 400 times larger luminosity than the Belle experiment.

Generic background MC
For the remaining background components, the MC sample with generic B meson decays ($B\bar{B}$ MC) is produced. In this MC sample, distributions in two-body decay distributions are modeled based on quantum numbers of the B meson and the decay particles. For decays with multiple final-state particles, intermediate resonances are taken into account based on the experimental measurements. If there is no past experimental study, the distribution of decay particles is determined based on the phase space. The continuum $e^+e^- \to q\bar{q}$ processes ($q\bar{q}$ MC) are also prepared. We use the $B\bar{B}$ and $q\bar{q}$ MC samples with the statistics equivalent to ten and five times larger luminosity than the Belle experiment, respectively.

3.3 Reconstruction of the Tag Side

The $B_{\rm tag}$ candidates are reconstructed using 32 decay modes shown in Table 3.3. The total branching fraction of these decay modes is 11.6% for B^- and 9.4% for $\bar{B}^0$. The daughter mesons $D^{(*)0}$, $D_{(s)}^{(*)+}$ and J/ψ are reconstructed from the decay modes in Tables 3.4 and 3.5. In total 1104 decay chains are used.

The $B_{\rm tag}$ reconstruction is performed using the multivariate analysis (MVA) [70] based on the NeuroBayes neural-network package [71]. Since we use the beam-energy-constraint

Table 3.2: Branching fractions of the decays $\bar{B} \to D^{**} \tau^- \bar{\nu}_\tau$ in the MC simulation.

Decay mode	Branching fraction
$\bar{B} \to D_0^* \tau^- \bar{\nu}_\tau$	0.057%
$\bar{B} \to D_1 \tau^- \bar{\nu}_\tau$	0.061%
$\bar{B} \to D_1' \tau^- \bar{\nu}_\tau$	0.022%
$\bar{B} \to D_2^* \tau^- \bar{\nu}_\tau$	0.041%
Total	0.18%

mass $M_{\rm bc} \equiv \sqrt{E^{*2}_{\rm beam}/c^4 - |\vec{p}^{\,*}_B|^2/c^2}$ ($E^*_{\rm beam}$ and $\vec{p}^{\,*}_B$ are the beam energy and the reconstructed $B_{\rm tag}$ momentum, respectively) to separate correctly-reconstructed $B_{\rm tag}$ candidates from combinatorial background, variables uncorrelated to $M_{\rm bc}$ are chosen as inputs. To effectively maximize the $B_{\rm tag}$ reconstruction efficiency (tag efficiency), the hierarchical MVA is performed. At each step of the following four-stage MVA, daughter particles are selected based on loose requirements and the network is trained.

Step 1. Charged particles, K^0_S, γ and π^0 reconstruction

Charged particles are required to have $|dr| < 2$ cm and $|dz| < 4$ cm to reject background-induced particles produced outside the IP. This requirement is not applied to daughter $\pi^\pm$ candidates for K^0_S because their vertices are detached from the IP. Candidate γ and K^0_S are selected based on the reconstruction procedure in Secs. 2.4.5 and 2.4.7, respectively. Neutral pion candidates are reconstructed from a pair of photons with energy greater than 30 MeV. All of available information such as the PID and the reconstructed invariant mass are used as input variables for the the MVA. After the MVA process, the NeuroBayes output classifier $O^1_{\rm NB,i}$ is obtained for the ith particle. The output classifier represents a probability that the

Table 3.3: List of the B decay modes for the $B_{\rm tag}$ reconstruction. The columns BF show the branching fraction of each decay mode reported in Ref. [10].

B^-		$\bar{B}^0$	
Mode	BF (%)	Mode	BF (%)
$B^- \to D^{*0}\pi^-$	0.52	$\bar{B}^0 \to D^{*+}\pi^-$	0.27
$B^- \to D^{*0}\pi^-\pi^0$	1.0	$\bar{B}^0 \to D^{*+}\pi^-\pi^0$	1.5
$B^- \to D^{*0}\pi^-\pi^-\pi^+$	1.0	$\bar{B}^0 \to D^{*+}\pi^-\pi^-\pi^+$	0.70
$B^- \to D^{*0}\pi^-\pi^-\pi^+\pi^0$	1.8	$\bar{B}^0 \to D^{*+}\pi^-\pi^-\pi^+\pi^0$	1.8
$B^- \to D^0\pi^-$	0.48	$\bar{B}^0 \to D^+\pi^-$	0.25
$B^- \to D^0\pi^-\pi^0$	1.3	$\bar{B}^0 \to D^+\pi^-\pi^0$	0.75
$B^- \to D^0\pi^-\pi^-\pi^+$	0.57	$\bar{B}^0 \to D^+\pi^-\pi^-\pi^+$	0.6
$B^- \to D^+\pi^-\pi^-$	0.11	$\bar{B}^0 \to D^0\pi^0$	0.026
$B^- \to D^0K^-$	0.037	$\bar{B}^0 \to D^{*+}D^{*-}_s$	1.8
$B^- \to D^{*0}D^{*-}_s$	1.7	$\bar{B}^0 \to D^{*+}D^-_s$	0.80
$B^- \to D^{*0}D^-_s$	0.82	$\bar{B}^0 \to D^+D^{*-}_s$	0.061
$B^- \to D^0D^{*-}_s$	0.76	$\bar{B}^0 \to D^+D^-_s$	0.72
$B^- \to D^0D^-_s$	0.90	$\bar{B}^0 \to J/\psi K^0_S$	0.044
$B^- \to J/\psi K^-$	0.10	$\bar{B}^0 \to J/\psi K^-\pi^+$	0.12
$B^- \to J/\psi K^-\pi^0$	0.11	$\bar{B}^0 \to J/\psi K^0_S\pi^+\pi^-$	0.022
$B^- \to J/\psi K^0_S\pi^-$	0.22		
$B^- \to J/\psi K^-\pi^+\pi^-$	0.081		
Total $BF = 11.6\%$		Total $BF = 9.4\%$	

particle is correctly identified or reconstructed.

Table 3.4: List of the $D^{(*)}$ decay modes used for the $B_{\rm tag}$ reconstruction. The columns BF show the branching fraction of each decay mode reported in Ref. [10].

$D^{(*)0}$		$D^{(*)+}$	
Mode	BF (%)	Mode	BF (%)
$D^{*0} \to D^0\pi^0$	64.7	$D^{*+} \to D^0\pi^+$	67.7
$D^{*0} \to D^0\gamma$	35.3	$D^{*+} \to D^+\pi^0$	30.7
Total $BF = 100\%$		Total $BF = 98.4\%$	
$D^0 \to K^-\pi^+$	3.9	$D^+ \to K^-\pi^+\pi^+$	9.5
$D^0 \to K^-\pi^+\pi^0$	14	$D^+ \to K^-\pi^+\pi^+\pi^0$	6.1
$D^0 \to K^-\pi^+\pi^+\pi^-$	8.1	$D^+ \to K^0_S\pi^+$	1.5
$D^0 \to K^0_S\pi^0$	1.2	$D^+ \to K^0_S\pi^+\pi^0$	7.2
$D^0 \to K^0_S\pi^+\pi^-$	2.9	$D^+ \to K^0_S\pi^+\pi^+\pi^-$	3.1
$D^0 \to K^0_S K^+K^-$	0.45	$D^+ \to K^+K^-\pi^+$	1.0
$D^0 \to K^+K^-$	4.0	$D^+ \to K^+K^-\pi^+\pi^0$	2.6
$D^0 \to \pi^+\pi^-$	0.14		
$D^0 \to \pi^+\pi^-\pi^0$	1.5		
$D^0 \to K^0_S\pi^+\pi^-\pi^0$	5.2		
Total $BF = 41.6\%$		Total $BF = 31.0\%$	

Table 3.5: List of the $D_s^{(*)+}$ and J/ψ decay modes used for the $B_{\rm tag}$ reconstruction. The columns BF show the branching fraction of each decay mode reported in Ref. [10].

$D_s^{(*)+}$		J/ψ	
Mode	BF (%)	Mode	BF (%)
$D_s^{*+} \to D_s^+\gamma$	93.5%	$J/\psi \to e^-e^+$	5.97
$D_s^+ \to K^+K^-\pi^+$	5.5	$J/\psi \to \mu^-\mu^+$	5.96
$D_s^+ \to K^0_S K^+$	1.5		
$D_s^+ \to \pi^+\pi^+\pi^-$	1.1		
$D_s^+ \to K^+K^-\pi^+\pi^0$	6.3		
$D_s^+ \to K^0_S K^-\pi^+\pi^+$	1.7		
$D_s^+ \to K^0_S K^+\pi^+\pi^-$	1.0		
$D_s^+ \to K^+\pi^+\pi^-$	0.66		
$D_s^+ \to K^+K^-\pi^+\pi^+\pi^-$	0.87		
Total $BF = 18.6\%$		Total $BF = 11.9\%$	

Step 2. $D^{\pm}_{(s)}$, D^0 **and** J/ψ **reconstruction**

Next, combining the reconstructed particles, $D^{\pm}_{(s)}$, D^0 and J/ψ candidates are reconstructed from the decay modes in Tables 3.4 and 3.5. To reject D mesons from the $e^+e^- \to c\bar{c}$ process, $p^*_D < 2.6$ GeV/c is required, where p^*_D denotes the D momentum in the CM frame. Each reconstructed candidate is assigned a value $O^1_{\rm NB,prod} = \prod_i^N O^1_{\rm NB,i}$, where N is the number of daughter particles. The relation between the retained numbers of correctly-reconstructed (true) and wrong combination (fake) candidates is determined by adjusting the requirement on $O^1_{\rm NB,prod}$ as shown in Fig. 3.3. Each line has a steep slope at the right end. It indicates that only the number of retained fake candidates increases without gaining the number of true candidates. We set the pre-requirements at the points indicated in Fig. 3.3 and reject obviously wrong candidates. The retained candidates are processed with the second step of the MVA, where all the usable variables such as $O^1_{\rm NB,prod}$, the momentum and the reconstructed invariant mass of the candidate are input. The output classifier of the second step, $O^2_{\rm NB}$, is calculated for each $D^{\pm}_{(s)}$, D^0 or J/ψ candidate.

Step 3. $D^{*\pm}_{(s)}$, D^{*0} **reconstruction**

Candidate $D^{*\pm}_{(s)}$ and D^{*0} are reconstructed from the modes in Tables 3.4 and 3.5. The same procedure as Step 2 is employed and the output classifier $O^3_{\rm NB}$ is obtained.

Step 4. $B_{\rm tag}$ **formation**

Finally, $B_{\rm tag}$ candidates are reconstructed from the decay modes in Table 3.3. The same procedure as Steps 2 and 3 is employed and the output classifier $O_{\rm NB}$ is obtained for each $B_{\rm tag}$ candidate.

In order to suppress the continuum $e^+e^- \to q\bar{q}$ background, one more step is developed using differences in the distribution of final-state particles; it is jet-like for $q\bar{q}$ events and spherical for $B\bar{B}$ events. The difference is quantitatively characterized by the event-

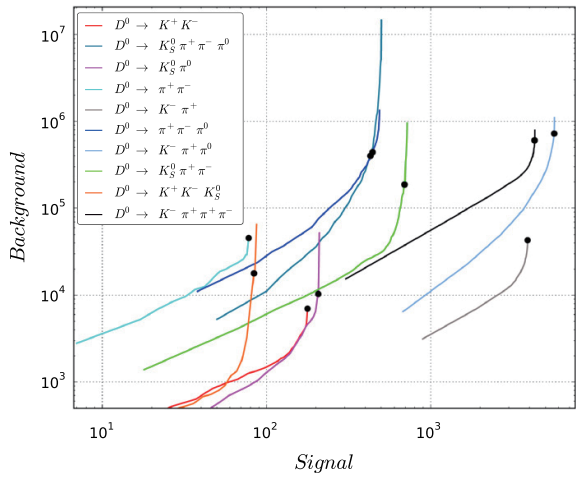


Figure 3.3: Relation between the numbers of true and fake D^0 candidates for each D^0 mode, obtained from the MC study [70]. The pre-requirements for Step 2 are indicated by the black dots.

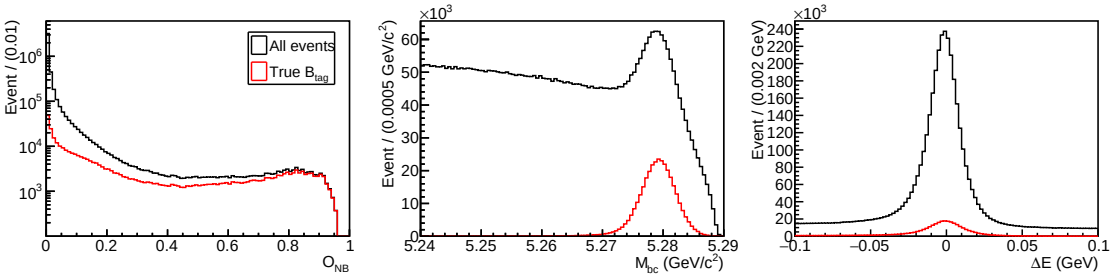


Figure 3.4: $B_{\rm tag}$ distributions obtained from the MC sample, amount of which is equivalent to the full data sample at Belle. From left to right, $O_{\rm NB}$, $M_{\rm bc}$ and ΔE distributions.

shape variables such as the reduced second Fox-Wolfram Moment [72], Super Fox-Wolfram Moment [73] and $\cos\theta_B$ (cosine of the polar angle of the $B_{\rm tag}$ momentum).

Figure 3.4 shows the distributions of $O_{\rm NB}$, $M_{\rm bc}$ and ΔE, where $\Delta E \equiv E^*_{\rm tag} - E^*_{\rm beam}$ and $E^*_{\rm tag}$ is the reconstructed energy of the $B_{\rm tag}$ candidate in the CM frame. The true $B_{\rm tag}$ candidates tend to have larger values of $O_{\rm NB}$, peak at $M_{\rm bc} = 5.28$ GeV/c^2 and distribute around $\Delta E = 0$ GeV. The ΔE peak in the background is significantly enhanced since ΔE is used as one of the input variables for the MVA.

If at least one $B_{\rm tag}$ candidate is reconstructed, we require

- $O_{\rm NB} > 0.008(0.006)$ for $B^-_{\rm tag}$ ($\bar{B}^0_{\rm tag}$),
- $M_{\rm bc} > 5.272$ GeV/c^2,
- $-150 < \Delta E < 100$ MeV.

At this requirement of $O_{\rm NB}$, about 90% of the true $B_{\rm tag}$ candidates are retained while 70% of the wrong candidates are rejected. If there are two or more $B_{\rm tag}$ candidates retained, we select only one with the highest $O_{\rm NB}$ value.

Since the network is trained using the MC simulation, discrepancies between the data and the MC sample must be corrected. According to the previous studies, for example, Refs. [42] and [74], the discrepancy in the tag efficiency is by 20 to 30%. It is therefore essential to correct the tag efficiency with the data, as discussed later in Sec. 3.5.1.

3.4 Reconstruction of the Signal Side

We reconstruct the signal mode and the normalization mode using the particle candidates not used for the $B_{\rm tag}$ reconstruction. The particle decay modes in Table 3.6 are used. The signal selection requirements have been optimized to maximize the figure of merit

$$FOM \equiv \frac{S}{\sqrt{S+B}}, \tag{3.13}$$

where S and B denote the number of correctly-reconstructed signal events and all the other events, respectively. The FOM is calculated in the region $E_{\rm ECL} < 0.5$ GeV, where most of the signal events distribute. In this section, we describe optimized event selection procedures; see Appendix D for the details of the optimization.

Table 3.6: List of the decay modes used in the $B_{\rm sig}$ reconstruction. The branching fractions are obtained from Ref. [10].

Decay mode	BF	Decay mode	BF
$D^0 \to K^-\pi^+\pi^0$	$(14.3 \pm 0.8)\%$	$D^+ \to K^-\pi^+\pi^+$	$(9.46 \pm 0.24)\%$
$D^0 \to K^-\pi^+\pi^+\pi^-$	$(8.06 \pm 0.23)\%$	$D^+ \to K^0_S\pi^+\pi^0$	$(7.24 \pm 0.17)\%$
$D^0 \to K^0_S\pi^+\pi^-\pi^0$	$(5.2 \pm 0.6)\%$	$D^+ \to K^-\pi^+\pi^+\pi^0$	$(6.14 \pm 0.16)\%$
$D^0 \to K^-\pi^+$	$(3.93 \pm 0.04)\%$	$D^+ \to K^0_S\pi^+\pi^+\pi^-$	$(3.05 \pm 0.09)\%$
$D^0 \to K^0_S\pi^+\pi^-$	$(2.85 \pm 0.20)\%$	$D^+ \to K^0_S\pi^+$	$(1.53 \pm 0.06)\%$
$D^0 \to K^0_S\pi^0$	$(1.20 \pm 0.04)\%$	$D^+ \to K^-K^+\pi^+$	$(0.996 \pm 0.026)\%$
$D^0 \to K^-K^+$	$(0.401 \pm 0.007)\%$	$D^+ \to K^0_S K^+$	$(0.295 \pm 0.015)\%$
$D^0 \to \pi^-\pi^+$	$(0.1421 \pm 0.0025)\%$	$D^{*+} \to D^0\pi^+$	$(67.7 \pm 0.5)\%$
$D^{*0} \to D^0\pi^0$	$(61.9 \pm 2.9)\%$	$D^{*+} \to D^+\pi^0$	$(30.7 \pm 0.5)\%$
$D^{*0} \to D^0\gamma$	$(38.1 \pm 2.9)\%$	$\rho^- \to \pi^-\pi^0$	$\sim 100\%$

3.4.1 Particle Selection

First, daughter particles of D^* and τ: $K^\pm$, $\pi^\pm$, K^0_S, γ, π^0 and $\rho^\pm$, and charged leptons $e^\pm$, $\mu^\pm$ are reconstructed. The charged leptons are needed for the reconstruction of normalization events. For the $B_{\rm sig}$ reconstruction, we apply different particle selections from those for the $B_{\rm tag}$ reconstruction. Details of the reconstruction procedure are described below.

$K^\pm$ and $\pi^\pm$

Charged tracks with $|dr| < 0.5$ cm and $|dz| < 2.0$ cm are selected. This selection is applied to all the other charged particles described later, except for the K^0_S daughter candidates. Candidate $K^\pm$ ($\pi^\pm$) are required to have $\mathcal{P}_K > 0.1$ ($\mathcal{P}_\pi > 0.1$). To reject fake candidates from charged leptons, $\mathcal{P}_e < 0.9$ and $\mathcal{P}_\mu < 0.9$ are also imposed. For $\pi^\pm$ used as τ-daughter candidates, additional veto requirements are employed: $\mathcal{P}_e < 0.1$, $\mathcal{P}_\mu < 0.8$ and $\mathcal{P}_p < 0.6$. The vetoes effectively reject background events from specific decay modes such as $\bar{B} \to D^{*(*)}\ell^-\bar{\nu}_\ell$ and $\bar{B} \to D^*\bar{p}n$.

K^0_S

Candidate K^0_S are formed with oppositely-charged two tracks with the pion mass hypothesis. The K^0_S reconstruction method in Sec. 2.4.7 is employed, and candidates with an invariant mass within ± 30 MeV/c^2 of the nominal K^0_S mass (498 MeV/c^2 [10]) are selected.

γ

Photon candidates are reconstructed from energy clusters in the ECL, as described in Sec. 2.4.5. Photon-energy (E_γ) thresholds of 50, 100 and 150 MeV are imposed in the barrel, the forward-endcap and the backward-endcap regions, respectively.

Normal π^0

Neutral pions are reconstructed from pairs of photons. The invariant mass of the π^0 candidates must lie between 115 and 150 MeV/c^2, corresponding to about $\pm 3\sigma$ around the nominal π^0 mass (135 MeV/c^2 [10]). For π^0 candidates from the D and the ρ decays (normal π^0), we impose the same photon energy thresholds described above. There are many fake π^0 candidates due to many combinations of two photons. In order to reduce the number of fake π^0 candidates, we employ the following π^0 candidate-selection procedure, utilizing the tendency that true π^0 candidates have energetic photons. The reconstructed π^0 candidates in one event are ranked in descending order according to the energy of the more energetic daughter. If a given photon is the more energetic daughter for two or more candidates, they are ranked based on the energy of the second photon. Table 3.7 shows one example of the π^0 candidate ranking. We select π^0 candidates from the one with a higher rank and discard the other candidates which shares at least one of the photons with the selected candidate.

Soft π^0

Soft neutral pions are π^0 from the D^* decays. They have a low momentum less than 300 MeV/c because the mass difference between the D and the D^* mesons is almost equal to the pion mass. First, the same requirement on the invariant mass as normal π^0 is imposed. The requirements on the daughter-photon energy is relaxed to be 22 MeV in the entire ECL region. We additionally require an energy-asymmetry

$$A_\gamma = \frac{E_\gamma^{\rm h} - E_\gamma^{\rm l}}{E_\gamma^{\rm h} + E_\gamma^{\rm l}}, \tag{3.14}$$

to be less than 0.6, where E_h and E_l are the energies of the high- and low-energy photon daughters, respectively, in the laboratory frame.

$\rho^\pm$

Candidate ρ mesons are formed with a $\pi^\pm$ and a π^0 candidates. Figure 3.5 shows the distribution of the ρ-candidate invariant mass $M_{\pi\pi^0}$. The candidates with an invariant mass in $660 < M_{\pi\pi^0} < 960$ MeV/c^2 are selected.

Table 3.7: Example of the π^0 candidate selection based on the candidate ranking. The indices of π^0 and γ indicate the candidate numbers in this event. The candidates with ✓in the "Selected" column are retained in this case. The π^0_2 is discarded because it shares γ_2 with the π^0_1. The π^0_3 is next selected, and the π^0_4 is discarded due to the γ_4 shared with the π^0_3.

Selected	Rank	Candidate	More energetic γ	Less energetic γ
✓	1	π^0_1	γ_1 (100 MeV)	γ_2 (70 MeV)
	2	π^0_2	γ_3 (90 MeV)	γ_2 (70 MeV)
✓	3	π^0_3	γ_3 (90 MeV)	γ_4 (60 MeV)
	4	π^0_4	γ_4 (60 MeV)	γ_5 (50 MeV)

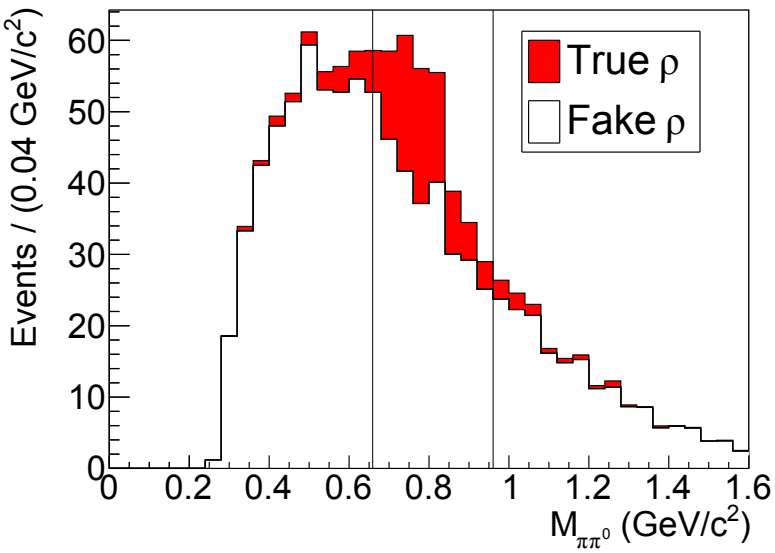


Figure 3.5: Distribution of $M_{\pi\pi^0}$ for the ρ candidates (MC). Vertical black lines define the signal region.

$e^{\pm}$ and $\mu^{\pm}$

Charged leptons $e^{\pm}$ and $\mu^{\pm}$ are selected from charged tracks with $\mathcal{P}_e > 0.9$ and $\mathcal{P}_\mu > 0.9$, respectively.

3.4.2 $D^{(*)}$ Reconstruction

The D candidates are formed by combining the reconstructed particle candidates in Table 3.6. By associating a D candidate with a $\pi^{\pm}$, a π^0 or a γ, D^* candidates are reconstructed. Figures 3.6 and 3.7 show the distributions of the invariant mass M_D of the D candidate and the mass difference $\Delta M \equiv M_{D^*} - M_D$, respectively, where M_{D^*} is the invariant mass of the D^* candidate. In Fig. 3.6 (3.7), ΔM (M_D) is required to be within $\pm 4\sigma_{D^*}$ ($\pm 3\sigma_D$) around the nominal value of the mass difference Δm calculated based on Ref. [10] (the nominal D meson mass m_D in Ref. [10]), where σ_{D^*} (σ_D) denotes the ΔM (M_D) resolution. Since each $D^{(*)}$ mode contains different final state particles, M_D and ΔM resolutions are significantly different depending on $D^{(*)}$ mode, as summarized in Table 3.8. The $D^{(*)}$ modes containing π^0 or γ tend to have worse resolution compared to those containing only $K^{\pm}$, $\pi^{\pm}$ and K^0_S.

The signal-to-noise ratio (SNR) is significantly different among the $D^{(*)}$ modes. To gain the signal significance, it is better to apply a looser (tighter) requirement on the high (low) SNR modes. We optimize the requirements separately for (i) the high- and the low-SNR D modes, and (ii) each D^* mode. Categorization of the high-SNR and low-SNR D modes is as below.

- The low-SNR D^0 mode contains $D^0 \to \pi^+\pi^-, K^-\pi^+\pi^0$ and $K^0_S\pi^-\pi^+\pi^0$.
- The high-SNR D^0 mode contains $D^0 \to K^0_S\pi^0, K^-\pi^+, K^-K^+, K^0_S\pi^-\pi^+$ and $K^-\pi^+\pi^+\pi^-$.
- The low-SNR D^+ mode contains $D^+ \to K^0_S\pi^+\pi^0, K^-\pi^+\pi^+\pi^0, K^-K^+\pi^+$ and $K^0_S\pi^+\pi^+\pi^-$.

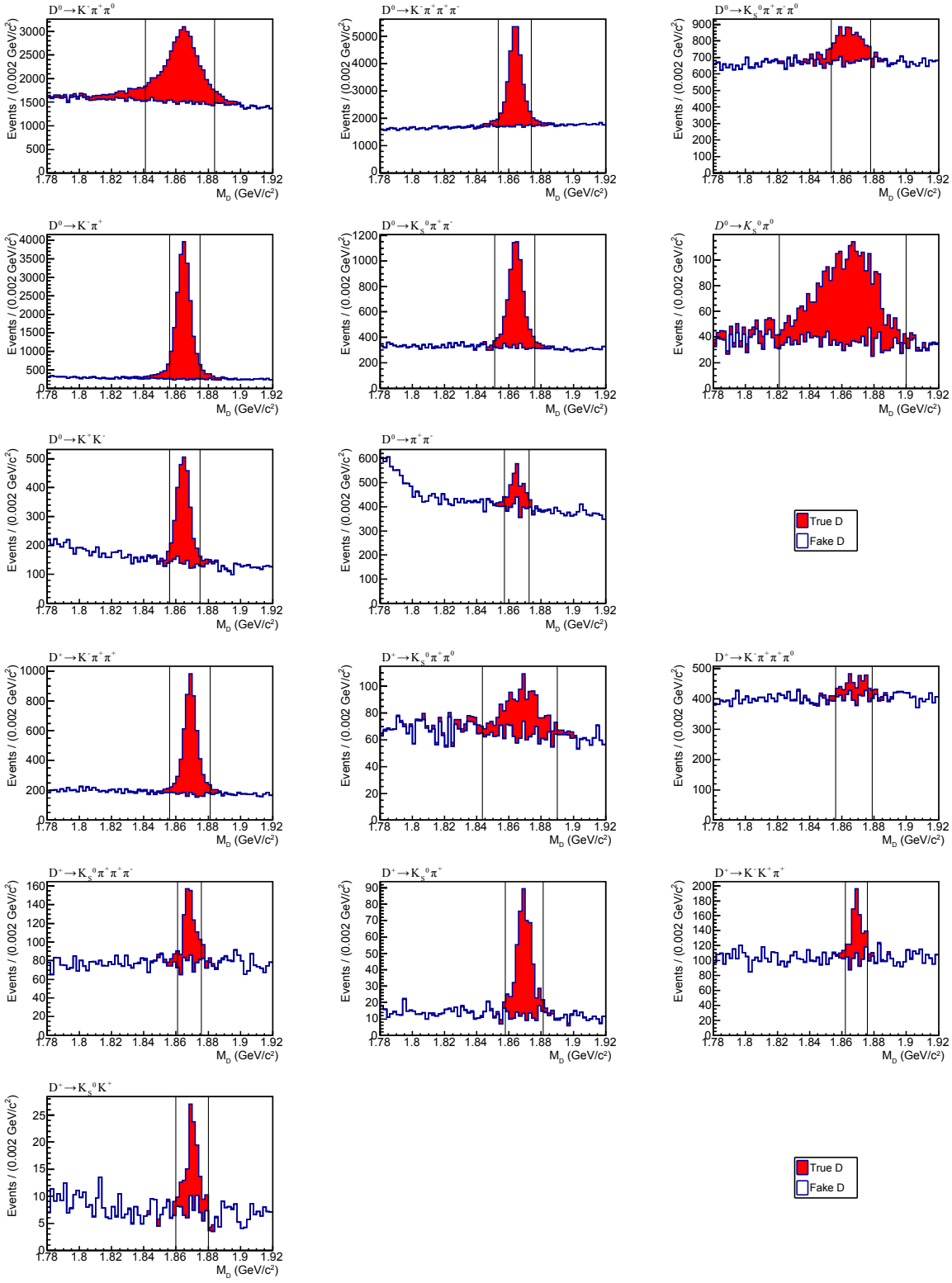


Figure 3.6: Distributions of M_D (MC) for the eight D^0 decays (top half) and the seven D^+ decays (bottom half). Vertical black lines define the signal region in Table 3.9.

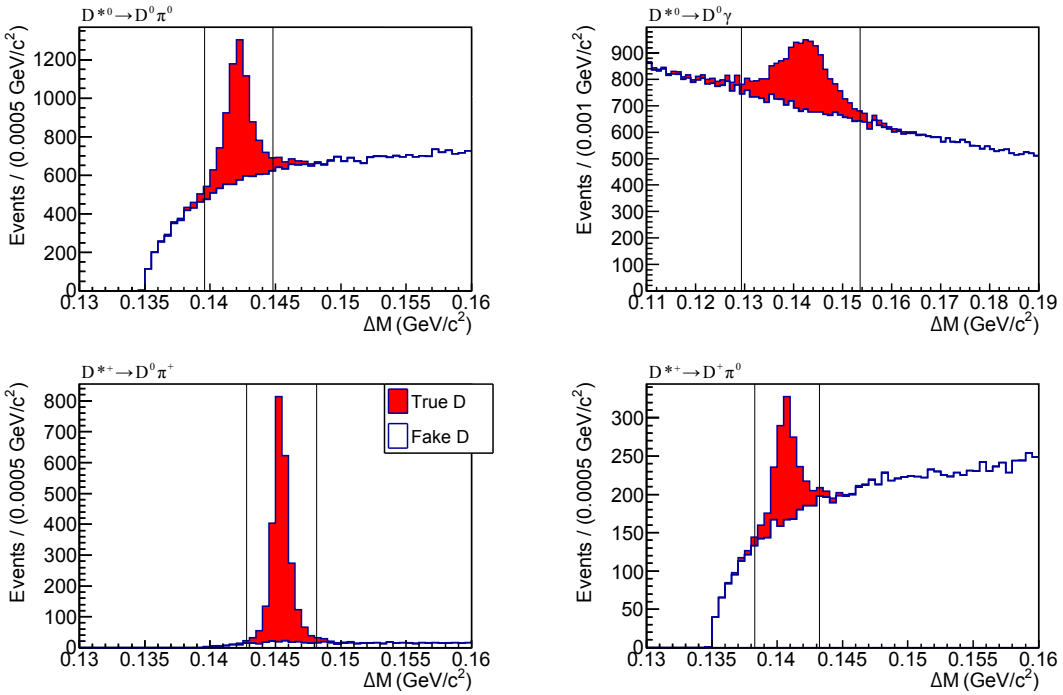


Figure 3.7: Distributions of ΔM (MC) for the two D^{*0} decays (top two) and the two D^{*+} decays (bottom two). Vertical black lines define the signal region in Table 3.9.

- The high-SNR D^+ mode contains $D^+ \to K^0_S\pi^+$ and $K^0_S K^+, K^-\pi^+\pi^+$.

We determine the requirements with respect to σ_D and σ_{D^*} as summarized in Table 3.9.

Table 3.8: Resolutions σ_D and σ_{D^*} for each mode. The column LS (HS) means the value for the lower (higher) side of the M_D or ΔM peak.

Decay mode	LS (MeV/c^2)	HS (MeV/c^2)
$D^+ \to K^-\pi^+\pi^+$	4.7	4.1
$D^+ \to K^0_S\pi^+\pi^0$	14	11
$D^+ \to K^-\pi^+\pi^+\pi^0$	11	8.2
$D^+ \to K^0_S\pi^+\pi^+\pi^-$	5.3	3.7
$D^+ \to K^0_S K^+$	4.0	4.6
$D^+ \to K^0_S\pi^+$	4.7	4.6
$D^+ \to K^-K^+\pi^+$	4.2	3.6
$D^0 \to K^-\pi^+\pi^0$	13	12
$D^0 \to K^-\pi^+\pi^+\pi^-$	4.1	4.3
$D^0 \to K^0_S\pi^+\pi^-\pi^0$	10	9.4
$D^0 \to K^-\pi^+$	4.4	4.5
$D^0 \to K^0_S\pi^+\pi^-$	4.5	4.5
$D^0 \to K^0_S\pi^0$	22	15
$D^0 \to K^-K^+$	4.1	4.1
$D^0 \to \pi^-\pi^+$	4.7	4.7
$D^{*+} \to D^0\pi^+$	0.68	0.70
$D^{*+} \to D^0\pi^+$	0.96	1.1
$D^{*0} \to D^0\pi^0$	1.1	1.2
$D^{*0} \to D^0\gamma$	6.9	6.3

Table 3.9: List of the $D^{(*)}$ requirements, where Δm denotes the nominal Δm values calculated from the D and D^* meson masses in Ref. [10].

D mode	D^* mode	$\lvert M_D - m_D\rvert$	$\lvert\Delta M - \Delta m\rvert$
Low-SNR D^0 mode	$D^{*0} \to D^0\pi^0$	$< 1.5\sigma_D$	$< 2.0\sigma_{D^*}$
High-SNR D^0 mode	$D^{*0} \to D^0\pi^0$	$< 2.0\sigma_D$	$< 2.0\sigma_{D^*}$
Low-SNR D^0 mode	$D^{*0} \to D^0\gamma$	$< 1.5\sigma_D$	$< 1.5\sigma_{D^*}$
High-SNR D^0 mode	$D^{*0} \to D^0\gamma$	$< 2.0\sigma_D$	$< 1.5\sigma_{D^*}$
Low-SNR D^0 mode	$D^{*+} \to D^0\pi^+$	$< 2.0\sigma_D$	$< 3.5\sigma_{D^*}$
High-SNR D^0 mode	$D^{*+} \to D^0\pi^+$	$< 4.0\sigma_D$	$< 3.5\sigma_{D^*}$
Low-SNR D^+ mode	$D^{*+} \to D^+\pi^0$	$< 1.5\sigma_D$	$< 2.0\sigma_{D^*}$
High-SNR D^+ mode	$D^{*+} \to D^+\pi^0$	$< 2.5\sigma_D$	$< 2.0\sigma_{D^*}$

3.4.3 $B_{\rm sig}$ Selection

The $\bar{B}_{\rm sig}$ candidates are formed by associating a τ-daughter meson ($\pi^{\pm}$ or $\rho^{\pm}$) with a D^* candidate. For the normalization mode, a charged lepton is associated instead of a τ daughter. The allowed combinations are

- $D^{*-} + (\pi^+, \rho^+ \text{ or } \ell^+)$ for $B^0_{\rm sig}$,
- $D^{*+} + (\pi^-, \rho^- \text{ or } \ell^-)$ for $\bar{B}^0_{\rm sig}$,
- $\bar{D}^{*0} + (\pi^+, \rho^+ \text{ or } \ell^+)$ for $B^+_{\rm sig}$,
- $D^{*0} + (\pi^-, \rho^- \text{ or } \ell^-)$ for $B^-_{\rm sig}$.

Next, flavors of the $B_{\rm sig}$ and $B_{\rm tag}$ candidates are compared. We allow one of the following combinations: $(B^0_{\rm sig}, \bar{B}^0_{\rm tag})$, $(\bar{B}^0_{\rm sig}, B^0_{\rm tag})$, $(B^+_{\rm sig}, B^-_{\rm tag})$ and $(B^-_{\rm sig}, B^+_{\rm tag})$. If at least one possible candidate is found in an event, the following requirements are imposed on them.

q^2 (for the signal mode)

Signal events must have q^2 greater than $m_\tau^2 c^2 = 3.1\ \text{GeV}^2/c^2$ to produce a τ lepton. Considering that the efficiency is low around the threshold $q^2 = m_\tau^2 c^2$, we require q^2 to exceed $4\ \text{GeV}^2/c^2$. This requirement effectively reduces the background events from the two-body hadronic B decays such as $\bar{B} \to D^* \pi^-$ ($q^2 = 0.02\ \text{GeV}^2/c^2$) and $\bar{B} \to D^* \rho^-$ ($0.6\ \text{GeV}^2/c^2$), or from $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$ ($0 < q^2 < 11\ \text{GeV}^2/c^2$).

$\cos\theta_{\rm hel}$ (for the signal mode)

We require that the events lie in the physical region of $|\cos\theta_{\rm hel}| < 1$, where the most of the signal events distribute. Since π^- or ρ^- is assumed as a τ daughter in the $\cos\theta_{\rm hel}$ calculation, the background events with a prompt π^- or a ρ^- in the B decay do not necessarily have $\cos\theta_{\rm hel}$ in the physical region. The requirement on $\cos\theta_{\rm hel}$ effectively reduces such background events. As shown in Fig. 3.8, there is a significant $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$ peak around $\cos\theta_{\rm hel} \sim 1$ in the $\tau^- \to \pi^- \nu_\tau$ sample due to mis-identification of $\ell^{\pm}$ as $\pi^{\pm}$. To avoid this background, we use the region

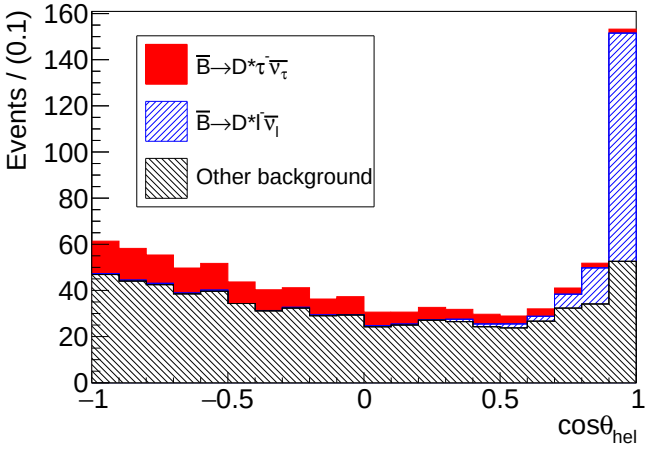


Figure 3.8: Distribution of $\cos\theta_{\rm hel}$ (MC) for the signal (red), $\bar{B} \to D^* \ell \bar{\nu}_\ell$ (blue-hatched) and the other background (black-hatched) in the $\tau^- \to \pi^- \nu_\tau$ sample.

$\cos\theta_{\rm hel} < 0.8$. For the $\tau^- \to \rho^-\nu_\tau$ sample, the entire $\cos\theta_{\rm hel}$ region is used because the $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ background is small compared to the $\tau^- \to \pi^-\nu_\tau$ sample.

$E_{\rm ECL}$ **(for the signal and the normalization modes)**
Both signal and normalization modes are expected to have a smaller value of $E_{\rm ECL}$. It is required to be less than 1.5 GeV.

$M^2_{\rm miss}$ **(for the normalization mode)**
Since the normalization events peak around $M^2_{\rm miss} = 0$ GeV$^2/c^4$, they are required to lie in $-0.5 < M^2_{\rm miss} < 0.5$ GeV$^2/c^4$.

Finally, we require that there be no extra charged tracks with $|dr| < 5$ cm and $|dz| < 20$ cm and normal π^0 candidates.

3.4.4 Best Candidate Selection

After all the signal selection criteria are applied, 8 to 10% (2 to 3%) events have more than one $B^-_{\rm sig}$ ($\bar{B}^0_{\rm sig}$) candidates. There are less multi-candidate events in the $\bar{B}^0$ mode than in the B^- mode because many of the $\bar{B}^0$ events are reconstructed from $D^{*+} \to D^0\pi^+$. Owing to the good ΔM resolution and the smaller number of fake $\pi^\pm$, extra candidates arising from a fake D^{*+} are suppressed. In events where two or more candidates are reconstructed, 2.1 candidates are found on average.

To select the best candidate in an event, the three-step best-candidate selection criteria are applied. Here, it should be noted that multi-$B_{\rm tag}$ candidates are not allowed at this stage since we have selected only one $B_{\rm tag}$ candidate based on $O_{\rm NB}$. Multi-candidate events therefore arise from D^* candidates and τ daughters.

D^* Decay-mode Selection

In the case of D^{*0} candidates, sometimes a $D^{*0} \to D^0\pi^0$ event is reconstructed both as the $D^{*0} \to D^0\pi^0$ mode and the $D^{*0} \to D^0\gamma$ mode. The latter occurs by missing one γ of the soft π^0. We therefore choose the $D^{*0} \to D^0\pi^0$ candidate if one D^{*0} is found in both D^{*0} modes. In the D^{*+} events, an exchange between π^+ in $D^{*+} \to D^0\pi^+$ and π^0 in $D^{*+} \to D^+\pi^0$ rarely happens. The candidate of $D^{*+} \to D^0\pi^+$, which has the larger branching fraction than $D^{*+} \to D^+\pi^0$, is preferentially selected.

D^* Candidate Selection

Even after only one D^* mode is selected, there are still some events having multi candidates reconstructed from the same D^* mode. We select only one D^* candidate among them. For the $D^{*0} \to D^0\pi^0$ and $D^{*+} \to D^+\pi^0$ ($D^{*0} \to D^0\gamma$) modes, the value of the π^0 invariant mass (the photon energy) is used. As shown in Fig. 3.9, true $D^{*0} \to D^0\pi^0$ and $D^{*+} \to D^+\pi^0$ ($D^{*0} \to D^0\gamma$) candidates tend to have a π^0 with an invariant mass close to the nominal π^0 mass (an energetic photon). We select

- a $D^{*0} \to D^0\pi^0$ or $D^{*+} \to D^+\pi^0$ candidate with an invariant mass nearest the nominal π^0 mass, and
- a $D^{*0} \to D^0\gamma$ candidate with the most energetic γ.

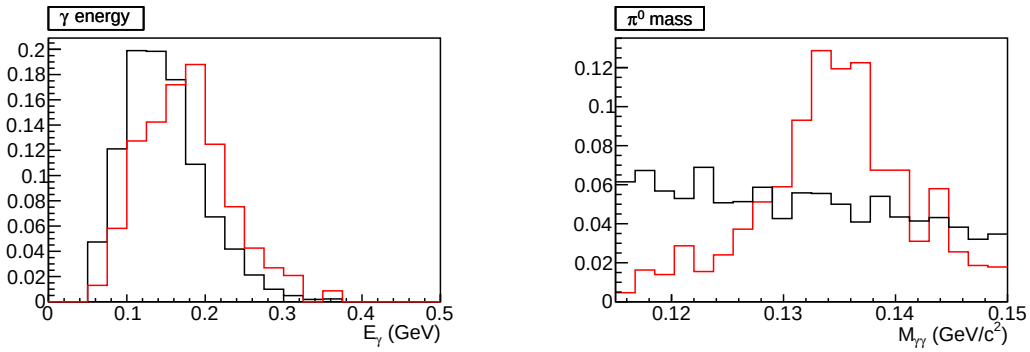


Figure 3.9: Distributions of E_γ for the $D^{*0} \to D^0\gamma$ candidates (left) and of $M_{\gamma\gamma}$ of the soft π^0 from the $D^{*0} \to D^0\pi^0$ and the $D^{*+} \to D^+\pi^0$ candidates (right). The red (black) histogram shows the distribution of true (fake) γ or π^0.

For the $D^{*+} \to D^0\pi^+$ mode, a multi-candidate probability is only 0.05%. We therefore select one candidate randomly if such a multi-candidate event is found.

In the process of the D^*-candidate selection, 69% of true D^* candidates are retained while 58% of fake D^* candidates are discarded. The performance is better than 50% (50%) for the survival (discarded) rate of the true (fake) D^* candidates expected from the random candidate selection.

τ Mode

After the D^* candidate selection, about 2% of the events are found both in the $\tau^- \to \pi^-\nu_\tau$ sample and in the $\tau^- \to \rho^-\nu_\tau$ sample. The MC study indicates that about 80% of such events originate from the $\tau^- \to \rho^-\nu_\tau$ mode and also reconstructed as the $\tau^- \to \pi^-\nu_\tau$ event by missing a π^0 from the ρ^- decay. We assign them to the $\tau^- \to \rho^-\nu_\tau$ sample.

3.4.5 Summary of the Event Reconstruction

Figures 3.10 and 3.11 show the $E_{\rm ECL}$ and the $\cos\theta_{\rm hel}$ distributions, respectively, for the signal sample after the event selection. Events are categorized as below.

Signal

Correctly-reconstructed signal events, which originate from $\tau^- \to \pi^-(\rho^-)\nu_\tau$ events reconstructed correctly as the $\tau^- \to \pi^-(\rho^-)\nu_\tau$ sample, are categorized in this component. The yield is treated as a free parameter associated with $R(D^*)$ and $P_\tau(D^*)$.

$\rho \leftrightarrow \pi$ cross feed

Cross-feed events, where the $\tau^- \to \rho^-\nu_\tau$ events are reconstructed as the $\tau^- \to \pi^-\nu_\tau$ sample due to misreconstruction of one π^0, or the $\tau^- \to \pi^-\nu_\tau$ events are reconstructed in the $\tau^- \to \rho^-\nu_\tau$ sample by adding a random π^0, comprise this component. Since these events originate from $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$, they are used in the $R(D^*)$ determination. They are also used to increase a sensitivity to $P_\tau(D^*)$ after the bias on $P_\tau(D^*)$ is corrected using MC information. This is further discussed is in Chapter 4.

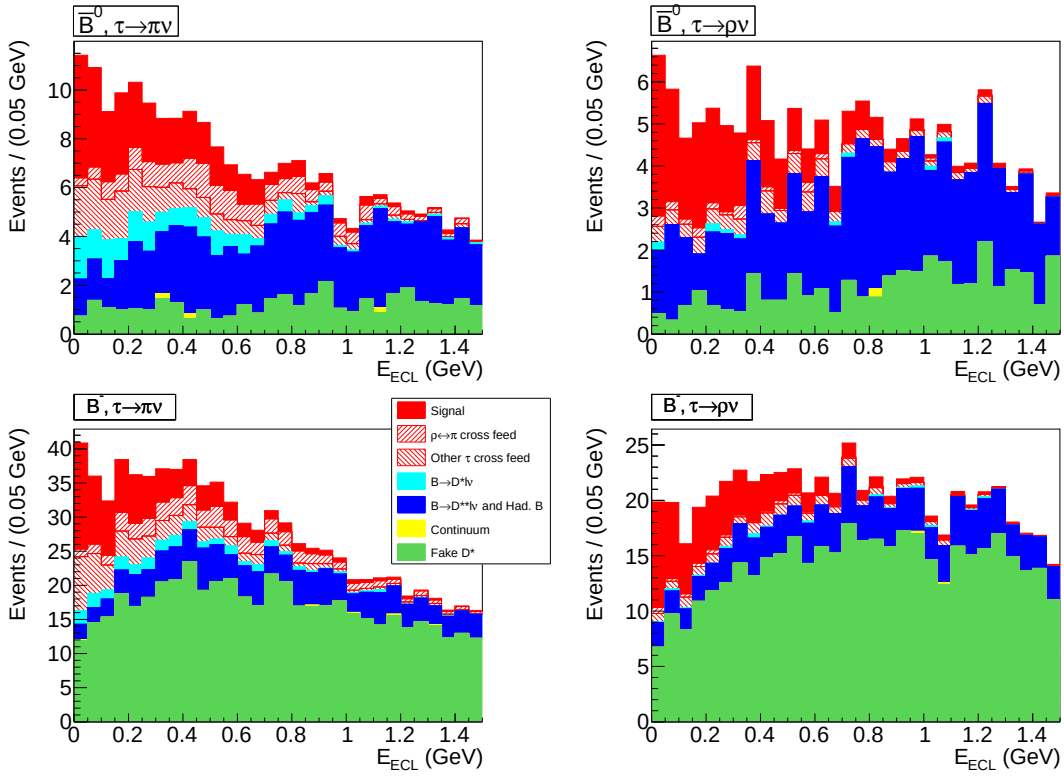


Figure 3.10: Distribution of $E_{\rm ECL}$ in the MC simulation: (top-left) $\bar{B}^0$ and $\tau^- \to \pi^-\nu_\tau$, (top-right) $\bar{B}^0$ and $\tau^- \to \rho^-\nu_\tau$, (bottom-left) B^- and $\tau^- \to \pi^-\nu_\tau$ and (bottom-right) B^- and $\tau^- \to \rho^-\nu_\tau$.

Other τ cross feed

Events of $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ from other τ decay modes also contribute to the signal sample. They originate mainly from $\tau^- \to a_1^-(\to \pi^-\pi^0\pi^0)\nu_\tau$ with one or two missing π^0, or $\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$ with a low-momentum μ^- which does not reach the KLM. There two modes occupy about 80% of this component. These are less sensitive to $P_\tau(D^*)$ since the heavy a_1 mass makes α in Eq. 3.3 almost equal to 0, and events with two neutrinos in the $\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$ mode wash out the $P_\tau(D^*)$ information. In the fit, the yield is associated with $R(D^*)$.

$\bar{B} \to D^*\ell^-\bar{\nu}_\ell$

The decay $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ contaminates the signal sample due to misassignment of ℓ^- as π^-. We fix the $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ yield in the signal sample from the fit to the $M^2_{\rm miss}$ distribution of the normalization sample.

$\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ **and hadronic B decays**

The $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ ($\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau$ is also included in this category) and hadronic B decays are the most uncertain component due to insufficient experimental knowledge. By missing a few particles such as π^0, the event topology resembles the signal event. We combine these decay modes into the same component. The fractions of the $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ decays and hadronic B decays in this category are about 10% and 90%, respectively, according to the MC study. Since it is difficult to estimate the yield of this component using the MC simulation or to fix the yield using control

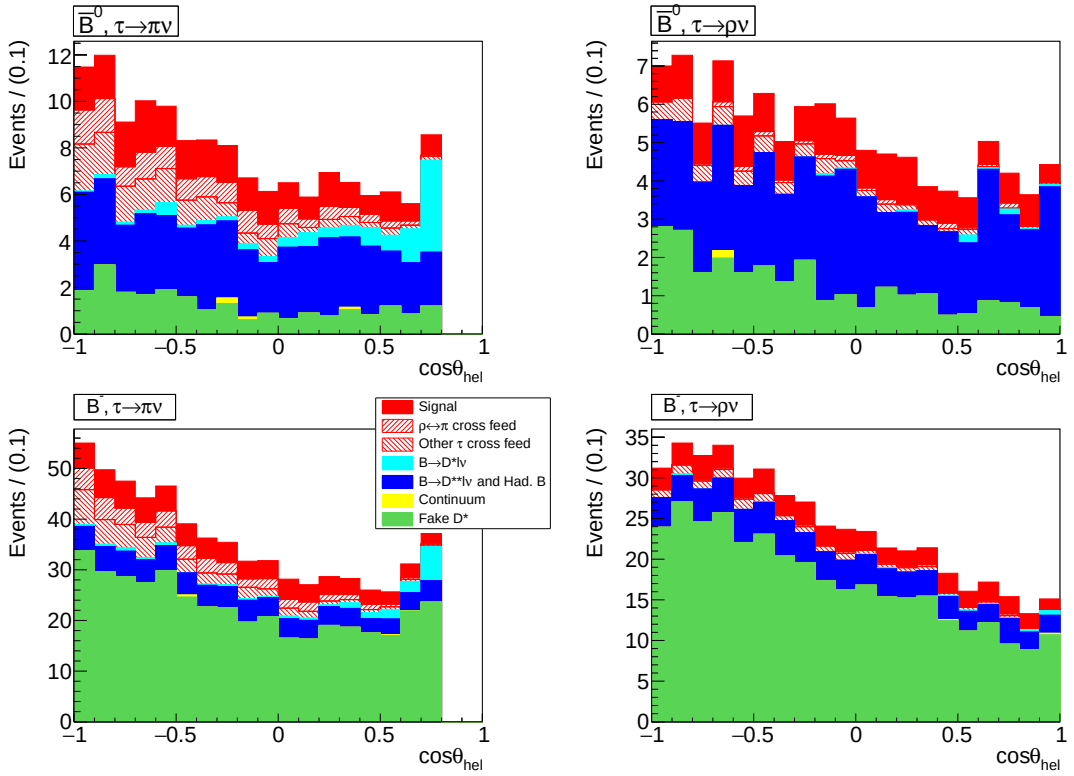


Figure 3.11: Distribution of $\cos\theta_{\mathrm{ECL}}$ in the MC simulation: (top-left) $\bar{B}^0$ and $\tau^- \to \pi^-\nu_\tau$, (top-right) $\bar{B}^0$ and $\tau^- \to \rho^-\nu_\tau$, (bottom-left) B^- and $\tau^- \to \pi^-\nu_\tau$ and (bottom-right) B^- and $\tau^- \to \rho^-\nu_\tau$.

data samples, we float the yield in the final fit. One exception is the modes with two D mesons such as $\bar{B} \to D^* D_s^{(*)-}$ and $\bar{B} \to D^* \bar{D}^{(*)} K^-$. The branching fractions of these modes are experimentally studied. We fix their yield based on the MC expectation.

Continuum

Continuum events from the $e^+e^- \to q\bar{q}$ process provide a minor contribution at $\mathcal{O}(0.1\%)$. We fix the yield using the MC expectation.

Fake D^*

All events containing fake D^* candidates are categorized in this component. This is the main background source in the B^- sample. For the $\bar{B}^0$ sample, many D^{*+} candidates are reconstructed from combination of a D^0 with a soft $\pi^\pm$, which is much cleaner than the D^* modes with π^0 or γ. The fake D^* component is therefore less significant. The yield is determined from the comparison of the data and the MC sample in the ΔM sideband region.

In Fig. 3.10, the fluctuation on the MC distributions is significant due to the limited MC statistics. Possible bias in the fit is evaluated later in Sec. 5.2.

Figure 3.12 shows the M^2_{miss} distribution for the normalization sample. As expected, the clear peaks of the normalization events is observed around $M^2_{\mathrm{miss}} = 0$ GeV$^2/c^4$.

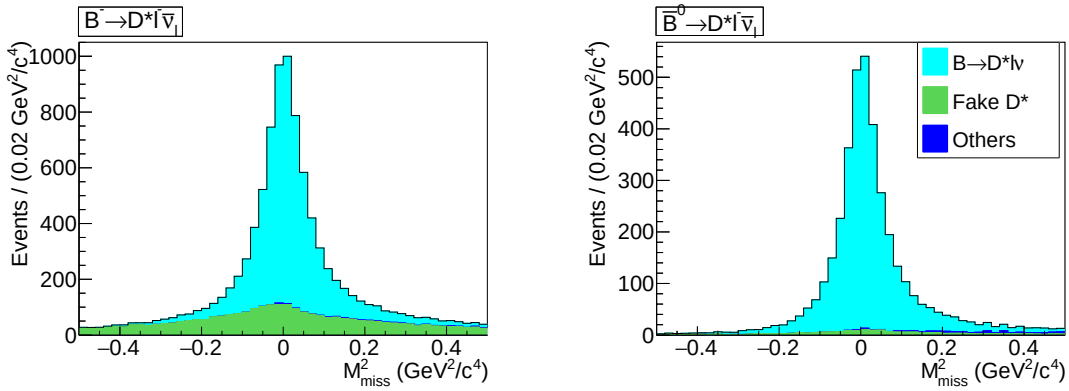


Figure 3.12: Distribution of $M^2_{\rm miss}$ distribution in the normalization sample (MC simulation). The left and the right panels are the B^- sample and $\bar{B}^0$ sample, respectively.

3.5 MC Calibration

In the previous section, we obtain the distributions of the signal and the normalization samples using the MC simulation. However, the MC simulation does not always reproduce the data perfectly. We therefore calibrate and validate the MC distributions using the calibration data sample.

3.5.1 Tag Efficiency Correction

To correct the tag efficiency, the decays $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ are reconstructed as the signal-side mode, as listed in Table 3.10. The signal is extracted using the $M^2_{\rm miss}$ distribution. A correction factor is obtained as a ratio of the yield of the data to that in the MC sample. Details of the method are described in Ref. [74]. The correction factor for the tag efficiency is found to be 0.73 (0.71) for the B^- ($\bar{B}^0$) sample.

We assign the systematic uncertainty of 4.6% for the B^- sample and 5.1% for the $\bar{B}^0$ sample. It mainly arises from the statistics of the calibration samples, the uncertainties in the branching fractions of $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ and the D meson decays, and the particle reconstruction efficiencies for $K^\pm$, $\pi^\pm$, π^0 and $\ell^\pm$. In the normalization mode, the uncertainty in the lepton ID is common between the tag efficiency and the signal reconstruction

Table 3.10: Calibration modes for the tag efficiency correction.

B^- sample	$\bar{B}^0$ sample
$B^- \to D^0\ell^-\bar{\nu}_\ell$	$\bar{B}^0 \to D^+\ell^-\bar{\nu}_\ell$
$D^0 \to K^-\pi^+$	$D^+ \to K^-\pi^+\pi^+$
$D^0 \to K^-\pi^+\pi^0$	$D^+ \to K^-\pi^+\pi^+\pi^0$
$D^0 \to K^-\pi^+\pi^+\pi^-$	$D^+ \to K^-\pi^+\pi^+\pi^+\pi^-$
$B^- \to D^{*0}\ell^-\bar{\nu}_\ell$	$\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$
$D^{*0} \to D^0\pi^0$, $D^0 \to K^-\pi^+$	$D^{*+} \to D^0\pi^+$, $D^0 \to K^-\pi^+$
$D^{*0} \to D^0\gamma$, $D^0 \to K^-\pi^+$	$D^{*+} \to D^+\pi^0$, $D^+ \to K^-\pi^+\pi^+$

efficiency. To avoid double counting of the same systematic uncertainty, we exclude the uncertainty in the lepton ID efficiency from the uncertainty in the tag efficiency correction. The systematic uncertainty in the tag efficiency is estimated to be 4.2% (4.6%) for the B^- ($\bar{B}^0$) sample for the normalization mode.

3.5.2 Resolution Correction

M_D and ΔM Resolution

In Sec. 3.4.2, we have defined the $D^{(*)}$-mode-dependent signal regions by 1.5 to $4\sigma_D$ and the 1.5 to $3.5\sigma_{D^*}$ around the M_D and ΔM peaks, respectively. Since, as shown in Fig. 3.13, the M_D and the ΔM resolutions in the MC simulation are slightly different from that in the data, these requirements possibly cause significant efficiency bias.

The resolution in the MC sample is corrected to match to that in the data for each $D^{(*)}$ decay. First, we fit the PDFs in Table 3.11 to the data and the MC simulation to obtain a ratio of the resolutions. The shape of the true $D^{(*)}$ distribution is modeled by three Gaussians. For the D modes containing π^0, one of the Gaussian component is replaced by the empirically parameterized function developed by the Crystal Ball Collaboration [75] (Crystal Ball function) to take into account the lower M_D tail, which is visible in Fig. 3.13 (left). In the fake $D^{(*)}$ component, a linear function or a quadratic function is used for the modes except for $D^* \to D\pi$, and an empirical threshold function $(\Delta M - 0.140)^a e^{-b\Delta M}$

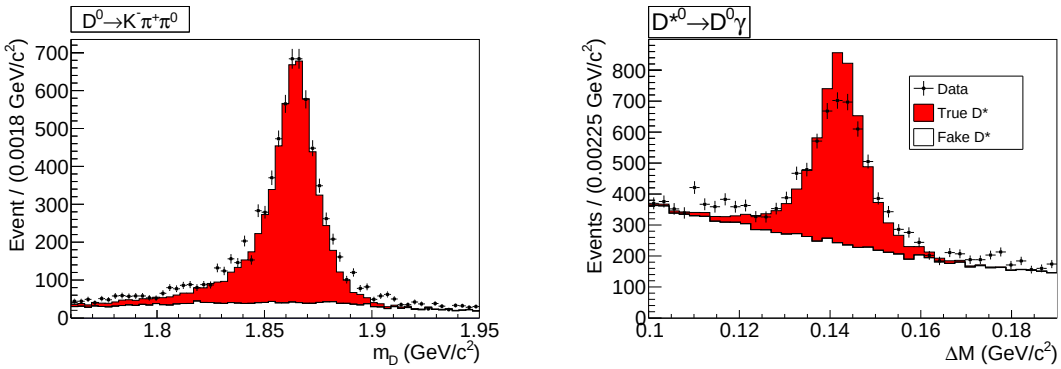


Figure 3.13: Examples of the comparisons for the M_D distribution in the $D^0 \to K^-\pi^+\pi^0$ mode (left) and the ΔM distribution in the $D^{*0} \to D^0\gamma$ mode (right) between the data and the MC sample. The solid histogram and the black dots are the distributions of the MC and the data, respectively.

Table 3.11: Signal and background PDFs for the M_D and the ΔM distributions. (The C.B. function indicates the Crystal Ball function.)

Decay mode	True $D^{(*)}$ shape	Fake $D^{(*)}$ shape
D decays with π^0	C.B. function + Double Gaussian	Linear function
D decays without π^0	Triple Gaussian	Linear function
$D^{*0} \to D^0\gamma$	Triple Gaussian	Quadratic function
Other D^* decays	Triple Gaussian	Threshold function

(where a and b are the free parameters) is used to model the ΔM shape with a threshold at the pion mass in the $D^* \to D\pi$ modes. After the function parameters except for the M_D or the ΔM resolution are fixed by the fits to the MC sample, a fit to the data is performed. One example of the fit results is shown in Fig. 3.14. The resolution ratio

$$r_{\rm res} \;=\; \frac{\sigma_i^{\rm data}}{\sigma_i^{\rm MC}}, \tag{3.15}$$

is extracted from the fit. Here, σ_i denotes the standard deviation of the ith Gaussian or Crystal Ball component in the true $D^{(*)}$ PDF. Figure 3.15 summarizes the obtained values of $r_{\rm res}$ for each $D^{(*)}$ mode. The resolution discrepancy by up to 25% is observed. The uncertainty in the $D^{(*)}$ selection efficiency due to the errors in $r_{\rm res}$ is estimated to be about 1% and sufficiently small.

$M^2_{\rm miss}$ Resolution

Figure 3.16 shows the comparison of the $M^2_{\rm miss}$ distribution between the data and the MC simulation for the normalization sample. We observe a slight difference in the $M^2_{\rm miss}$ peak width. This difference presumably originates from a discrepancy in the momentum resolution of low-momentum pions and photons. We correct the resolution discrepancy

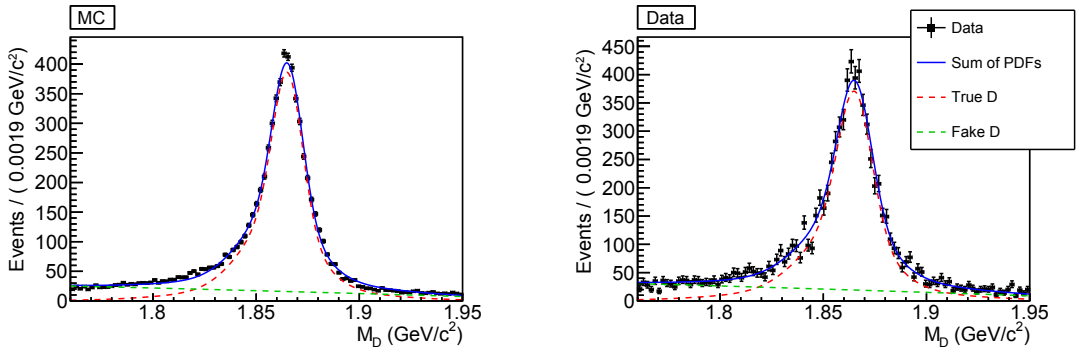


Figure 3.14: Fit result to the $D^0 \to K^-\pi^+\pi^0$ sample in the MC simulation (left) and the data (right). The dashed red and green lines show the $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ signal and the fake D^* components, respectively, and the solid blue line represents the sum of the two components.

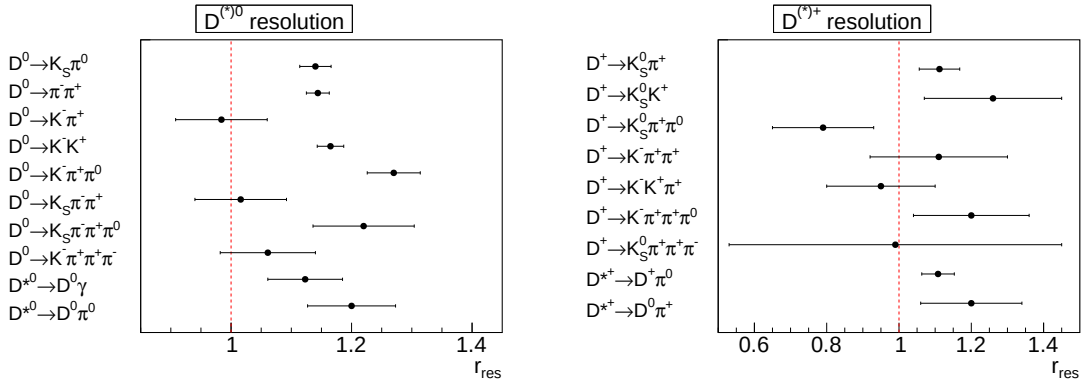


Figure 3.15: Obtained $r_{\rm res}$ for each $D^{(*)}$ mode.

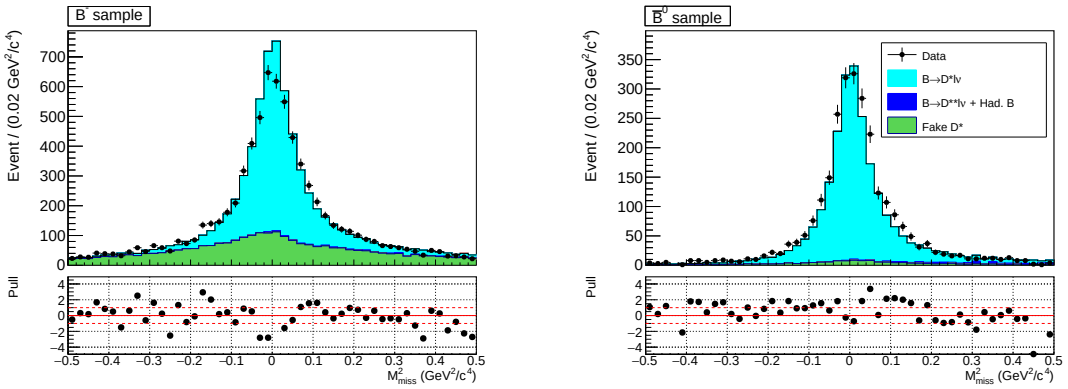


Figure 3.16: Comparison of the $M^2_{\rm miss}$ distribution between the data and the MC sample (top) and their pull in each bin (bottom).

Table 3.12: Resolution calibration factors $R^i_{\rm res}$ for the $M^2_{\rm miss}$ resolution.

D^* mode	$R^i_{\rm res}$
$D^{*0} \to D^0\pi^0$	1.21 ± 0.08
$D^{*0} \to D^0\gamma$	1.23 ± 0.08
$D^{*+} \to D^0\pi^+$	1.11 ± 0.03
$D^{*+} \to D^+\pi^0$	1.21 ± 0.15

depending on the D^* mode. The method for the $M^2_{\rm miss}$ resolution correction is similar to that applied to the M_D and ΔM resolution correction; we extract the resolution calibration factors $R^i_{\rm res}$ (where i denotes the index of the D^* mode) from the fit to the $M^2_{\rm miss}$ distribution. A model function is constructed using two or three Gaussians for the $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ and for the fake D^* components, respectively. Since contribution from the other background components is negligibly small, we consider only these two components. The fit result to the $D^{*0} \to D^0\pi^0$ sample is shown in Fig. 3.17. The function parameters except for the $M^2_{\rm miss}$ resolution are determined in the fit to the MC sample, and the obtained function is fitted to the data to extract the resolution ratio using Eq. 3.15 (with replacing $r_{\rm res}$ by $R^i_{\rm res}$). The obtained values of $R^i_{\rm res}$ are summarized in Table 3.12.

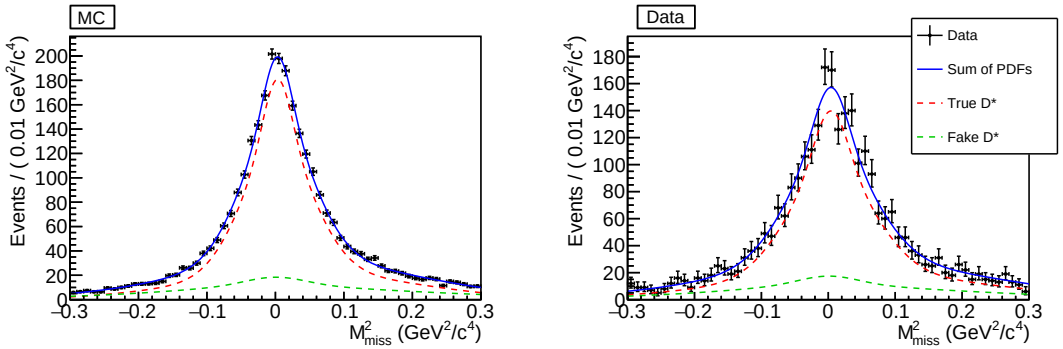


Figure 3.17: Fit result to the $M^2_{\rm miss}$ distribution of the $D^{*0} \to D^0\pi^0$ mode in the MC sample (left) and the data (right). Meaning of each line is the same as Fig. 3.14.

The factor $R^i_{\rm res}$ is applied event-by-event to the $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$ events of the normalization sample in the MC simulation as

$$(M^{\rm corr}_{\rm miss})^2 = R^i_{\rm res}(M^{\rm org}_{\rm miss})^2, \tag{3.16}$$

where $(M^{\rm corr}_{\rm miss})^2$ and $(M^{\rm org}_{\rm miss})^2$ denote the corrected and the original values of $M^2_{\rm miss}$, respectively. The histogram PDFs is then constructed from $(M^{\rm corr}_{\rm miss})^2$.

3.5.3 Fake D^* Yield

The fake D^* yield in the signal region is scaled from the comparison between the data and the MC sample in the ΔM sideband region. The sideband regions are defined as

- $0.140 < \Delta M < 0.142, 0.148 < \Delta M < 0.170$ GeV/c^2 for $D^{*+} \to D^0\pi^+$,
- $0.135 < \Delta M < 0.136, 0.146 < \Delta M < 0.170$ GeV/c^2 for $D^{*+} \to D^+\pi^0$,
- $0.135 < \Delta M < 0.137, 0.148 < \Delta M < 0.170$ GeV/c^2 for $D^{*0} \to D^0\pi^0$,
- $0.100 < \Delta M < 0.108, 0.173 < \Delta M < 0.190$ GeV/c^2 for $D^{*0} \to D^0\gamma$,

where about $\pm 4\sigma_{D^*}$ region from the nominal ΔM value is excluded. To extract the fake D^* yield, as shown in Fig. 3.18, we fit the background model functions in Table 3.11 in the sideband region. An yield calibration factor is obtained as a ratio of the yield in the data to that in the MC sample. The extracted calibration factors are summarized in Table 3.13.

3.5.4 Composition of Hadronic B Decays

Since each B decay mode has a different $E_{\rm ECL}$ shape, it is essential to calibrate the composition of the hadronic B background. We first divide the hadronic B modes into

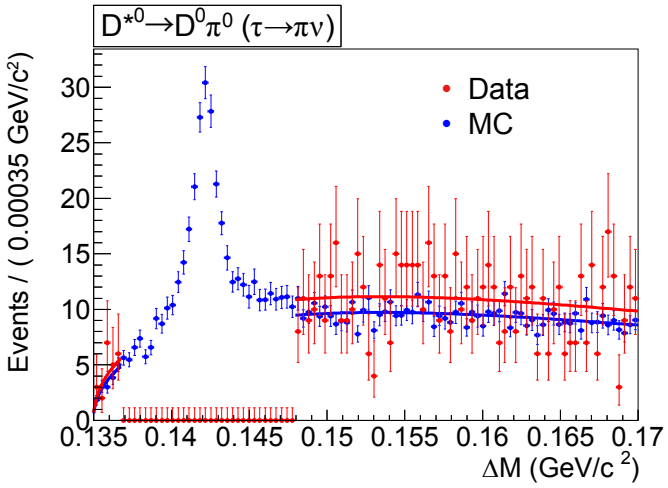


Figure 3.18: Comparison of the ΔM distribution in the sideband regions between the data (red dots) and the MC sample (blue dots). The fit results with the background model functions are indicated by the solid lines. The $D^{*0} \to D^0\pi^0$ mode in the $\tau^- \to \pi^-\nu_\tau$ sample is shown as an example. For the data, the signal region is masked.

Table 3.13: Calibration factors for the fake D^* yield.

	$\tau^- \to \pi^-\nu_\tau$	$\tau^- \to \rho^-\nu_\tau$	$\bar{B} \to D^*\ell^-\bar{\nu}_\ell$
$D^{*0} \to D^0\pi^0$	1.15 ± 0.07	1.14 ± 0.08	1.14 ± 0.03
$D^{*0} \to D^0\gamma$	1.05 ± 0.06	1.11 ± 0.06	1.08 ± 0.03
$D^{*+} \to D^0\pi^+$	0.75 ± 0.20	0.98 ± 0.27	1.04 ± 0.17
$D^{*+} \to D^+\pi^0$	0.99 ± 0.13	1.18 ± 0.15	1.15 ± 0.07

three sub-categories based on their similarity: the two-D mode, the K^0_L mode and the other hadronic B modes. As shown in Fig. 3.19, each sub-category has a different $E_{\rm ECL}$ shape. Table 3.14 summarizes the fraction of each category estimated based on the MC simulation.

Two-D mode

The main modes in this category are $\bar{B} \to D^* D_s^{(*)-}$, $\bar{B} \to D^* D^{(*)-} \bar{K}^0$ and $\bar{B} \to D^* \bar{D}^{(*)0} K^-$. Most of these decay modes are well measured experimentally. We therefore correct the branching fractions in the MC simulation based on the measured value

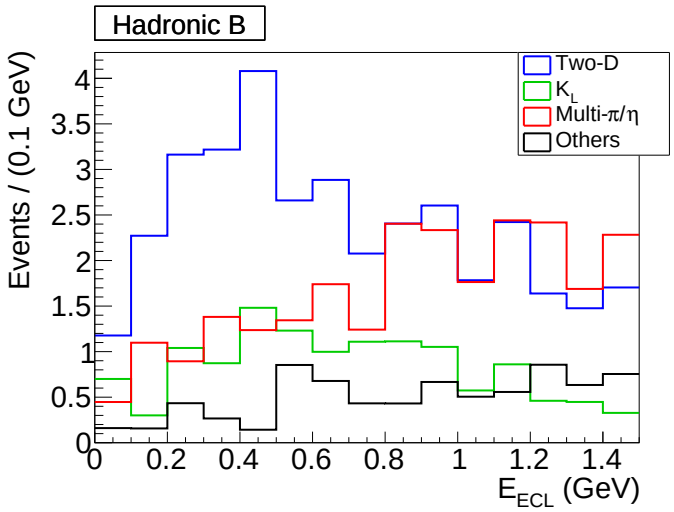


Figure 3.19: Comparison of the $E_{\rm ECL}$ shapes in the four categories of the hadronic B decay modes. For the purpose of visualization, twice wider bins than the actual PDF is used.

Table 3.14: Fraction of each category to the total amount of hadronic B events.

	$\tau^- \to \pi^-\nu_\tau$	$\tau^- \to \rho^-\nu_\tau$)
Two-D	43%	40%
K^0_L	15%	13%
Multi-π/η mode	32%	24%
Other decay modes	9%	22%

and assign the uncertainty in the branching fractions based on the errors reported in Ref. [10].

K_L^0 mode

The K_L^0 mode category contains B decay modes with at least one K_L^0 in the final state such as $\bar{B} \to D^* \pi^- K_L^0$ and $\bar{B} \to D^* K^- K_L^0$. (An event in which K_L^0 is produced through a decay of the D or the D_s meson is included in the two-D mode.) We correct the branching fractions in the MC simulation based on Ref. [10, 76]. For the decay modes without experimental measurements, we assign 100% uncertainty in the branching fraction in the MC simulation.

Multi-π/η mode

Remaining background modes mainly have a pair of missing charged pions and/or missing neutral particles such as π^0 and η. Since various B decay modes contribute to this component, it is difficult to treat them only with the branching fraction correction. Our strategy is, using the signal-side particles, to reconstruct the seven calibration modes $\bar{B} \to D^* \pi^- \pi^- \pi^+$, $D^* \pi^- \pi^- \pi^+ \pi^0$, $D^* \pi^- \pi^- \pi^+ \pi^0 \pi^0$, $D^* \pi^- \pi^0$, $D^* \pi^- \pi^0 \pi^0$, $D^* \pi^- \eta$ and $D^* \pi^- \eta \pi^0$ in the same kinematic region as the signal, and extract the yields from the the data and the MC sample. A calibration factor is then obtained as a ratio of the yields.

The selection criteria for the calibration samples are as below.

- $O_{\rm NB} > 0.001$, $M_{\rm bc} > 5.24$ GeV/c^2 and $|\Delta E| < 200$ MeV for the $B_{\rm tag}$ selection.
- $|M_D - m_D| < 3\sigma_D$ and $|\Delta M - \Delta m| < 3\sigma_{D^*}$ for the D^* selection.
- $500 < M_{\gamma\gamma} < 600$ MeV/c^2 for the η selection. The invariant mass distribution for the η candidates by the MC simulation is shown in Fig. 3.20.
- $q^2 > 4$ GeV$^2/c^4$, $|\cos\theta_{\rm hel}| < 1$ and $E_{\rm ECL} < 1.5$ GeV. To calculate $\cos\theta_{\rm hel}$, we regard (one of) the pion(s) as a τ daughter.

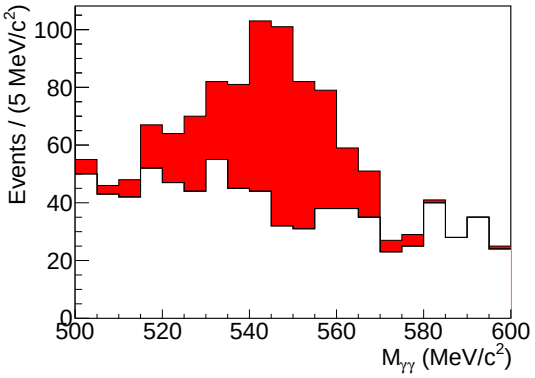


Figure 3.20: Distribution of $M_{\gamma\gamma}$ for the η candidates (MC simulation). To enhance the number of events for visibility, the candidates are selected with a looser requirement $|\Delta E^{\rm sig}| < 500$ MeV.

- $|\Delta E^{\rm sig}| < 100$ MeV or $M_{\rm bc}^{\rm sig} > 5.27$ GeV/c^2, where $\Delta E^{\rm sig}$ and $M_{\rm bc}^{\rm sig}$ are calculated from the signal-side B candidate.

The obtained $M_{\rm bc}^{\rm sig}$ and $\Delta E^{\rm sig}$ distributions from the $\bar{B} \to D^*\pi^-\eta$ and the $\bar{B} \to D^*\pi^-\pi^-\pi^+\pi^0$ samples are illustrated in Fig. 3.21. Especially in the $\bar{B} \to D^*\pi^-\eta$ sample, a significant discrepancy between the data and the MC sample is observed. In the MC simulation, this mode is produced by the hadronization in PYTHIA with a sizable uncertainty. The fit results to Fig. 3.21 are shown in Fig. 3.22. Details of the fit procedure and the fit results to the other samples are summarized in Appendix E. If no signal event is observed, a 68% C.L. upper limit is assigned to the signal yield. The obtained calibration factors are summarized in Table 3.15.

Other decay modes

For the further minor components not contained in the above three categories such as $\bar{B} \to D^*\omega\pi^-$ and $\bar{B} \to D^*\bar{p}n$, we correct the branching fractions based on the experimental measurements [10].

Table 3.16 shows the remaining B decay modes according to the MC simulation. Most of them contain multiple missing neutral particles such as π^0, η and n. It is therefore impossible to reconstruct these modes from the data. We use the branching fractions in the MC simulation, and assign 100% uncertainties. To evaluate the validity of the 100%-uncertainty assumption, we compare the branching fractions in the MC simulation with that reported in Ref. [10] using some B decay modes, which have been experimentally measured, and have similar multiplicity and the final-state particles to the modes in Table 3.16. We choose five multi-pion modes and two baryonic modes as shown in Fig. 3.23. As all the modes have the data-MC ratios within 2, we conclude that our assumption is acceptable.

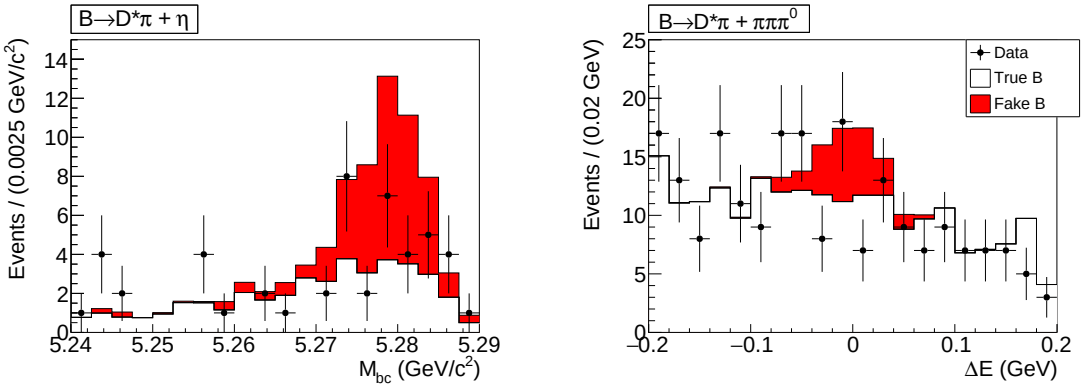


Figure 3.21: Examples of the $M_{\rm bc}^{\rm sig}$ distribution in the $B^- \to D^{*0}\pi^-\eta$ sample (left) and the $\Delta E^{\rm sig}$ distribution in the $B^- \to D^{*0}\pi^-\pi^-\pi^+\pi^0$ sample. The solid histogram and the black dots show the MC and the data distributions, respectively.

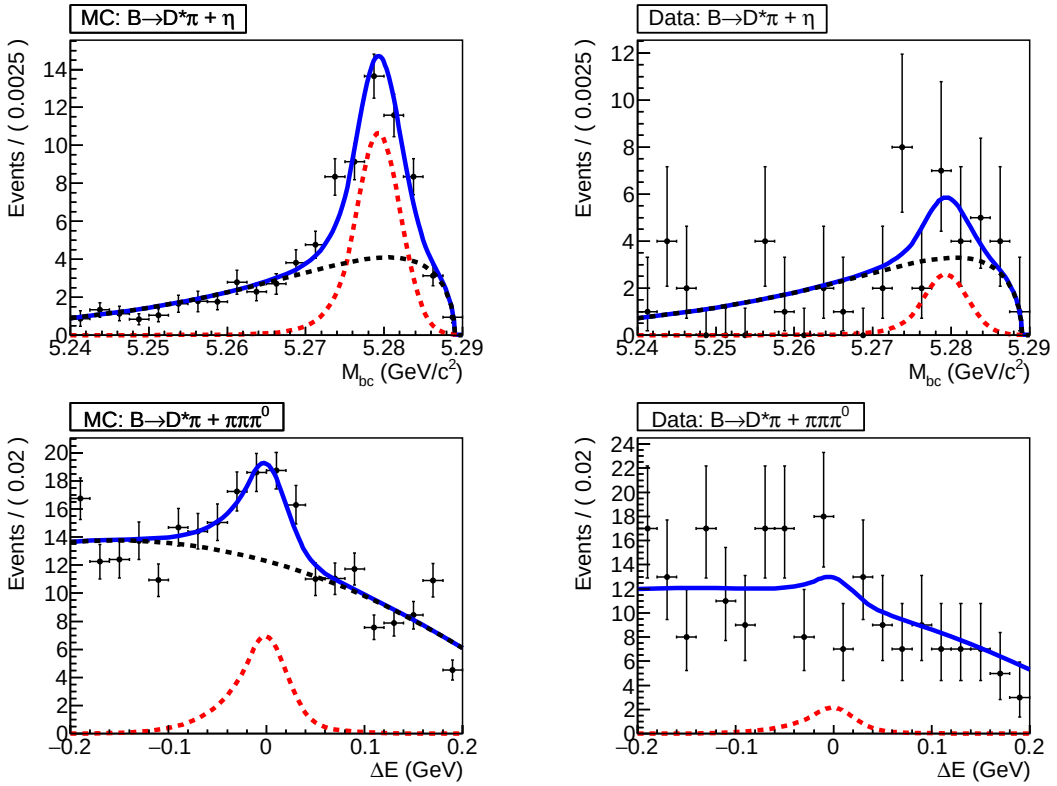


Figure 3.22: Fit results to the distributions in Fig. 3.21. The left and the right panels show the distributions in the MC simulation and the data, respectively. The black dots are the data distribution, and the solid blue, dashed red and the dashed black lines are the sum of the fitted function, the signal and the background components, respectively.

Table 3.15: Calibration factors for the multi-π/η hadronic B modes.

Calibration Mode	B^- sample	$\bar{B}^0$ sample
$D^*\pi^-\pi^-\pi^+$	< 0.51	$0.62^{+0.67}_{-0.49}$
$D^*\pi^-\pi^-\pi^+\pi^0$	$0.31^{+0.43}_{-0.40}$	$0.59^{+0.45}_{-0.39}$
$D^*\pi^-\pi^-\pi^+\pi^0\pi^0$	$2.15^{+1.70}_{-1.60}$	$2.60^{+6.95}_{-2.24}$
$D^*\pi^-\pi^0$	$0.06^{+0.33}_{-0.28}$	< 0.47
$D^*\pi^-\pi^0\pi^0$	$0.09^{+1.04}_{-0.98}$	$1.63^{+0.74}_{-0.69}$
$D^*\pi^-\eta$	$0.24^{+0.21}_{-0.18}$	$0.15^{+0.16}_{-0.10}$
$D^*\pi^-\eta\pi^0$	$0.74^{+0.79}_{-0.75}$	$0.89^{+1.04}_{-0.88}$

Table 3.16: Main decay modes of the remaining B decay modes. The column "Fraction" indicates the fraction of each mode to the total amount of the hadronic B background in the MC sample.

Decay mode	Fraction
Multiple light meson mode	
$\bar{B} \to D^* \pi^- \pi^0 \pi^0 \pi^0$	3.3%
$\bar{B} \to D^* \pi^- \pi^0 \pi^0 \gamma$	1.2%
$\bar{B} \to D^* \pi^- \pi^0 \pi^0 \pi^0 \pi^0 \pi^0$	0.5%
Total	13%
Baryonic mode	
$\bar{B} \to D^* \pi^- n \bar{n}$	1.1%
$\bar{B} \to D^* \pi^- \pi^0 n \bar{n}$	0.8%
Total	2.6%

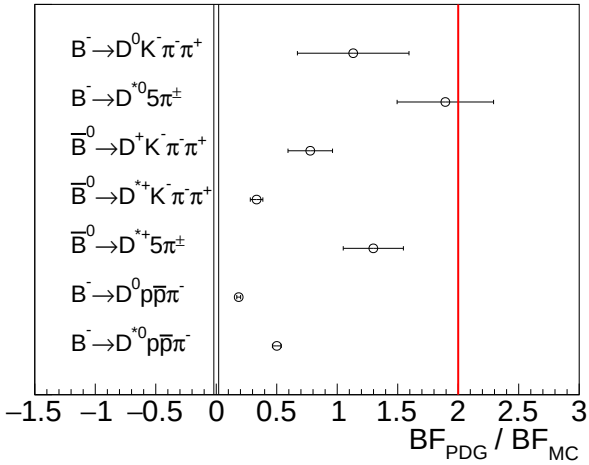


Figure 3.23: Ratio of the branching fractions in Ref. [10] and in the MC simulation.

3.6 Validation

Since the $E_{\rm ECL}$ shape is mostly determined by background photons, its validation is important. We investigate the $E_{\rm ECL}$ shapes of the signal and the fake D^* components, which are the main components in our signal sample. For the fake D^* component, the shape of the $\cos\theta_{\rm hel}$ in the signal sample and the $M^2_{\rm miss}$ distribution in the normalization sample are also checked.

3.6.1 $E_{\rm ECL}$ Distribution

Signal component

To validate the $E_{\rm ECL}$ shape of the signal component, we use the normalization mode as the control sample. It has similar $E_{\rm ECL}$ properties to the signal component: no extra photon from the $B_{\rm sig}$ decay and the $E_{\rm ECL}$ shape is determined by the background photons. This sample contains roughly 50 times more events than the expected number of signal events. Figure 3.24 shows the $E_{\rm ECL}$ comparison between the data and the MC sample. The fake D^* yield is scaled based on the calibration in Sec. 3.5.3, and the $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$ yield is scaled so that the total numbers of events become the same between the data and the MC simulation. The pull distribution is indicated in the bottom panels of Fig. 3.24. Since the data and the MC sample show good agreement, we conclude that the PDF shape of the signal component in the MC simulation well reproduces the data.

Fake D^* component

The PDF shape of the fake D^* component in the data and the MC sample are compared using the ΔM sideband region defined in Table 3.17. We have at least 5 times more events than the number of fake D^* events in the signal region. Figure 3.25 shows the $E_{\rm ECL}$ shape comparison for the D^{*+} and the D^{*0} modes. The figures in the separate comparison for each signal sample are shown in Appendix F. Within the statistical uncertainty, the distribution in the MC sample agrees with the data. We thus apply no correction for the PDF shape.

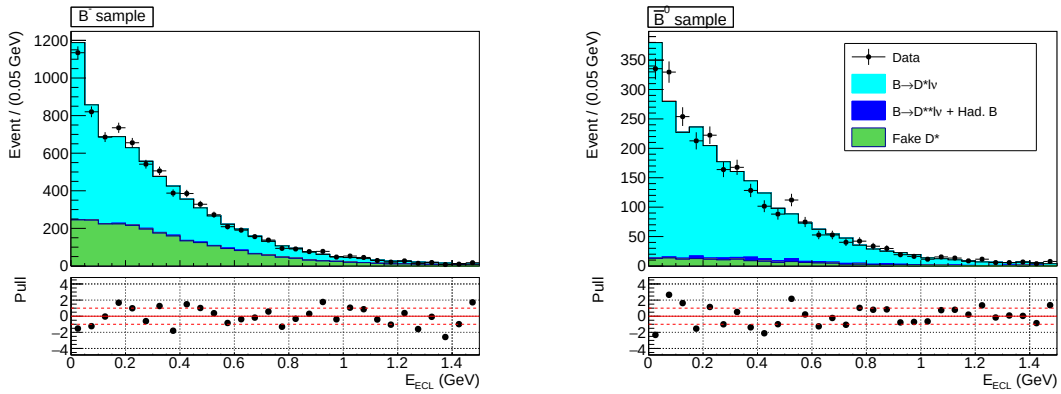


Figure 3.24: Comparison of the $E_{\rm ECL}$ distribution of the normalization mode between the data and the MC (top) and their pull value in each bin (bottom panel).

Table 3.17: Definition of the ΔM sideband region for the validation of the fake D^* component. The column "Stat." shows the rough statistics of the fake D^* events in the sideband region relative to that in the signal region.

D^* decay	Lower sideband (GeV/c^2)	Upper sideband (GeV/c^2)	Stat.
$D^{*+} \to D^0\pi^+$	$0.140 < \Delta M < 0.143$	$0.148 < \Delta M < 0.500$	$\times 47$
$D^{*+} \to D^+\pi^0$	$0.135 < \Delta M < 0.137$	$0.144 < \Delta M < 0.190$	$\times 10$
$D^{*0} \to D^0\pi^0$	$0.135 < \Delta M < 0.138$	$0.146 < \Delta M < 0.190$	$\times 8.5$
$D^{*0} \to D^0\gamma$	$0.050 < \Delta M < 0.117$	$0.166 < \Delta M < 0.500$	$\times 4.7$

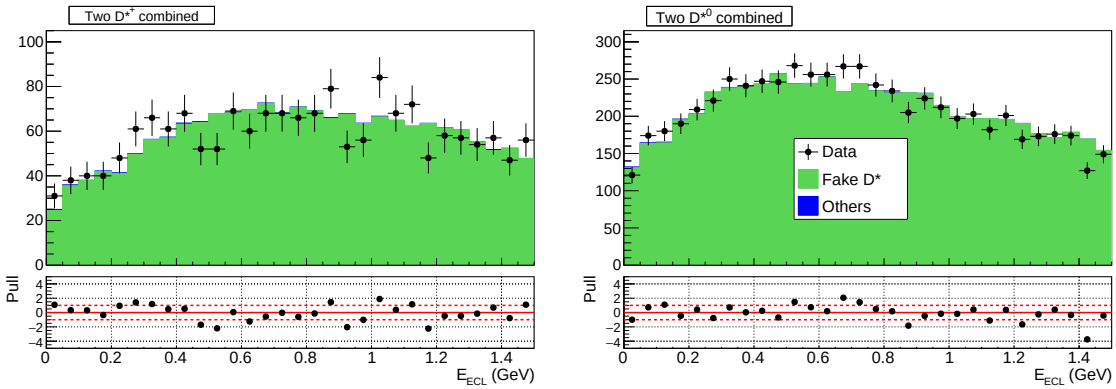


Figure 3.25: Comparison of the $E_{\rm ECL}$ distribution of the D^{*+} sample (left) and the D^{*0} sample (right) between the data and the MC simulation in the ΔM sideband region. The solid histogram shows the MC distribution with two components: fake D^* (green) and all the other event with a true D^* (blue).

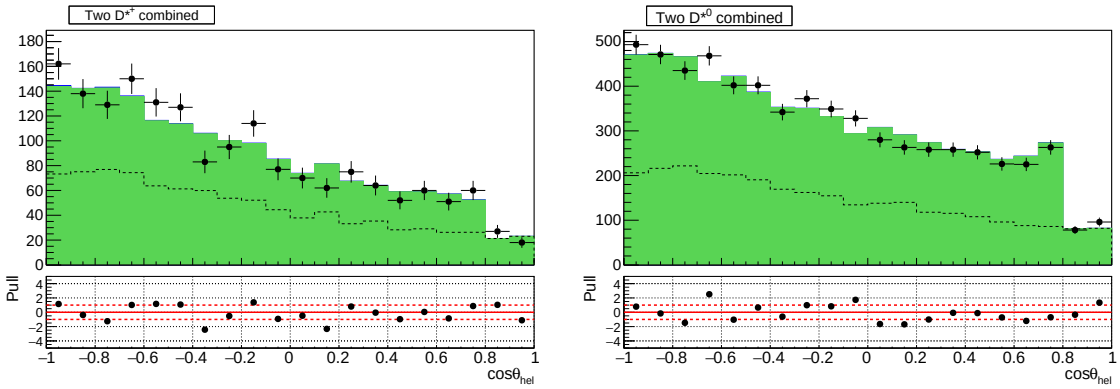


Figure 3.26: Comparison of the $\cos\theta_{\rm hel}$ distribution of the D^{*+} sample (left) and the D^{*0} sample (right) between the data and the MC simulation in the ΔM sideband region. The meaning of each distribution is the same as Fig. 3.25. The component above (below) the dashed line is events in the $\tau^- \to \pi^-\nu_\tau$ ($\tau^- \to \rho^-\nu_\tau$) sample.

3.6.2 $\cos\theta_{\rm hel}$ Distribution (Fake D^* Component)

Figure 3.26 shows the comparison of the $\cos\theta_{\rm hel}$ distributions between the data and the MC simulation in the ΔM sideband region. The distribution in the MC simulation well

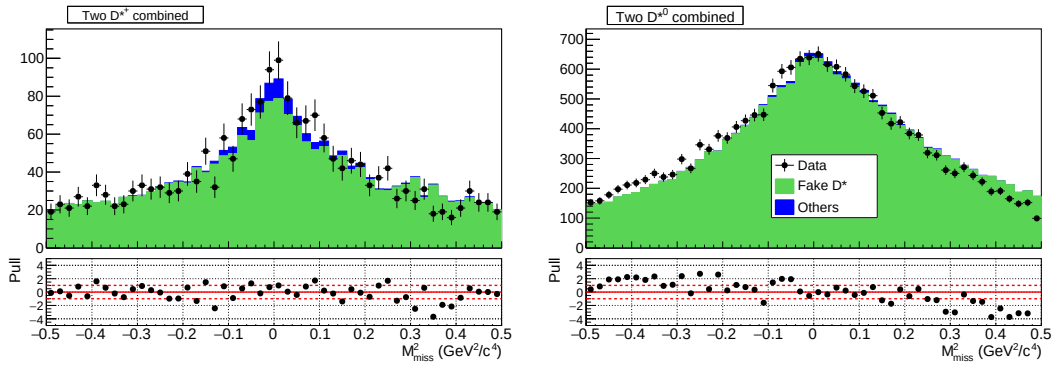


Figure 3.27: Comparison of the $M^2_{\rm miss}$ distribution of the D^{*+} sample (left) and the D^{*0} sample (right) between the data and the MC simulation in the ΔM sideband region. The meaning of each distribution is the same as Fig. 3.25.

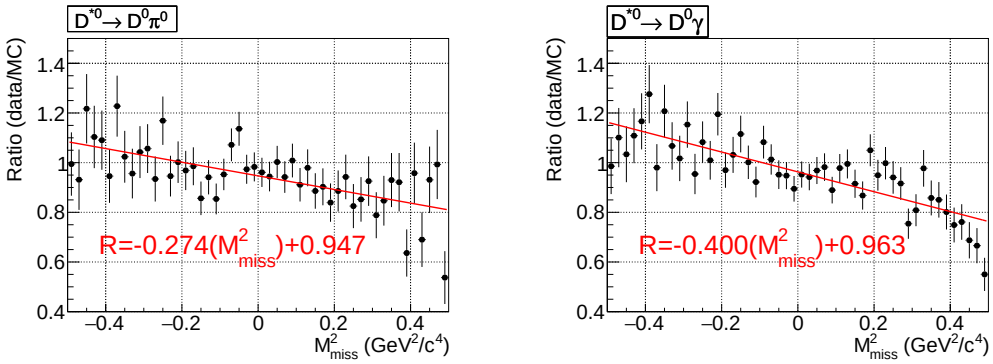


Figure 3.28: Ratio of the data and the MC sample for $D^{*0} \to D^0\gamma$ (left) and $D^{*0} \to D^0\pi^0$ (right). The formula R denotes the shape correction function.

reproduces that in the data.

3.6.3 $M^2_{\rm miss}$ Distribution (Fake D^* Component)

The $M^2_{\rm miss}$ distribution of the fake D^* component is validated using samples in the ΔM sideband region as shown in Fig. 3.27. While the D^{*+} sample shows good agreement between the data and the MC sample, the D^{*0} sample shows a discrepancy. We correct the distribution shape using the D^*-mode-dependent linear function, which are obtained by the fits to the ratio of the distributions as shown in Fig. 3.28,

3.6.4 Summary of the MC Calibration

Up to the previous section, we have completed the calibration and the validation of the MC simulation. Table 3.18 shows the expected yield for each component in the signal sample.

3.7 Measurement of $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$ Branching Fraction

As a cross check, we measure the branching fraction of the decay $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$. We perform a fit to the normalization sample using the obtained $M^2_{\rm miss}$ distributions as histogram PDFs. The fit result is shown in Fig. 3.29. The p-value calculated from the agreement between the data and the fitted PDFs is 15%. The normalization yields are measured to be 4711 ± 81 events for the B^- sample and 2502 ± 52 events for the $\bar{B}^0$ sample, where the errors are statistical. From the yield, the branching fraction is calculated as

$$BF(B^- \to D^* \ell^- \bar{\nu}_\ell) \;=\; \frac{N_{\rm norm}}{\epsilon_{\rm norm} N_{B\bar{B}}}, \tag{3.17}$$

where $N_{B\bar{B}}$ is the total number of $B\bar{B}$ pairs, and $N_{\rm norm}$ and $\epsilon_{\rm norm}$ denote the yield and the efficiency of the normalization mode, respectively. Using the MC simulation, the efficiency is estimated to be $(5.54 \pm 0.02) \times 10^{-5}$ for the B^- sample and $(3.15 \pm 0.02) \times 10^{-5}$ for the $\bar{B}^0$ sample, where the errors arise from the MC statistics.

In addition to the statistical uncertainties, we consider the following major systematic uncertainties: the number of $B\bar{B}$ pairs (2.0%, 2.0%), the tag efficiency (4.2%, 4.6%), the branching fractions of the D meson (1.9%, 1.7%), and the particle reconstruction and identification efficiencies (3.3%, 3.2%). The first and the second values in the parentheses are the uncertainties for the B^- mode and the $\bar{B}^0$ mode, respectively. The obtained

Table 3.18: Expected number of events for each category from the calibrated MC.

		$\tau^- \to \pi^- \nu_\tau$		$\tau^- \to \rho^- \nu_\tau$	
B^-	Signal	69.3	10.7%	46.5	9.7%
	$\rho \leftrightarrow \pi$ cross feed	37.4	5.8%	2.8	0.6%
	Other τ cross feed	40.7	6.3%	12.5	2.6%
	$\bar{B} \to D^* \ell^- \bar{\nu}_\ell$	17.9	2.8%	2.0	0.4%
	$\bar{B} \to D^{**} \ell^- \bar{\nu}_\ell$ and had. B	73.8	11.4%	67.0	14.0%
	Continuum	0.8	0.1%	0.4	0.1%
	Fake D^*	409.6	63.1%	347.3	72.6%
	Total	649.5		478.4	
$\bar{B}^0$	Signal	25.5	17.9%	19.2	18.5%
	$\rho \leftrightarrow \pi$ cross feed	13.0	9.1%	1.5	1.4%
	Other τ cross feed	14.6	10.3%	5.1	4.9%
	$\bar{B} \to D^* \ell^- \bar{\nu}_\ell$	10.9	7.7%	0.7	0.6%
	$\bar{B} \to D^{**} \ell^- \bar{\nu}_\ell$ and had. B	54.6	38.4%	52.0	49.9%
	Continuum	0.5	0.3%	0.2	0.2%
	Fake D^*	23.1	16.2%	25.5	24.5%
	Total	142.1		104.1	

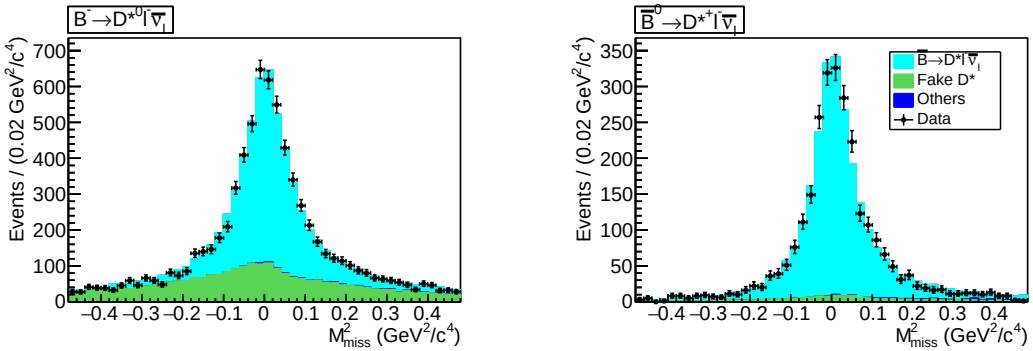


Figure 3.29: Fit result to the normalization sample.

branching fractions are,[1]

$$BF(B^- \to D^{*0}\ell^-\bar{\nu}_\ell) = (10.72 \pm 0.19 \pm 0.64)\%, \quad (3.18)$$

$$BF(\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell) = (10.60 \pm 0.22 \pm 0.66)\%. \quad (3.19)$$

where the first and the second errors show the statistical and the systematic uncertainties, respectively. The obtained branching fractions are consistent with the world averages $BF(B^- \to D^{*0}\ell^-\bar{\nu}_\ell) = (11.38 \pm 0.38)\%$ and $BF(\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell) = (9.86 \pm 0.22)\%$ [10].

[1]The values are the sum of $\bar{B} \to D^* e^- \bar{\nu}_e$ and $\bar{B} \to D^* \mu^- \bar{\nu}_\mu$.

Chapter 4

Signal Extraction Method

Using the corrected MC distributions as histogram PDFs, we perform a fit to the data to measure $R(D^*)$ and $P_\tau(D^*)$. In the fit, possible bias arising from the detector efficiency needs to be considered. Non-uniformity in the efficiency over $\cos\theta_{\rm hel}$ causes a bias on the $P_\tau(D^*)$ measurement. Modification of the kinematic distributions due to NP effects induces changes of the reconstruction efficiency. In this chapter, we describe formulation of $R(D^*)$ and $P_\tau(D^*)$ for the fit and discuss the efficiency estimation. We then test the fitting method using pseudo data and evaluate validity of the fit strategy.

4.1 Fit Strategy

4.1.1 $R(D^*)$ and $P_\tau(D^*)$ Parameterization

To extract $R(D^*)$ and $P_\tau(D^*)$ from data, we relate the signal yields to these observables. Using $BF(\bar{B} \to D^* \tau^- \bar{\nu}_\tau) = N_{\rm sig}/(\epsilon_{\rm sig} N_{B\bar{B}} BF_\tau)$ and $BF(\bar{B} \to D^* \ell^- \bar{\nu}_\ell) = (N_{\rm norm})/(\epsilon_{\rm norm} N_{B\bar{B}})$, $R(D^*)$ is represented by

$$R(D^*) \;=\; \frac{1}{BF_\tau^i} \frac{2\epsilon_{\rm norm}}{\epsilon_{\rm sig}^i} \frac{N_{\rm sig}^i}{N_{\rm norm}} \tag{4.1}$$

where BF_τ, $\epsilon_{\rm sig(norm)}$ and $N_{\rm sig(norm)}$ are the branching fraction of τ, the efficiency and the observed yield of the signal (normalization) mode, respectively. For the variables of the signal mode, the index i denotes the τ mode: $i = 1$ for $\tau^- \to \pi^- \nu_\tau$ and $i = 2$ for $\tau^- \to \rho^- \nu_\tau$. In the discussion below, we omit the index i for simplicity.

For the $P_\tau(D^*)$ measurement, we use the parameterization in Eq. 3.6, where $N_{\rm F} + N_{\rm B} = N_{\rm sig}$. However, this formula does not take the detector efficiency into account. If the efficiency has dependence on $\cos\theta_{\rm hel}$, bias is induced in the measured $P_\tau(D^*)$ value. We evaluate the detector effects using the MC simulation, as discussed later in Sec. 4.2.

In the $\tau^- \to \pi^- \nu_\tau$ sample, we require $\cos\theta_{\rm hel}$ to be less than 0.8. This requirement causes the efficiency variation as a function of $P_\tau(D^*)$. For example, as illustrated in Fig. 4.1, a larger value of $P_\tau(D^*)$ increases the number of discarded signal events $N_{\rm dis}$ in the region $0.8 < \cos\theta_{\rm hel} < 1$. The signal efficiency with $P_\tau(D^*) = +0.5$ is by about 10% smaller than that with $P_\tau(D^*) = -0.5$ (SM). To take the efficiency variation into account, we employ the following correction. Integrating Eq. 3.2 from 0.8 to 1 and substituting α

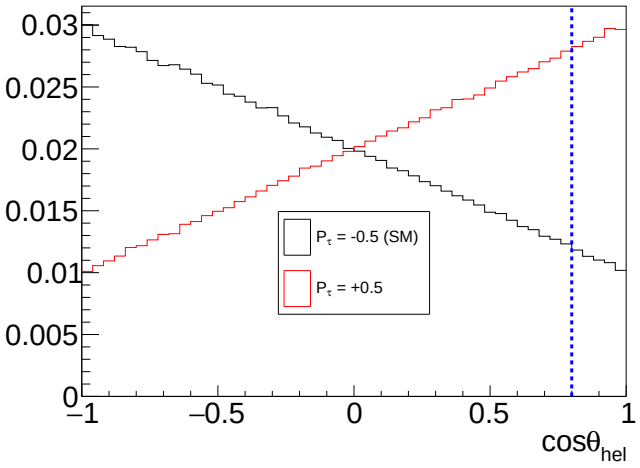


Figure 4.1: Comparison of the $\cos\theta_{\rm hel}$ distributions in the $\tau^- \to \pi^-\nu_\tau$ mode with $P_\tau(D^*) = -0.5$ (black) and $+0.5$ (red). The dashed blue line shows the threshold value for $\cos\theta_{\rm hel}$; the signal events in the right side of the blue line are discarded.

by 1, a linear equation with respect to $N_{\rm dis}$ is obtained as

$$N_{\rm dis} = \frac{N'_{\rm F} + N_{\rm B} + N_{\rm dis}}{2}\left[0.2 + 0.18P_\tau(D^*)\right], \tag{4.2}$$

where $N'_{\rm F}$ is the observed signal yield in $0 < \cos\theta_{\rm hel} < 0.8$. Solving the equation,

$$N_{\rm dis} = \frac{N'_{\rm F} + N_{\rm B}}{2}\frac{0.2 - 0.18P_\tau(D^*)}{0.9 - 0.09P_\tau(D^*)} \tag{4.3}$$

is obtained. The formula of $R(D^*)$ is then represented by

$$R(D^*) = \frac{1}{BF_\tau}\frac{\epsilon_{\rm norm}}{\epsilon_{\rm sig}}\frac{N_{\rm sig} + N_{\rm dis}(N_{\rm sig}, P_\tau(D^*))}{N_{\rm norm}} \quad (N_{\rm sig} = N'_{\rm F} + N_{\rm B}), \tag{4.4}$$

where $\epsilon_{\rm sig}$ is the signal reconstruction efficiency in the entire $\cos\theta_{\rm hel}$ region from -1 to $+1$. It should be noted that $R(D^*)$ and $P_\tau(D^*)$ are not independent anymore. This method is valid only in the case that the efficiency does not strongly depend on $\cos\theta_{\rm hel}$. The uniformity of the efficiency is investigated later in Sec. 4.2.1.

For the $P_\tau(D^*)$ measurement, we introduce the re-defined forward-backward asymmetry

$$A'_{\rm FB} = \frac{N'_{\rm F} - N_{\rm B}}{N'_{\rm F} + N_{\rm B}}. \tag{4.5}$$

Using the integration of Eq. 3.2 from 0 to 0.8 and -1 to 0, N'_F and N_B are obtained as

$$N'_{\rm F} = \frac{N'_{\rm F} + N_{\rm B} + N_{\rm dis}}{2}\left[0.8 + 0.32P_\tau(D^*)\right], \tag{4.6}$$

$$N_{\rm B} = \frac{N'_{\rm F} + N_{\rm B} + N_{\rm dis}}{2}\left[1 + 0.5P_\tau(D^*)\right]. \tag{4.7}$$

$$\tag{4.8}$$

By substituting Eqs. 4.6 and 4.7 for Eq. 3.5, the following relation is obtained:

$$\frac{N'_{\rm F} - N_{\rm B}}{N'_{\rm F} + N_{\rm B}} = \frac{[0.8 + 0.32P_\tau(D^*)] - [1 - 0.5P_\tau(D^*)]}{[0.8 + 0.32P_\tau(D^*)] + [1 - 0.5P_\tau(D^*)]}, \tag{4.9}$$

$$= -\frac{0.2 - 0.82P_\tau(D^*)}{1.8 - 0.18P_\tau(D^*)}. \tag{4.10}$$

Transforming the equation, $P_\tau(D^*)$ is obtained as

$$P_\tau(D^*) = \frac{2N'_{\rm F} - 1.6N_{\rm B}}{N'_{\rm F} + 0.64N_{\rm B}}. \tag{4.11}$$

For the $\tau^- \to \rho^-\nu_\tau$ mode, we do not employ the requirement on $\cos\theta_{\rm hel}$. The formulae 4.1 and 3.6 are used for the measurement of $R(D^*)$ and $P_\tau(D^*)$, respectively.

4.1.2 Fitting Method

Based on the formulae discussed in the previous section, the two-step fitting procedure is employed.

- Fit to the normalization sample and obtain $N_{\rm norm}$. (The result has been already shown in Sec. 3.7.)
- Simultaneous fit to the eight signal samples $(B^-, \bar{B}^0) \otimes (\pi^-\nu_\tau, \rho^-\nu_\tau)$ $\otimes$ (forward, backward).

An extended binned maximum likelihood fit is performed based on the likelihood

$$L = \prod_{i=1}^{8}\prod_{j=1}^{30}\left\{\frac{[F_{ij}]^{n_j}}{n_j!}e^{-[F_{ij}]^{n_j}}\right\}, \tag{4.12}$$

where F_{ij} denotes the total PDF of the ith sample in the jth $E_{\rm ECL}$ bin, and n_j is the number of events in the jth $E_{\rm ECL}$ bin. The PDF F_{ij} is the sum of the seven PDF components f_{ij}^k and represented by

$$F_{ij} = \sum_{k=1}^{7}[Y_i^k f_{ij}^k], \tag{4.13}$$

where Y_i^k is the yield of the kth PDF component in the ith signal sample.

The method of the yield determination is summarized in Table 4.1. The yields of the $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ components are associated with $R(D^*)$ and $P_\tau(D^*)$. Assuming the isospin symmetry, we use the common $R(D^*)$ and $P_\tau(D^*)$ between the B^- and the $\bar{B}^0$ samples. For the $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ and hadronic B component, we treat the yields as the fit parameters independent of the eight signal samples. The yields for the other components are fixed. We eventually have 10 fit parameters in total.

Table 4.1: Summary of the yield determination in the fit. The row "Hadronic B" implies " Hadronic B and $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$".

PDF component	Yield determination method	# of free param.
Signal	Free, associated with $R(D^*)$ and $P_\tau(D^*)$	2
$\rho \leftrightarrow \pi$ cross feed	Free, associated with $R(D^*)$ and $P_\tau(D^*)$	
Other τ cross feed	Free, associated with $R(D^*)$	
$\bar{B} \to D^*\ell^-\bar{\nu}_\ell$	Fixed	0
Hadronic B	Free	8
Continuum	Fixed	0
Fake D^*	Fixed	0

4.2 Signal Reconstruction Efficiency

Using the signal MC sample with the SM value of $P_\tau(D^*)$, the reconstruction efficiency is estimated as shown in Table 4.2. For the B^- sample, the efficiency ratio $\epsilon_{\rm norm}/\epsilon_{\rm sig}$ of the $\tau^- \to \pi^-\nu_\tau$ mode in the signal component is found to be 0.97 ± 0.02, which is consistent with 1. It is reasonable because both signal and normalization modes have the same final state containing one $B_{\rm tag}$, one D^* and one charged particle. For the $\bar{B}^0$ sample, the efficiency ratio is 1.21 ± 0.03 and significantly larger than 1. It arises from the negative correlation between q^2 and the D^* momentum. Compared to the normalization mode, the signal mode has higher q^2 and hence lower D^* momentum, decreasing the momentum of the soft π^+ from the $D^{*+} \to D^0\pi^+$ decay. As shown in Fig. 4.2, the soft π^+ efficiency is almost zero in the momentum region less than 60 MeV/c. The signal mode eventually has the lower efficiency than the normalization mode.

4.2.1 Acceptance

The efficiency variation due to the NP effects is precisely investigated using the MC simulation with sufficiently large statistics. Since the decays of $B_{\rm tag}$ and $B_{\rm sig}$ are independent, we generate the MC sample using only two $B_{\rm tag}$ decay modes which have the best

Table 4.2: Reconstruction efficiency estimated from the signal MC sample with the SM value of $P_\tau(D^*)$ (with a unit of $\times 10^{-5}$). The uncertainties arise from the MC statistical errors.

	B^- sample		$\bar{B}^0$ sample	
	$\tau^- \to \pi^-\nu_\tau$	$\tau^- \to \rho^-\nu_\tau$	$\tau^- \to \pi^-\nu_\tau$	$\tau^- \to \rho^-\nu_\tau$
Signal	5.73 ± 0.10	1.62 ± 0.03	2.60 ± 0.07	0.82 ± 0.02
$\rho \leftrightarrow \pi$ cross feed	1.28 ± 0.03	0.22 ± 0.02	0.55 ± 0.02	0.14 ± 0.02
Other τ cross feed	0.55 ± 0.01	0.171 ± 0.007	0.244 ± 0.008	0.085 ± 0.005
$\bar{B} \to D^*\ell^-\bar{\nu}_\ell$	5.54 ± 0.02		3.15 ± 0.02	

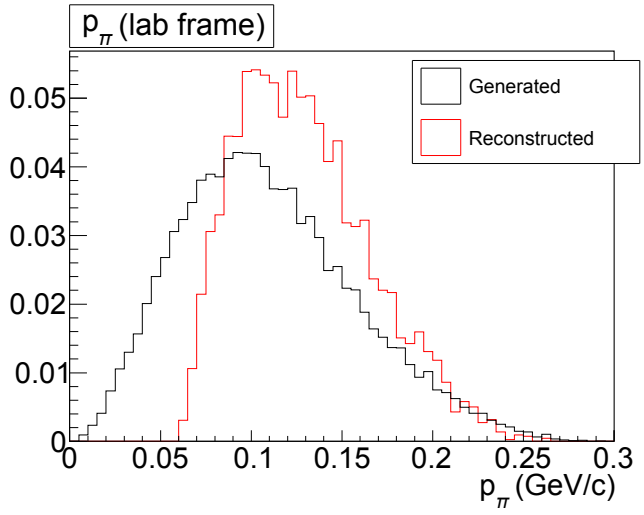


Figure 4.2: Comparison of the generated and the reconstructed momentum distribution of the soft π^+ from the $D^{*+} \to D^0\pi^+$ decay. Both histograms are normalized to be a unit area.

reconstruction efficiency (charge-conjugate modes are implied):

- $B^+_{\rm tag} \to \bar{D}^0\pi^+$, $\bar{D}^0 \to K^+\pi^-$,
- $B^0_{\rm tag} \to D^-\pi^+$, $D^- \to K^+\pi^-\pi^-$,

The number of reconstructed signal events is roughly 120 times more than the expected signal yield in the Belle experiment.

Figure 4.3 shows the reconstruction efficiency as a function of $\cos\theta_{\rm hel}$. It is almost constant, however, tends to be smaller in the backward region for the $\tau^- \to \rho^-\nu_\tau$ mode. The efficiency dependence originates from the correlation between $\cos\theta_{\rm hel}$ and the momentum of π^0 from the ρ decay, shown in Fig. 4.4. In the small $\cos\theta_{\rm hel}$ region, the π^0 momentum tends to be smaller, that induces the lower efficiency. In contrast to the $\tau^- \to \rho^-\nu_\tau$ mode, the efficiency is uniform in the $\tau^- \to \pi^-\nu_\tau$ mode. Our $P_\tau(D^*)$ extraction method in Eq. 4.4 is therefore applicable.

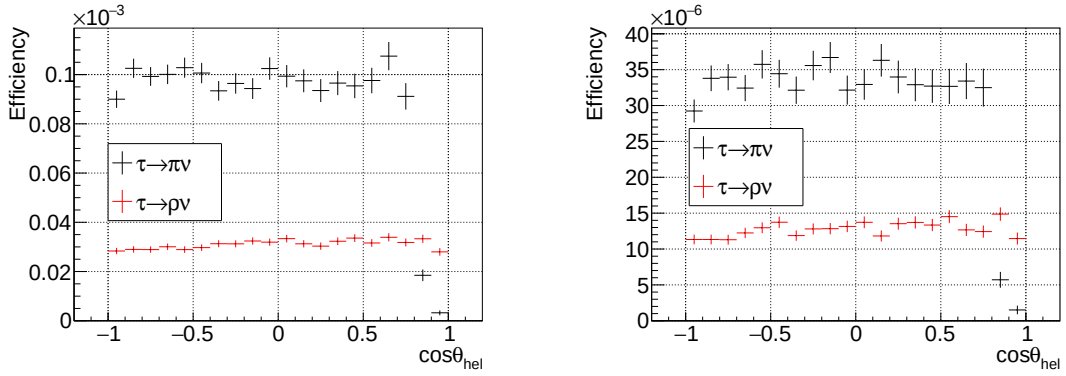


Figure 4.3: Reconstruction efficiency as a function of $\cos\theta_{\rm hel}$. The left and the right panels show the efficiencies for the B^- and the $\bar{B}^0$ samples, respectively.

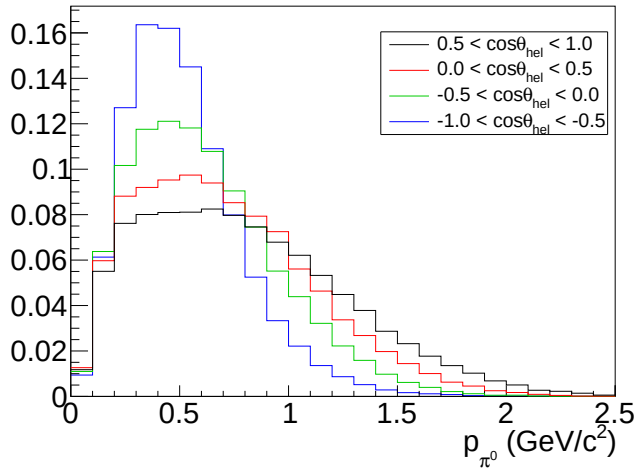


Figure 4.4: Momentum distribution of π^0 from the ρ meson decay in different $\cos\theta_{\rm hel}$ regions.

While the efficiency dependence on the other kinematic variables $\cos\theta_\tau$, q^2 and $\cos\theta_V$ does not cause a large modification in the efficiency uniformity for $\cos\theta_{\rm hel}$, it affects the total reconstruction efficiency. Figures 4.5 and 4.6 show the reconstruction efficiency for each variable. As obviously seen in the right two panels of Fig. 4.6, the q^2 and $\cos\theta_V$ dependence of the efficiency in the $D^{*+} \to D^0\pi^+$ mode is significant. This is induced by the strong momentum dependence of the reconstruction efficiency of the soft $\pi^\pm$, as discussed in the previous section. The impact of the q^2 and $\cos\theta_V$ dependence is investigated in the next section.

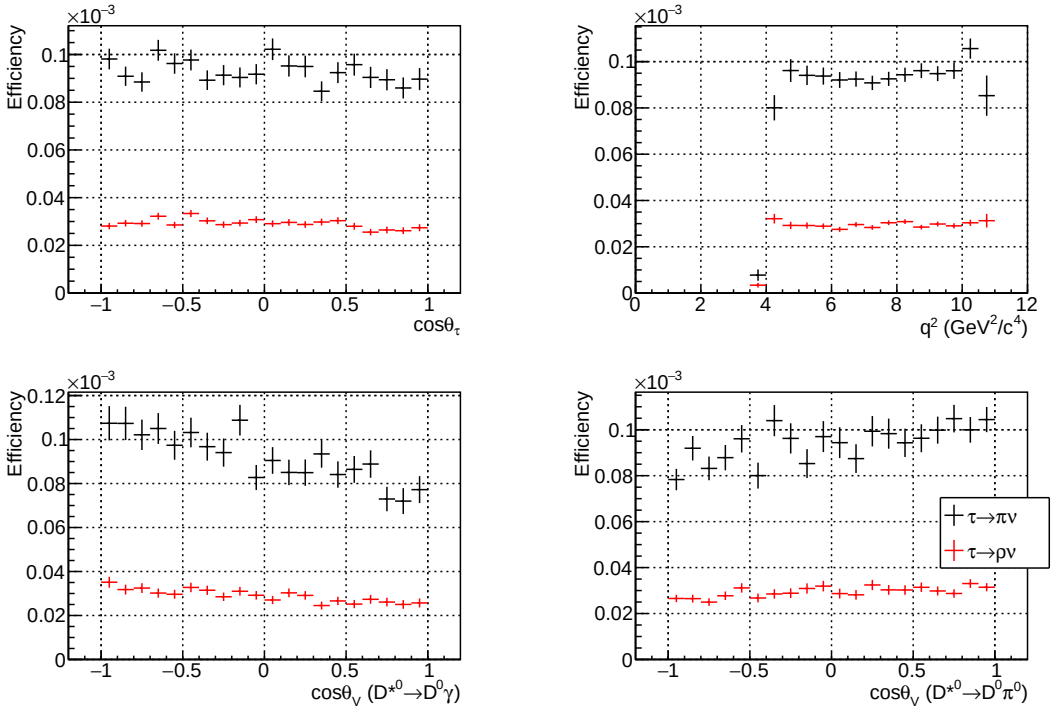


Figure 4.5: Reconstruction efficiency for $\cos\theta_\tau$, q^2 and $\cos\theta_V$ ($D^{*0} \to D^0\gamma$ and $D^{*0} \to D^0\pi^0$) in the B^- sample.

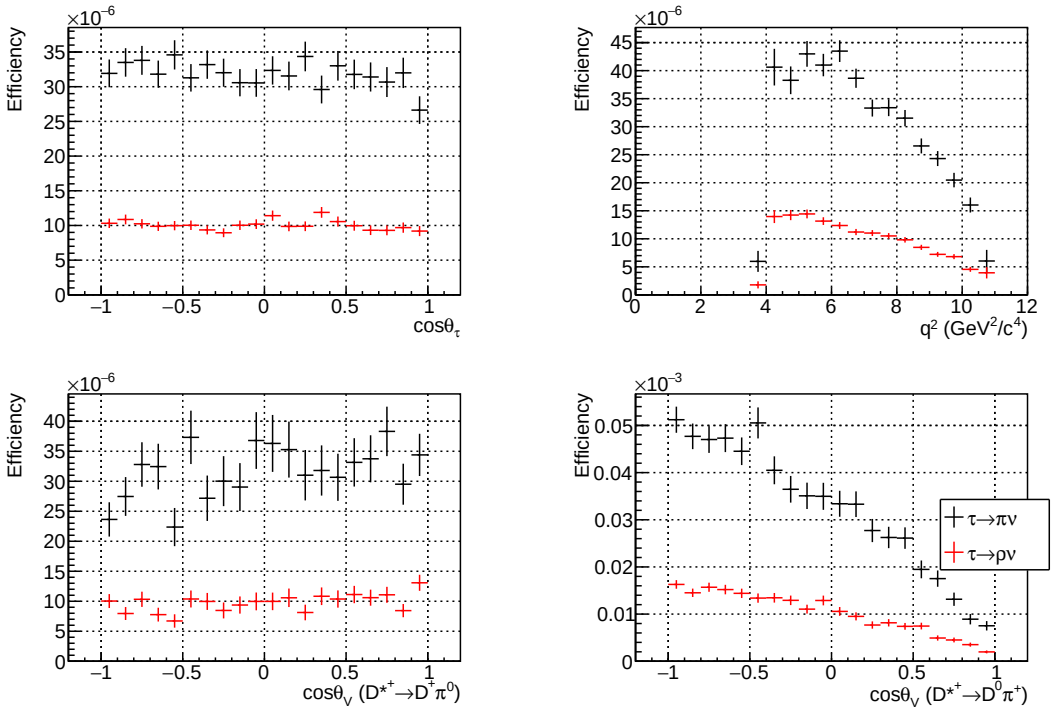


Figure 4.6: Same figure as Fig. 4.5 for the $\bar{B}^0$ sample.

4.2.2 Efficiency Variation in Different $P_\tau(D^*)$ Cases

To investigate the efficiency in different values of $P_\tau(D^*)$, we produce the signal MC samples with $P_\tau(D^*)$ from -0.9 to $+0.9$. The value of $P_\tau(D^*)$ is determined by the amplitude

$$A = C_+ h_+ + C_- h_-, \tag{4.14}$$

where $h_\pm$ denotes the amplitude with the τ helicity $\pm 1/2$. The values of the real coefficients $C_\pm$ ($0 < C_\pm < 1$) are adjusted to obtain a specific value of $P_\tau(D^*)$. Due to the modification of the helicity amplitudes, the distributions of q^2, $\cos\theta_V$ and $\cos\theta_\tau$ also vary, as shown in Fig. 4.7.

Figure 4.8 (left) shows the relative efficiency for the signal mode. The relative efficiency is defined as the efficiency at $P_\tau(D^*)$ with respect to that at $P_\tau(D^*)_{\rm SM} = -0.497$. As shown by the closed circles, the efficiency is almost constant for the $\tau^- \to \pi^-\nu_\tau$ mode and is varied by 5% at the maximum for the $\tau^- \to \rho^-\nu_\tau$ mode. Although it is found in the previous section that the efficiency for the $D^{*+} \to D^-\pi^+$ mode strongly depends on q^2 and $\cos\theta_V$, Fig. 4.8 does not indicate a large efficiency variation. The relative efficiency variation for the D^* mode is shown in Appendix G. For the $\tau^- \to \pi^-\nu_\tau$ mode, the relative efficiency without the correction for the $\cos\theta_{\rm hel}$ requirement is also shown by the open circles. Without the correction, the efficiency largely varies up to 15%.

Figure 4.8 (right) shows the efficiency for the $\rho \leftrightarrow \pi$ cross-feed component. Especially, misreconstructed events of $\tau^- \to \pi^-\nu_\tau$ as $\tau^- \to \rho^-\nu_\tau$ shows a significant efficiency variation up to 20%. However, according to Table 3.18, the contribution of this component to the total number of $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ events is only about 5%. The impact of the efficiency

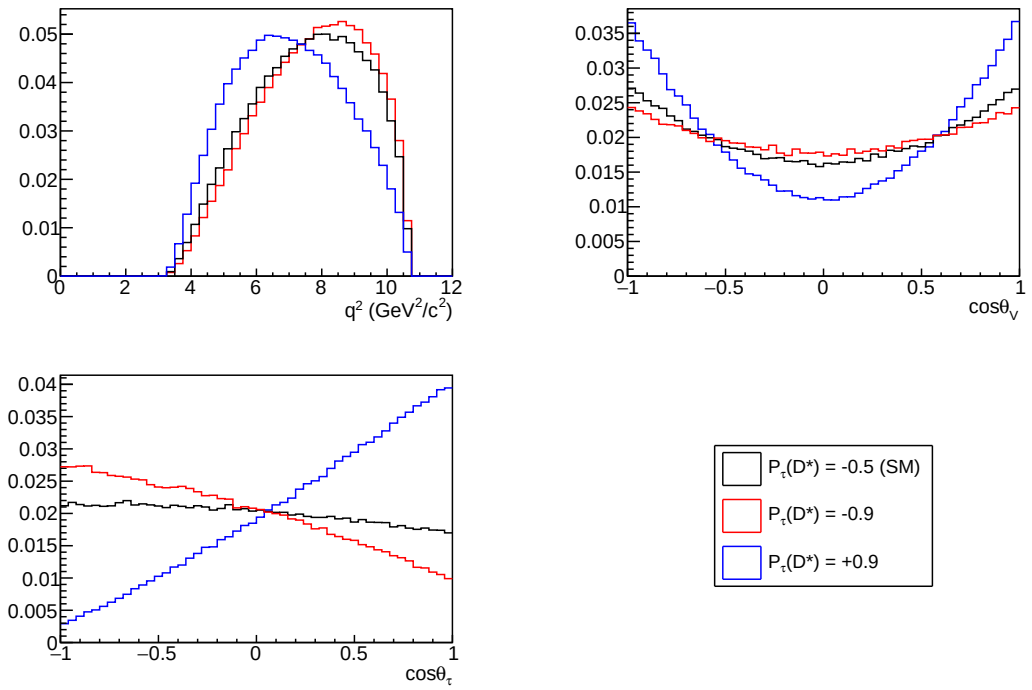


Figure 4.7: Distributions of q^2, $\cos\theta_V$ and $\cos\theta_\tau$ for the $D^{*+} \to D^0\pi^+$ mode in the MC simulation with different values of $P_\tau(D^*)$.

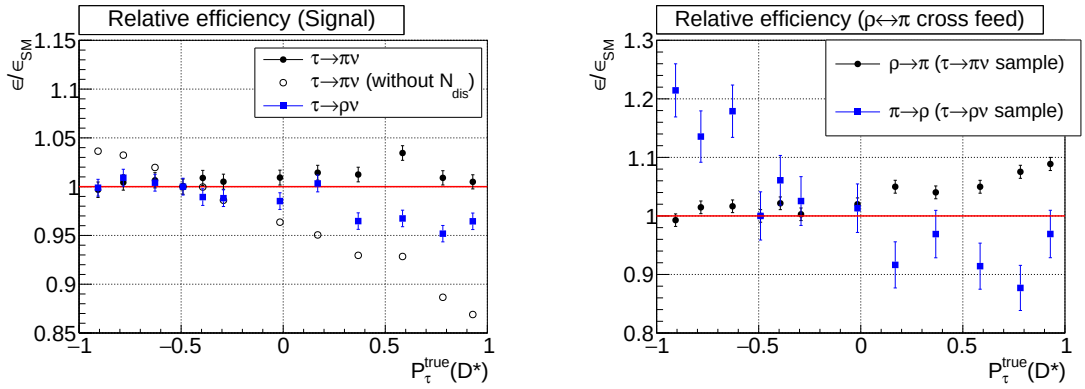


Figure 4.8: Reconstruction efficiency for the signal component relative to that of the SM case as a function of $P_\tau(D^*)$. (left) The efficiency for the signal component. The open circles for the $\tau^- \to \pi^-\nu_\tau$ mode shows the efficiency without correction for the $\cos\theta_{\rm hel}$ requirement. (right) The efficiency for the $\rho \leftrightarrow \pi$ cross-feed component.

variation is therefore expected to be less than 1% and negligibly small. Additional figures for the efficiency evaluation are shown in Appendix G.

4.2.3 Detector Bias Correction for $P_\tau(D^*)$

The detector bias on the $P_\tau(D^*)$ measurement is evaluated using the relation between the true polarization value $P_\tau^{\rm true}(D^*)$ and the extracted value $P_\tau^{\rm raw}(D^*)$ in the signal MC simulation, as shown in Fig. 4.9 (left). While there is no significant bias in the $\tau^- \to \pi^-\nu_\tau$ mode, the $\tau^- \to \rho^-\nu_\tau$ mode shows a visible bias up to -0.4 at $P_\tau^{\rm true}(D^*) = +1$. However, as shown later in Sec. 4.3.3, the expected statistical uncertainty in $P_\tau(D^*)$ from a measurement using only the $\tau^- \to \rho^-\nu_\tau$ mode is about 2. The shift by -0.4 is only 20% of the expected statistical uncertainty.

Figure 4.9 (right) shows the same relation for the $\rho \leftrightarrow \pi$ cross-feed component. It shows the significant bias because the $\rho \leftrightarrow \pi$ cross-feed events has the largely mis-estimated τ-daughter momentum due to a missing or an additional π^0. Linear functions

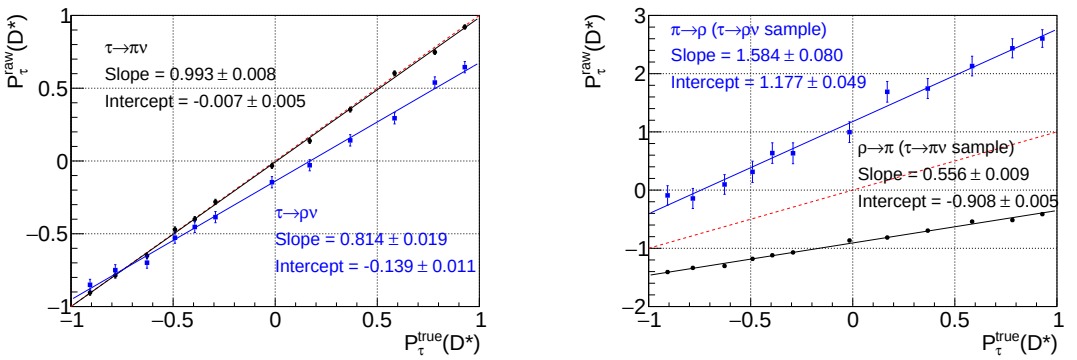


Figure 4.9: Relation between $P_\tau^{\rm true}(D^*)$ and $P_\tau^{\rm raw}(D^*)$ for the signal (left) and the $\rho \leftrightarrow \pi$ cross-feed components (right). The solid lines show the linear functions fitted to the data points. The dashed red line is the diagonal line as a reference.

are fitted to the relation between $P_\tau^{\rm true}(D^*)$ and $P_\tau^{\rm raw}(D^*)$ for each component. We call these functions "$P_\tau(D^*)$ correction functions". In the $P_\tau(D^*)$ measurement, we use them to remove the detector bias on $P_\tau(D^*)$.

4.2.4 Summary of the Efficiency Study

According to the studies so far, we find that

- the reconstruction efficiency variation is at most 5% (depending on the component of $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$) over the $P_\tau(D^*)$ range, and
- the detector bias on the $P_\tau(D^*)$ extraction is removed using the $P_\tau(D^*)$ correction function.

For the $R(D^*)$ measurement, we assume that the total reconstruction efficiency does not change from the estimation using the SM MC simulation. After fitting to the data, we assign a systematic uncertainty on $R(D^*)$ depending on the observed value of $P_\tau(D^*)$.

4.3 Fit Procedure Test

We construct a fitting program, called fitter, based on the strategy for the $R(D^*)$ and $P_\tau(D^*)$ measurement described in the previous section. In order to check validity of our fitter, we perform the fitter tests using pseudo-data sets. In this study, we use the SM predictions $R(D^*) = 0.252$ and $P_\tau(D^*) = -0.497$ as input.

4.3.1 Ensemble Test

Fit bias is evaluated using the pseudo-data sets composed of the MC sample. We prepare ten pseudo-data sets for the ensemble test. We use the same MC samples for the pseudo-data sets and the histogram PDFs. The statistical overlap between the pseudo data and the PDFs is 3 to 10%, which is sufficiently is small. One fit result to a pseudo-data set is shown in Figs. 4.10 and 4.11.

Figure 4.12 shows the measured values of $R(D^*)$ and $P_\tau(D^*)$ from the ten pseudo-data sets. All the values agree each other within the statistical error, and the averages of $R(D^*)$ and $P_\tau(D^*)$ are consistent with the input values. The result of the ensemble test indicates no bias in our fitter.

4.3.2 Linearity Test

To check the fitter validity at non-SM points, we perform a linearity test. The pseudo-data sets are generated by a toy MC simulation based on the histogram PDFs we constructed in Chapter 3. Five-thousand pseudo experiments are repeated at each point. Figure 4.13 shows the distribution of the pull

$$pull = \begin{cases} (O_{\rm meas} - O_{\rm input})/\sigma_{\rm neg}, & (O_{\rm meas} > O_{\rm input}), \\ (O_{\rm input} - O_{\rm meas})/\sigma_{\rm pos}, & (O_{\rm meas} < O_{\rm input}), \end{cases} \tag{4.15}$$

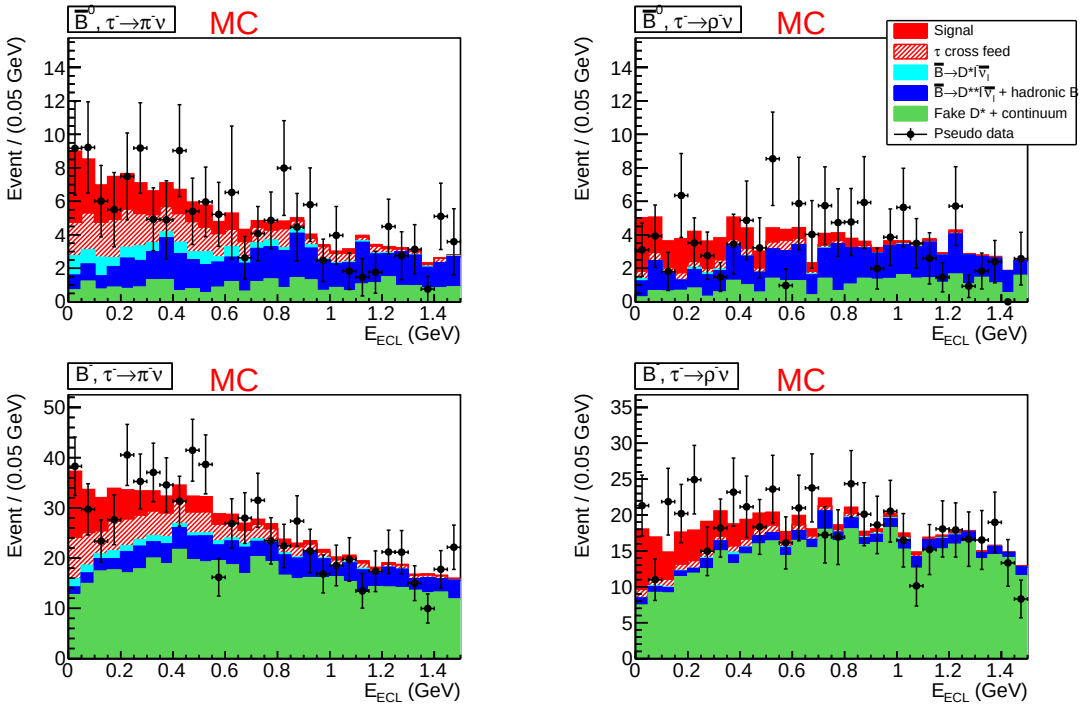


Figure 4.10: Fit result to the pseudo-data set, projected to the E_{ECL} axis. In the "τ cross-feed" component, the $\rho \leftrightarrow \pi$ cross-feed and the other τ cross-feed components are combined.

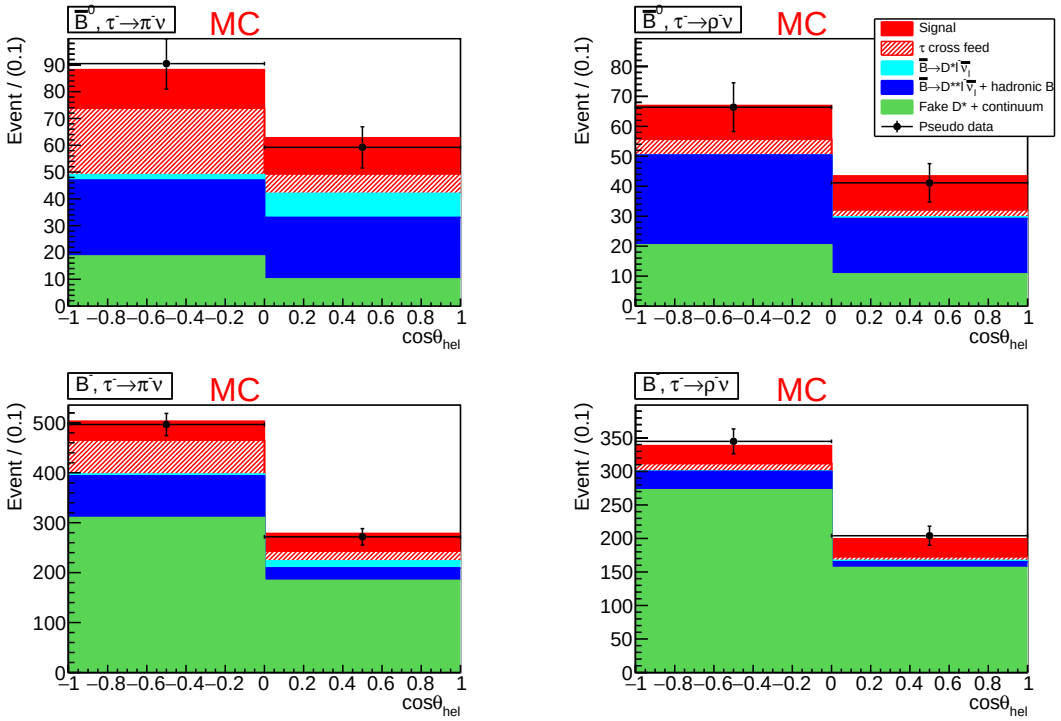


Figure 4.11: Fit result to the pseudo-data set, projected to the $\cos\theta_{\mathrm{hel}}$ axis. Used pseudo-data set and the meaning of each component are the same as Fig. 4.10.

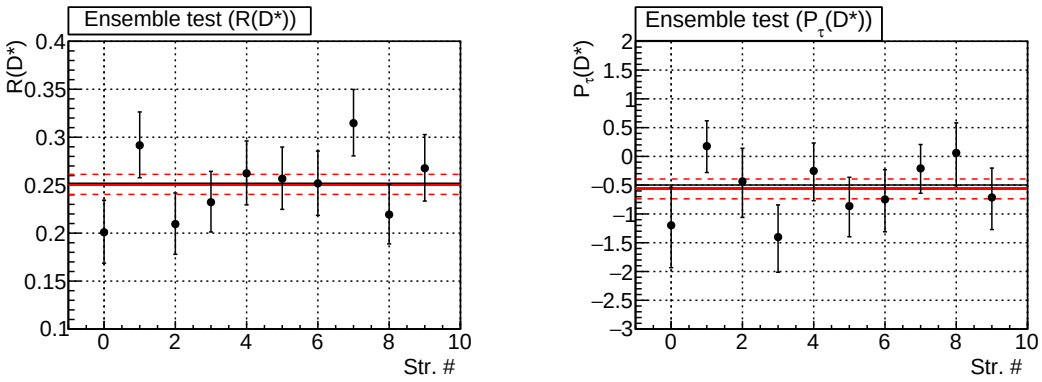


Figure 4.12: Result from the fits to the ten pseudo-data sets (black dots). The black line shows the input value, and the solid red lines is the average of the extracted values with the $\pm 1\sigma$ error.

at the SM point, where $O_{\mathrm{meas(input)}}$ denotes the measured (input) value of $R(D^*)$ or $P_\tau(D^*)$, and $\sigma_{\mathrm{pos(neg)}}$ is the positive (negative) error of the measurement. If there is no fit bias, the pull distribution becomes a Gaussian with a mean and a standard deviation of 0 and 1, respectively. As shown in Fig. 4.13, except for the mean from the $P_\tau(D^*)$ measurements, we find that the means and the standard deviations are by 3 to 4σ discrepant from the no-bias hypothesis. However, these discrepancies indicate only about 3% bias with respect to the expected statistical errors and are sufficiently small.

In the linearity test, we scan $R(D^*)$ from 0.126 to 0.403 and $P_\tau(D^*)$ from -0.9 to $+0.9$. Figure 4.14 shows the relation between the input value and the average of the measured $R(D^*)$ or $P_\tau(D^*)$. Bias on the $R(D^*)$ measurement is only at $O(0.1\%)$. For $P_\tau(D^*)$, we find a visible bias on $P_\tau(D^*)$ up to -0.05. However, the pull distribution for $P_\tau(D^*)$ in Fig. 4.13 (right) has the mean consistent with 0 at this input point. The $P_\tau(D^*)$ shift is presumably caused by the slightly asymmetric $P_\tau(D^*)$ distribution shown in Fig. 4.15. We rely on the result from the pull test, where the asymmetric error is taken into account.

According to the results of the fitter tests, we conclude that our fitter has no bias on the $R(D^*)$ and $P_\tau(D^*)$ measurement.

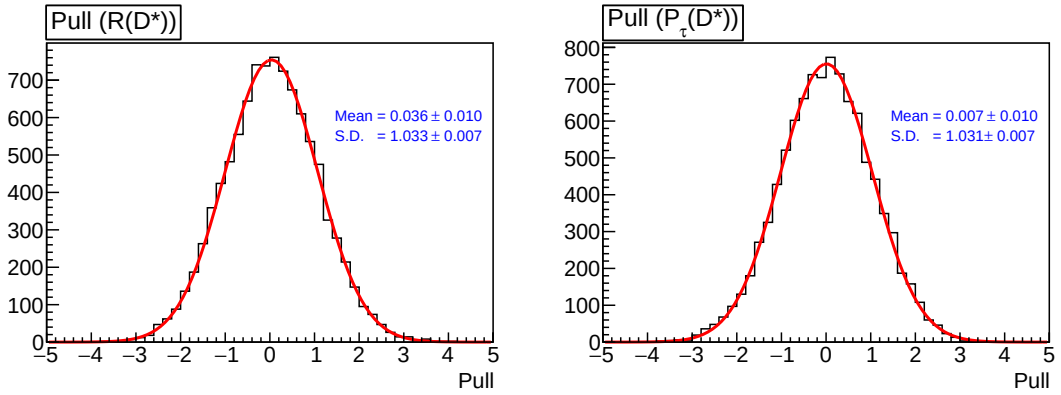


Figure 4.13: Pull distributions for $R(D^*)$ (left) and $P_\tau(D^*)$ at the SM point. The values Mean and S.D. are the mean and the standard deviation, respectively.

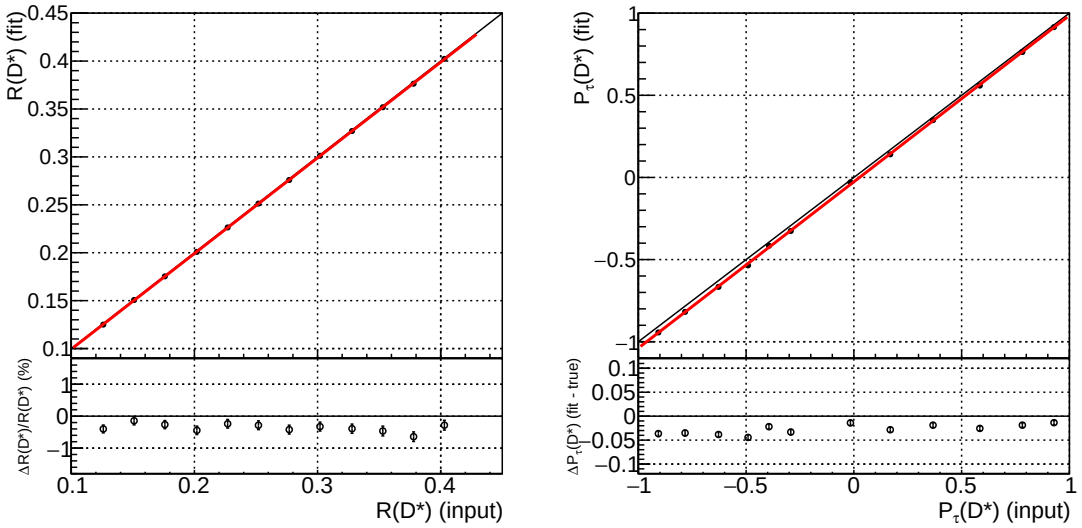


Figure 4.14: Comparison of the input values (horizontal axes) and the measured values (vertical axes) in the linearity test. The left and the right panels correspond to $R(D^*)$ and $P_\tau(D^*)$, respectively. The bottom panels show the difference between the true value and the extracted value from the fit.

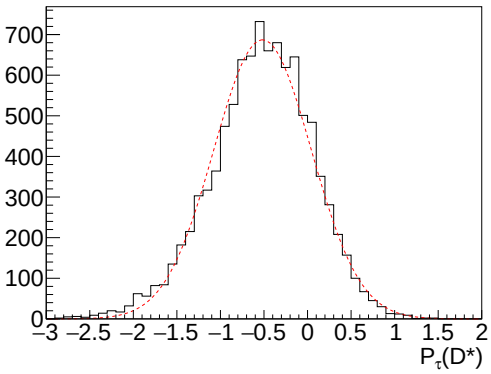


Figure 4.15: Distribution of $P_\tau(D^*)$ from the pseudo experiments at the SM point. As a reference, a fit of the Gaussian is indicated by the dashed red line.

4.3.3 Estimation of Statistical Errors

Using the toy MC study at the SM point, the expected statistical precision is estimated to be

$$\sigma_{[R(D^*)=0.252]} = \pm 0.034, \tag{4.16}$$

$$\sigma_{[P_\tau(D^*)=-0.50]} = \pm 0.56. \tag{4.17}$$

For the $R(D^*)$ measurement, we expect about the 13% statistical precision. This is comparable to the statistical uncertainties in the past measurements at Belle: 13% in the measurement with the hadronic tag [42] and 10% with the other analysis method, where $B_{\rm tag}$ is reconstructed from the semileptonic decays [43]. The statistical precision is also separately estimated for individual samples, as summarized in Table 4.3.

Table 4.3: Estimates of the statistical uncertainties for the individual samples. The top four (bottom two) rows are the expected statistical uncertainty for the individual samples of $(B^-, \bar{B}^0) \otimes (\pi^-\nu_\tau, \rho^-\nu_\tau)$ (B^- or $\bar{B}^0$ samples with both τ modes).

Sample	$\delta R(D^*)$	$\delta P_\tau(D^*)$
$(B^-, \pi^-\nu_\tau)$	0.05	0.8
$(\bar{B}^0, \pi^-\nu_\tau)$	0.06	1.0
$(B^-, \rho^-\nu_\tau)$	0.08	2.1
$(\bar{B}^0, \rho^-\nu_\tau)$	0.09	2.5
B^-	0.04	0.8
$\bar{B}^0$	0.05	0.9

Chapter 5

Systematic Uncertainty

In this chapter, we discuss the systematic uncertainties in $R(D^*)$ and $P_\tau(D^*)$. The basic strategy is to repeat 1000 fits to the data sample with varying one of the systematic uncertainty sources based on a Gaussian error. From the fit, 1000 $R(D^*)$ and $P_\tau(D^*)$ values are extracted. We regard standard deviations of the $R(D^*)$ and the $P_\tau(D^*)$ distributions

Table 5.1: List of the systematic uncertainties in $R(D^*)$ and $P_\tau(D^*)$. The values for $R(D^*)$ are indicated as the relative errors. The category "Common sources" is the common systematic uncertainty sources between the signal and the normalization modes. The column "Combined", "B^-" and "$\bar{B}^0$" show the uncertainties with combining the B^- and the $\bar{B}^0$ samples, only with the B^- and only with the $\bar{B}^0$ sample, respectively.

	Combined		B^-		$\bar{B}^0$	
Source	$R(D^*)$	$P_\tau(D^*)$	$R(D^*)$	$P_\tau(D^*)$	$R(D^*)$	$P_\tau(D^*)$
Hadronic B composition	$^{+7.6\%}_{-6.8\%}$	$^{+0.134}_{-0.103}$	$^{+6.1\%}_{-5.4\%}$	$^{+0.228}_{-0.196}$	$^{+12.7\%}_{-10.0\%}$	$^{+0.151}_{-0.127}$
MC statistics	$^{+4.0\%}_{-2.8\%}$	$^{+0.146}_{-0.108}$	$^{+4.2\%}_{-3.5\%}$	$^{+0.200}_{-0.162}$	$^{+7.5\%}_{-5.0\%}$	$^{+0.151}_{-0.127}$
Fake D^*	3.4%	0.018	5.2%	0.066	1.4%	0.017
$\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$	2.4%	0.048	3.1%	0.093	2.3%	0.014
$\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau$	1.1%	0.001	0.9%	0.013	1.5%	0.019
$\bar{B} \to D^*\ell^-\bar{\nu}_\ell$	2.3%	0.007	2.3%	0.010	2.2%	0.008
τ daughter and ℓ^- efficiency	1.9%	0.019	1.8%	0.015	1.9%	0.022
MC stat. for efficiency	1.0%	0.019	1.0%	0.024	1.7%	0.030
τBF	0.3%	0.002	0.3%	0.003	0.3%	0.001
$P_\tau(D^*)$ correction function	0.0%	0.010	0.1%	0.017	0.0%	0.005
Common sources						
Tagging efficiency correction	1.6%	0.018	1.6%	0.006	2.8%	0.000
D^* reconstruction	1.4%	0.006	1.3%	0.008	1.5%	0.005
D sub-decay BF	0.8%	0.007	0.7%	0.008	0.8%	0.010
Number of $B\bar{B}$	0.5%	0.006	0.5%	0.003	0.9%	0.003
Total systematic uncertainty	$^{+10.4\%}_{-9.4\%}$	$^{+0.21}_{-0.16}$	$^{+10.3\%}_{-9.7\%}$	$^{+0.33}_{-0.28}$	$^{+15.9\%}_{-12.6\%}$	$^{+0.26}_{-0.24}$

as the systematic uncertainties, unless otherwise specified.

The systematic uncertainties are determined based on the final fit results presented in Chapter 6. Table 5.1 is a summary of the estimated systematic uncertainties. The total systematic uncertainty is calculated by taking a quadratic sum of each component; the correlation between the τ-daughter efficiency and the D^* reconstruction efficiency is taken into account since they contain the common uncertainties arising from the π ID and the π^0 reconstruction efficiencies.

5.1 Composition of Hadronic B Decays

The uncertainty in the composition of the hadronic B decay modes causes bias on the $R(D^*)$ and $P_\tau(D^*)$ measurement through modification of the PDF shape. The systematic uncertainties are estimated based on the errors in the branching fractions or in the yield calibration factors in Table 3.15. For the B decay modes not having an experimentally-measured branching fraction nor covered by our calibration (uncalibrated mode), we vary the branching fraction continuously from 0% to 200% with respect to the value used in the MC simulation and take maximum differences of $R(D^*)$ and $P_\tau(D^*)$ from the nominal fit values as the systematic uncertainties.

Figure 5.1 shows the $R(D^*)$ and the $P_\tau(D^*)$ distributions from the two-D mode, the K_L^0 mode with experimental branching fractions, the $B^- \to D^*\pi^-\pi^-\pi^+$ mode in the multi-π/η mode and the other uncalibrated modes. In the calibrated B components in Fig. 5.1 (bottom-left), a strong asymmetry is observed due to the asymmetric uncertainties in the calibration factors. For such components, we take the standard deviation in the positive side and the negative side separately as the systematic uncertainty. The estimated uncertainties are summarized in Table 5.2.

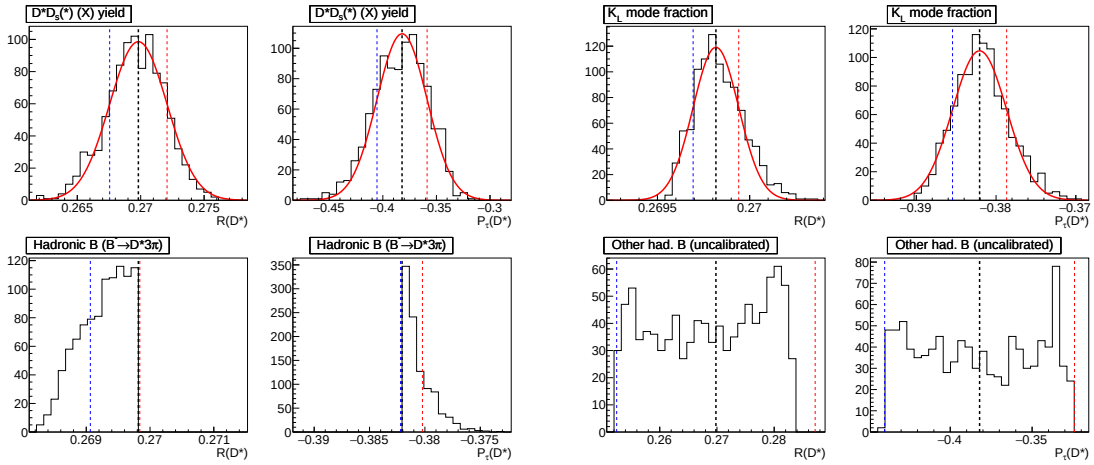


Figure 5.1: Distributions of $R(D^*)$ and $P_\tau(D^*)$ obtained by changing the hadronic B background composition. Two-D mode (top-left), K_L^0 mode with the experimental branching fractions (top-right), $B^- \to D^*\pi^-\pi^+\pi^-$ (bottom-left) and the uncalibrated mode (bottom-right) are shown. The dashed black line indicates the nominal fit value. The difference between the nominal value and the red (blue) line is regarded as the positive (negative) systematic uncertainty.

Table 5.2: Systematic uncertainties arising from the composition of the hadronic B decays.

Source	$R(D^*)$	$P_\tau(D^*)$
Two-D mode	0.8%	0.023
K^0_L mode	0.4%	0.066
Multi-π/η mode	$^{+4.0}_{-2.0}$%	$^{+0.098}_{-0.049}$
Other B decays	6.5%	0.058
Total	$^{+7.6}_{-6.8}$%	$^{+0.134}_{-0.103}$

5.2 MC Statistics

Due to the finite MC statistics, our PDF shapes have a significant statistical fluctuation as shown in Sec. 3.4.5. To estimate the fit bias arising from the statistical fluctuation on the PDFs, we re-shape the PDFs based on the statistical uncertainty in each PDF bin. This is performed by generating the same number of events randomly based on the original PDF. One of the re-shaped PDFs is illustrated in Fig. 5.2. This procedure is applied to each PDF component and every eight signal sample independently. Figure 5.3 shows the $R(D^*)$ and the $P_\tau(D^*)$ distributions obtained with changing all the PDF shapes at the same time. The asymmetric distribution is observed because the yield of the "signal" and the "$\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ and hadronic B components" are floated in the repetitive fit process. By taking a 68% interval of the distributions from the nominal fit values, we obtain the uncertainties of $^{+4.0}_{-2.8}$% for $R(D^*)$ and $^{+0.146}_{-0.108}$ for $P_\tau(D^*)$.

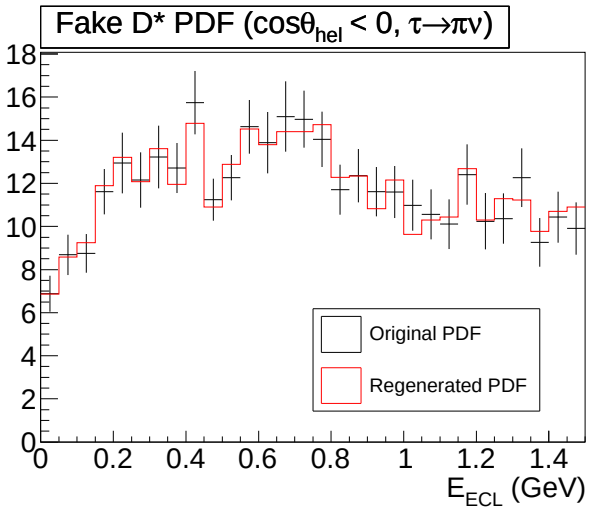


Figure 5.2: Example of the PDF regeneration for the fake D^* component of the $(B^-, \pi^-\nu_\tau, \text{backward})$ sample. The new PDF (red-solid) is generated from the original PDF shape (black dots with error bars).

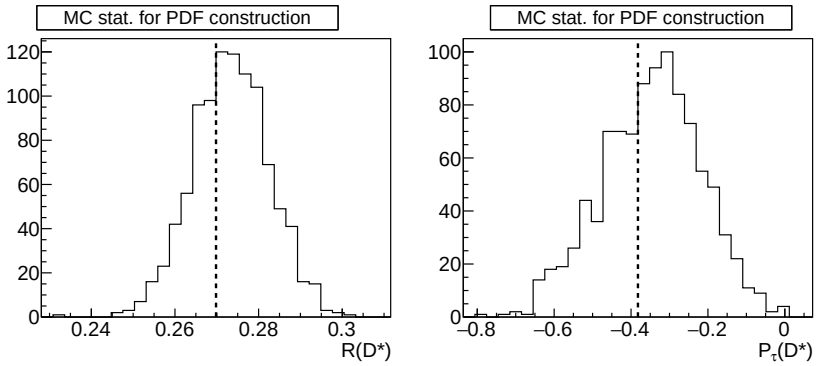


Figure 5.3: Distributions of $R(D^*)$ and $P_\tau(D^*)$ obtained by fluctuating the PDF shapes. The vertical dashed lines indicate the nominal fit values.

5.3 Fake D^* Component

Two uncertainty sources are taken into account in the fake D^* component. The first uncertainty arises from the yield calibration factors. The calibration factors have an uncertainty due to the sideband statistics. The second one arises from a peaking background in the fake D^* component originating from the decay $\bar{B} \to D\tau^-\bar{\nu}_\tau$. To estimate its effect, we vary the branching fraction based on the world averages $(0.77 \pm 0.25)\%$ for the B^- mode and $(1.03 \pm 0.22)\%$ for the $\bar{B}^0$ mode [10]. The $M^2_{\rm miss}$ shape correction for the normalization sample is also taken into account. Table 5.3 summarizes the estimated systematic uncertainties. In total, we estimate the uncertainties to be $\pm 3.4\%$ for $R(D^*)$ and ± 0.018 for $P_\tau(D^*)$.

5.4 $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$

The systematic uncertainty arising from the branching fractions of the $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ decays. For the light charged lepton modes $\ell^- = e^-, \mu^-$, it is estimated based on the uncertainties in the experimentally-measured branching fractions. For the decay modes which have not been experimentally measured, we change the branching fractions in the MC simulation continuously from 0% to 200% and take maximum differences of $R(D^*)$ and $P_\tau(D^*)$ from the nominal fit values as the systematic uncertainties. Consequently, we estimate the uncertainties to be $\pm 2.4\%$ for $R(D^*)$ and ± 0.020 for $P_\tau(D^*)$. Additionally, we take into account the uncertainty arising from the hadronic FFs, which is estimated to be $\pm 0.5\%$ for $R(D^*)$ and ± 0.033 for $P_\tau(D^*)$.

Table 5.3: Systematic uncertainty arising from the fake D^* component.

Source	$R(D^*)$	$P_\tau(D^*)$
Yield calibration factor	1.6%	0.016
Peaking background	3.0%	0.008
$M^2_{\rm miss}$ shape	0.1%	0.001
Total	3.4%	0.018

For the $\bar{B} \to D^{**}\tau^-\bar{\nu}_\tau$ modes, the branching fractions in Table 3.2 are varied continuously from 0% to 200%, and maximum differences of $R(D^*)$ and $P_\tau(D^*)$ from the nominal fit value are taken as the systematic uncertainties. They are estimated to be $\pm 1.1\%$ for $R(D^*)$ and ± 0.001 for $P_\tau(D^*)$.

5.5 $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$

There are two uncertainty sources related to $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$: the errors in the HQET parameters and the $M^2_{\rm miss}$ resolution correction. We assign the systematic uncertainties due to these two sources to be $\pm 2.3\%$ for $R(D^*)$ and ± 0.007 for $P_\tau(D^*)$.

5.6 Efficiency for τ daughters and Charged Leptons

Since the τ daughter is not the charged lepton in our analysis, the efficiency uncertainty in the τ daughter for the signal mode and the charged lepton for the normalization mode remains in $R(D^*)$. The uncertainties arising from the π ID, the π^0 reconstruction and the lepton ID efficiencies are estimated to be $\pm 1.9\%$ for $R(D^*)$ and ± 0.019 for $P_\tau(D^*)$.

5.7 $P_\tau(D^*)$ Correction Function

The parameters of the $P_\tau(D^*)$ correction function have uncertainties due to the MC statistics. The systematic uncertainties are estimated to be ± 0.01 for $P_\tau(D^*)$ and negligibly small for $R(D^*)$.

5.8 Total Reconstruction Efficiency

The total signal (normalization) reconstruction efficiencies have errors due to finite statistics of the MC simulation. These values have been shown in Table 4.2. We estimate the systematic uncertainties in the reconstruction efficiencies to be $\pm 1.0\%$ for $R(D^*)$ and ± 0.019 for $P_\tau(D^*)$.

5.9 Common Uncertainties between Signal and Normalization Modes

Uncertainties common for the signal and the normalization samples are largely canceled in Eq. 4.1 for $R(D^*)$ and in Eq. 3.6 for $P_\tau(D^*)$. These include the uncertainties in $N_{B\bar{B}}$, the tag efficiency correction, the D^* reconstruction efficiency and the branching fractions of the D decay modes. These uncertainties are 1.9%, 4.7%, 3.4% and 4.8%, respectively, and their quadratic sum is 7.8%. Since the yield of a part of the background mode such as $\bar{B} \to D^* D_s^{*-}$ is fixed using the MC expectation, we need to consider the residual effect from these uncertainties in our analysis. It is estimated to be $\pm 2.3\%$ for $R(D^*)$ and ± 0.021 for $P_\tau(D^*)$.

Chapter 6

Result and Discussion

After all the analysis procedures are complete, we perform a fit to the signal sample data. In the first section of this chapter, we present the fit results. In the second section, the obtained results are discussed and compared with the previous studies. In the third section, the results are interpreted in the context of the NP models. In the last section, we discuss the prospect for the Belle II experiment [78].

6.1 Fit result

Figures 6.1 and 6.2 show the fit results to the signal sample of the real data. (The fit to the normalization mode has been already shown in Fig. 3.29.) The p-value of the fit

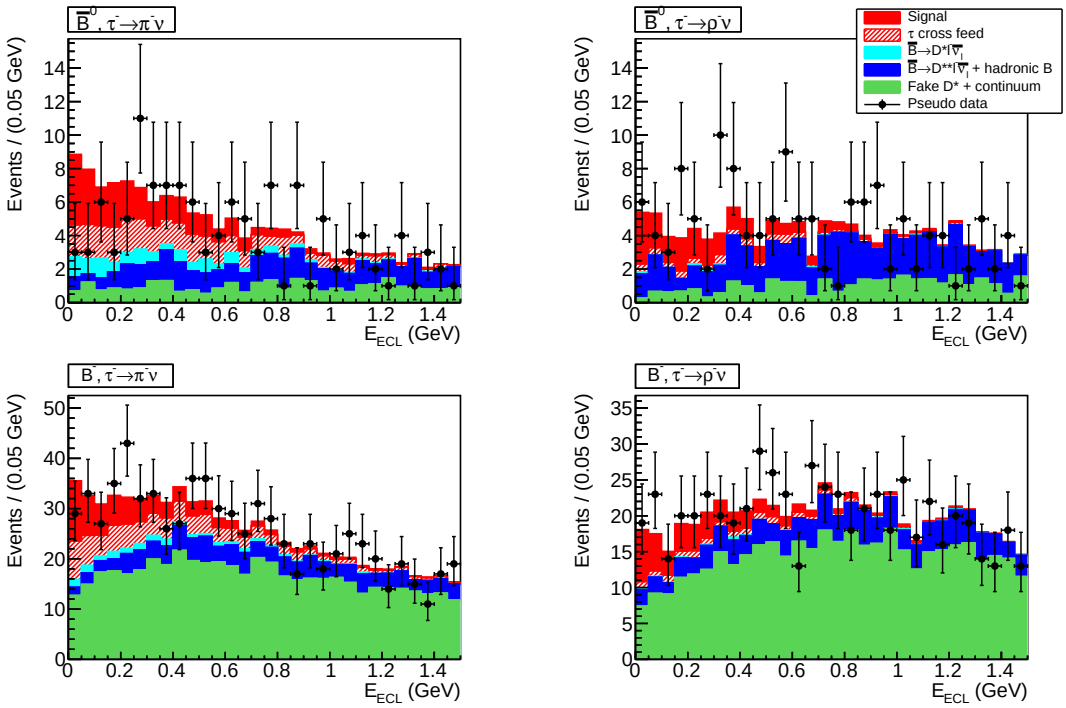


Figure 6.1: Fit results to the four signal sample, where two $\cos\theta_{\rm hel}$ regions are combined. In the "τ cross feed" component, the $\rho \leftrightarrow \pi$ cross feed and the other τ cross feed components are combined.

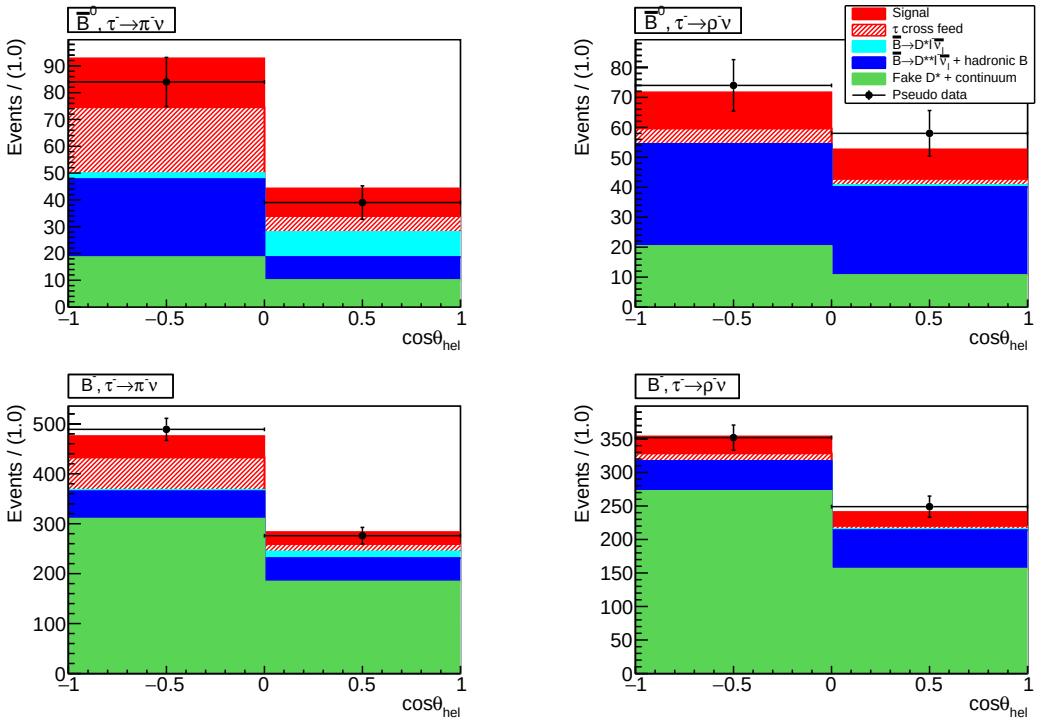


Figure 6.2: Projection of the fit result to the $\cos\theta_{\rm hel}$ axis.

to the signal sample is 29%. From the fit and the estimated systematic uncertainties in Table 5.1, we obtain

$$R(D^*) = 0.270 \pm 0.035(\text{stat.})^{+0.028}_{-0.025}(\text{syst.}), \tag{6.1}$$
$$P_\tau(D^*) = -0.38 \pm 0.51(\text{stat.})^{+0.21}_{-0.16}(\text{syst.}). \tag{6.2}$$

Our result is consistent with the SM prediction $R(D^*) = 0.252 \pm 0.003$ and $P_\tau(D^*) = -0.497 \pm 0.013$. With this value of $P_\tau(D^*)$, the efficiency variation discussed in Chapter 4 is estimated to be less than 1% and negligibly small.

The signal significance is calculated by $\sqrt{2\ln(L_{\rm max}/L_0)}$, where $L_{\rm max}$ and L_0 are the likelihood with the nominal fit and the null hypothesis, respectively. It is found to be 9.4σ only with the statistical error. Figure 6.3 shows the comparison of the $2\ln(L_{\rm max}/L)$ distribution, where L is a likelihood as a function of $R(D^*)$ and $P_\tau(D^*)$, and the quadratic function. Since the $2\ln(L_{\rm max}/L)$ distribution well agrees with the quadratic function, we assume that the uncertainties are Gaussian. The signal significance is therefore proportional to $1/\sigma^2$, where σ denotes the uncertainty of the measurement. Using $\sigma^2 = \sigma^2_{\rm stat.} + \sigma^2_{\rm syst.}$, we estimate the signal significance with the systematic uncertainty to be 7.0σ.

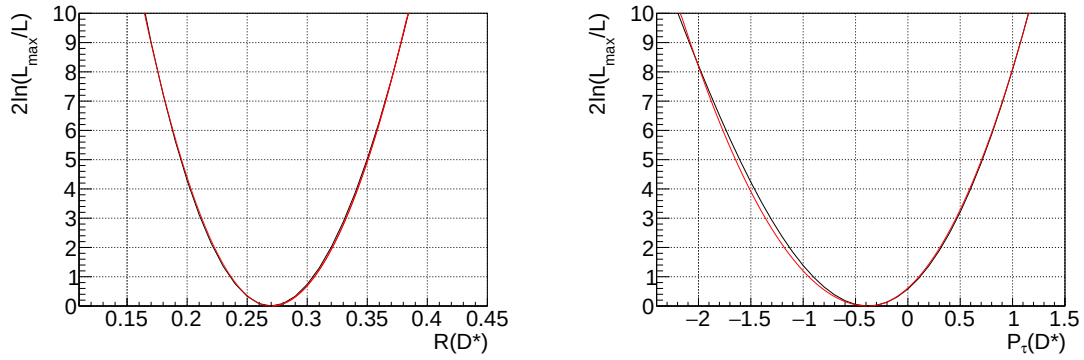


Figure 6.3: Comparison of the $2\ln(L_{\max}/L)$ distribution for $R(D^*)$ (left) and $P_\tau(D^*)$ (right), and the fitted quadratic function (solid red line). The quadratic functions are asymmetric with respect to the best-fit values of $R(D^*)$ and $P_\tau(D^*)$.

Separate Fit to B^- and $\bar{B}^0$

We also measure $R(D^*)$ and $P_\tau(D^*)$ by fitting to the B^- and the $\bar{B}^0$ samples independently. It results in

$$R(D^{*0}) = 0.296 \pm 0.048(\text{stat.}) \pm 0.031(\text{syst.}), \tag{6.3}$$
$$R(D^{*+}) = 0.227 \pm 0.050(\text{stat.})^{+0.037}_{-0.029}(\text{syst.}), \tag{6.4}$$
$$P_\tau(D^{*0}) = -0.64^{+0.64}_{-0.73}(\text{stat.})^{+0.33}_{-0.28}(\text{syst.}), \tag{6.5}$$
$$P_\tau(D^{*+}) = 0.16 \pm 0.85(\text{stat.}) \pm 0.025(\text{syst.}), \tag{6.6}$$

where the systematic uncertainties are estimated using the same method as discussed in Chapter 5. The obtained results from the B^- and the $\bar{B}^0$ samples are consistent.

Correlation between $R(D^*)$ and $P_\tau(D^*)$

In the fitting process, the statistical correlation is estimated to be 0.29. This correlation originates from Eq. 4.4. The systematics correlation is estimated using the relation

$$\rho_{\text{syst}} = \frac{\Sigma_i \rho_i \sigma_i^R \sigma_i^P}{\sigma_{\text{syst}}^R \sigma_{\text{syst}}^P}, \tag{6.7}$$

where ρ, σ^R and σ^P are the linear correlation factor, the $R(D^*)$ and the $P_\tau(D^*)$ uncertainties, respectively. The subscripts "syst" and "i" indicate the total systematic uncertainty and the ith source, respectively. Each correlation factor ρ_i is obtained in the estimation of the systematic uncertainties in Chapter 5. One example of the correlation in the calibration factor for the fake D^* yield is shown in Fig 6.4. We estimate the systematics correlation factor to be

$$\rho_{\text{syst}} = 0.55. \tag{6.8}$$

The total correlation factor is calculated by

$$\rho_{\text{exp}} = \frac{\rho_{\text{stat}} \sigma_{\text{stat}}^R \sigma_{\text{stat}}^P + \rho_{\text{syst}} \sigma_{\text{syst}}^R \sigma_{\text{syst}}^P}{\sigma_{\text{exp}}^R \sigma_{\text{exp}}^P}, \tag{6.9}$$
$$= 0.33, \tag{6.10}$$

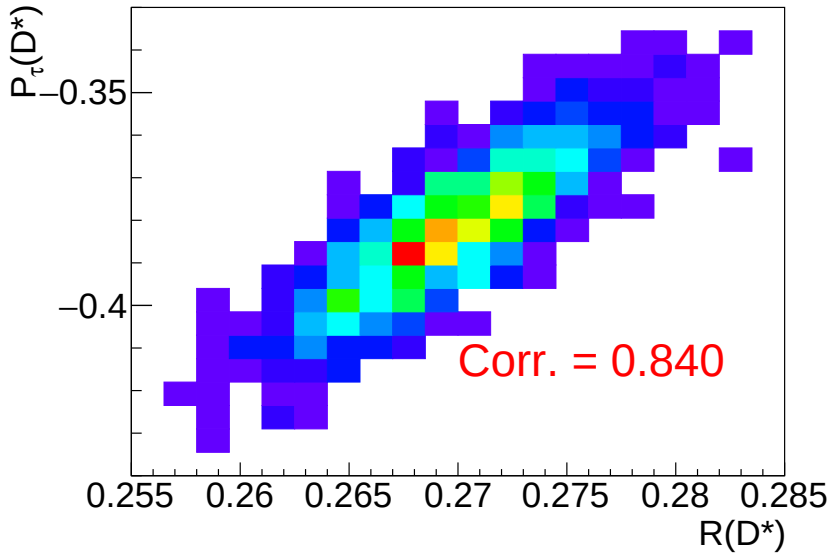


Figure 6.4: Correlation between $R(D^*)$ and $P_\tau(D^*)$, obtained by varying the yield calibration factor for the fake D^* yield. The value "Corr." is the linear correlation factor.

where $\sigma_{\rm exp}^{R(P)} = \sqrt{\left(\sigma_{\rm stat}^{R(P)}\right)^2 + \left(\sigma_{\rm syst}^{R(P)}\right)^2}$.

Using $\rho_{\rm exp}$, the χ^2 distribution is defined by

$$\chi^2 = \begin{pmatrix} \Delta^R & \Delta^P \end{pmatrix} \begin{pmatrix} (\sigma_{\rm exp}^R)^2 & \rho_{\rm exp}\sigma_{\rm exp}^R\sigma_{\rm exp}^P \\ \rho_{\rm exp}\sigma_{\rm exp}^R\sigma_{\rm exp}^P & (\sigma_{\rm exp}^P)^2 \end{pmatrix}^{-1} \begin{pmatrix} \Delta^R \\ \Delta^P \end{pmatrix}, \tag{6.11}$$

where

$$\Delta^R = [R(D^*)]_{\rm exp} - R(D^*), \tag{6.12}$$
$$\Delta^P = [P_\tau(D^*)]_{\rm exp} - P_\tau(D^*), \tag{6.13}$$

and $[R(D^*)]_{\rm exp}$ and $[P_\tau(D^*)]_{\rm exp}$ are the best-fit values of $R(D^*)$ and $P_\tau(D^*)$. Since the errors are treated as Gaussian, $\sqrt{\chi^2}$ is equivalent to the significance $\sqrt{2\ln L_{\rm max}/L}$. In the distribution of $\sqrt{\chi^2}$ illustrated in Fig. 6.5, we find that the region $P_\tau(D^*) > +0.5$ is excluded at 90% C.L. Taking into account the theoretical uncertainties in the SM predictions: $\sigma_{\rm th}^R = 0.003$ and $\sigma_{\rm th}^P = 0.013$, the consistency of our result with the SM prediction is estimated to be 0.4σ.

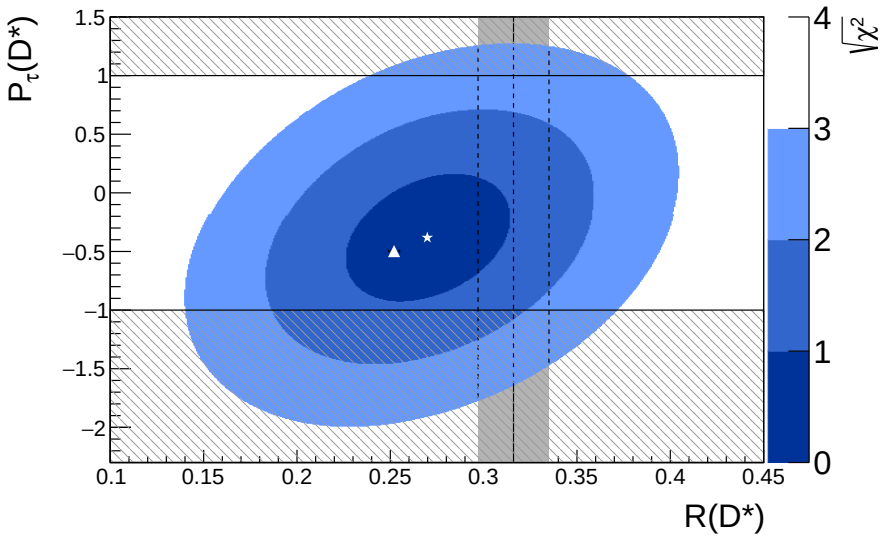


Figure 6.5: Comparison of our result (star for the best-fit value and the 1 to 3σ contours) with the SM prediction (triangle). The hatched regions are outside the physical $P_\tau(D^*)$ value. The vertical gray band shows the world-average $R(D^*)$ without our result [17].

6.2 Discussion

In Fig. 6.6, our $R(D^*)$ result is compared with the previous $R(D^*)$ measurements. Our result supports the world-average $R(D^*)$ within about 1σ and closest to the SM prediction among all the $R(D^*)$ measurements. Including our result, we calculate the new $R(D^*)$ average. The uncertainties arising from the semileptonic B decays are regarded as 100%-correlated among the measurements. The numerical values of $R(D^*)$ in the previous measurements and the estimated correlation factors are summarized in Table 6.1. (To use later in Sec. 6.3, the $R(D)$ and $P_\tau(D^*)$ measurements are also listed.) Our method gives $R(D^*) = 0.316 \pm 0.019$ as an $R(D^*)$ average without our result, and this is consistent with the world average reported by HFAG [17]. The new $R(D^*)$ average is calculated as

$$R(D^*) \quad = \quad 0.310 \pm 0.018. \tag{6.14}$$

This value is also shown in Fig. 6.6. The precision of the $R(D^*)$ average is improved by about 5%. The $R(D^*)$ discrepancy between the experiments and the SM prediction is estimated by $[R(D^*)_{\rm exp} - R(D^*)_{\rm SM}]/\sqrt{\sigma^2_{\rm exp} + \sigma^2_{\rm th}}$. It is slightly decreases from 3.3σ to 3.2σ. The combination of the three results at Belle is also calculated to be

$$R(D^*) \quad = \quad 0.292 \pm 0.022, \tag{6.15}$$

where the consistency with the SM is 1.8σ.

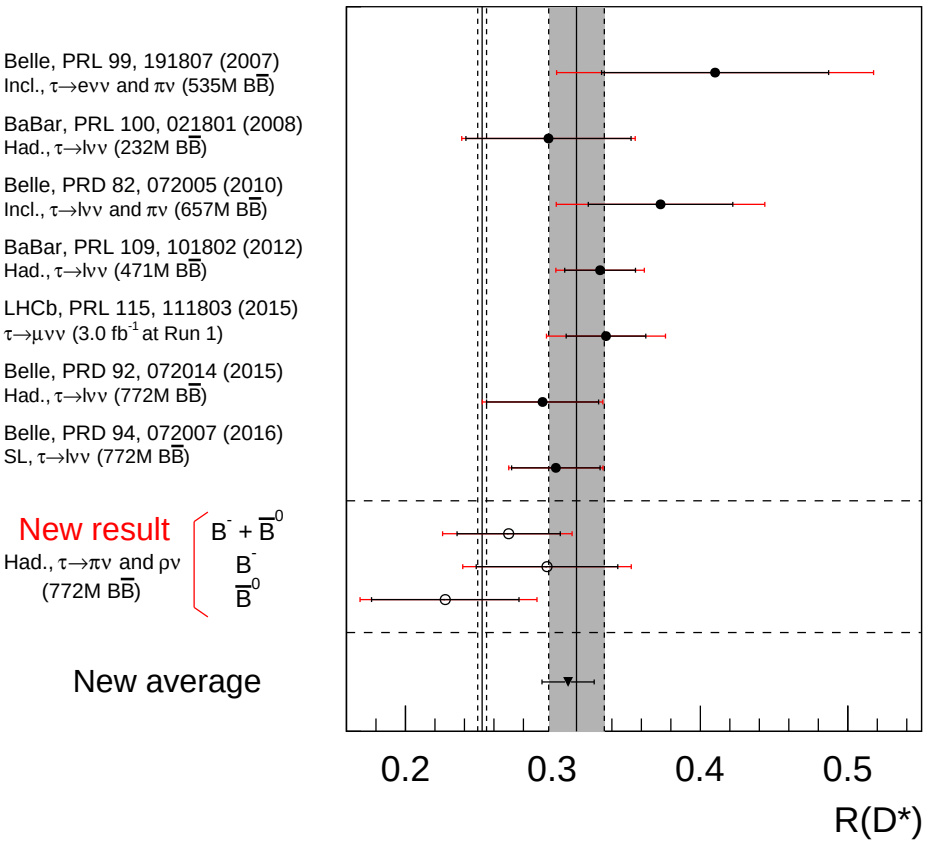


Figure 6.6: Comparison of our $R(D^*)$ result with the previous $R(D^*)$ measurements. The black and the red error bars show the statistical and the total uncertainties, respectively, except for the new average, where only the total uncertainty is indicated. For each measurement, the method (Incl., Had. and SL represent the inclusive tag, the hadronic tag, the semileptonic tag, respectively), the reconstructed τ decay modes and the size of the data sample are shown. LHCb uses proton-proton collisions at LHC and hence the measurement method is different from the tagging method. The integrated luminosities of 1.0 fb^{-1} at $\sqrt{s} = 7$ TeV and 2.0 fb^{-1} at $\sqrt{s} = 8$ TeV are used.

Table 6.1: Numerical values of $R(D)$, $R(D^*)$ and $P_\tau(D^*)$ in the measurements of $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$, and linear correlation factors among them. The column "Value" shows the result of each measurement, where the errors are the total uncertainties. BaBar $R(D^{(*)})$ is taken from Ref. [45], LHCb $R(D^*)$ from Ref. [49], Belle $R(D^{(*)})$ from Ref. [42] and Belle $R(D^*)$ (sl) from Ref. [43]. New $R(D^*)$ and $P_\tau(D^*)$ are our results.

	Value	BaBar $R(D)$	BaBar $R(D^*)$	LHCb $R(D^*)$	Belle $R(D)$	Belle $R(D^*)$	Belle $R(D^*)$ (sl)	New (D^*)	New $P_\tau(D^*)$
BaBar $R(D)$	0.440 ± 0.072	1.00	-0.27	0.05	0.02	0.05	0.03	0.04	0.02
BaBar $R(D^*)$	0.332 ± 0.030	-0.27	1.00	0.08	0.03	0.08	0.05	0.07	0.03
LHCb $R(D^*)$	0.336 ± 0.040	0.05	0.08	1.00	0.02	0.05	0.03	0.05	0.02
Belle $R(D)$	0.375 ± 0.069	0.02	0.03	0.02	1.00	-0.56	0.01	0.02	0.01
Belle $R(D^*)$	0.293 ± 0.041	0.05	0.08	0.05	-0.56	1.00	0.03	0.05	0.02
Belle $R(D^*)$ (sl)	0.302 ± 0.032	0.03	0.05	0.03	0.01	0.03	1.00	0.03	0.01
New (D^*)	$0.270 \pm^{+0.045}_{-0.043}$	0.04	0.07	0.05	0.02	0.05	0.03	1.00	0.33
New $P_\tau(D^*)$	$-0.38 \pm^{+0.55}_{-0.53}$	0.02	0.03	0.02	0.01	0.02	0.01	0.33	1.00

6.3 Constraint on NP Parameters

Using our result and the previous $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ measurements, we estimate constraints on NP parameters. In the presence of NP effects, the signal reconstruction efficiency varies due to modification of the kinematic distributions. However, as shown in Chapter 4, such effects are only $\mathcal{O}(1\%)$ in our analysis. This is sufficiently small compared to the total experimental uncertainty of 16%. We therefore assume that the efficiency does not change due to NP. For the measurements in Refs. [42, 43, 45, 49], we take into account the efficiency variation if it is reported in the reference. Otherwise, we assume that the efficiency is constant over the theoretical parameter space.

To test the NP scenarios, we define χ^2 as

$$\chi^2 = A^T C^{-1} A, \tag{6.16}$$

$$A = \begin{pmatrix} \Delta_1 \\ \Delta_2 \\ \vdots \\ \Delta_n \end{pmatrix}, \tag{6.17}$$

$$\Delta_i = (\mathcal{O}_{\rm exp} - \mathcal{O}_{\rm th}), \tag{6.18}$$

$$C = \begin{pmatrix} \left(\sigma^1_{\rm exp}\right)^2 + \left(\sigma^1_{\rm th}\right)^2 & \rho_{12}\sigma^1_{\rm exp}\sigma^2_{\rm exp} & \cdots & \rho_{1n}\sigma^1_{\rm exp}\sigma^n_{\rm exp} \\ \rho_{21}\sigma^2_{\rm exp}\sigma^1_{\rm exp} & \left(\sigma^2_{\rm exp}\right)^2 + \left(\sigma^2_{\rm th}\right)^2 & \cdots & \rho_{2n}\sigma^2_{\rm exp}\sigma^n_{\rm exp} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1}\sigma^n_{\rm exp}\sigma^1_{\rm exp} & \rho_{n2}\sigma^n_{\rm exp}\sigma^2_{\rm exp} & \cdots & \left(\sigma^n_{\rm exp}\right)^2 + \left(\sigma^n_{\rm th}\right)^2 \end{pmatrix} \tag{6.19}$$
$$(\rho_{ij} = \rho_{ji}),$$

where $\mathcal{O}_{\rm exp(th)}$ and $\sigma_{\rm exp(th)}$ denote the experimental (theoretical) value of the observable: $\mathcal{O} = R(D)$, $R(D^*)$, $P_\tau(D^*)$, and their uncertainty, respectively. The linear correlation factors ρ_{ij} are given in Table 6.1.

6.3.1 Type-II 2HDM

As introduced in Sec. 1.2.3, type-II 2HDM is the NP model with the charged Higgs boson. Although BaBar reported that this model was excluded at 99.8% C.L. [45], the exclusion is currently determined mostly by the single measurement. It is therefore important to consider the possibility of type-II 2HDM with including our result and the other measurements.

Type-II 2HDM has the Wilson coefficient C_{S_1} as shown in Eq. 1.44. We use $m_b = 4.2$ GeV/c^2 and $m_c = 0.901$ GeV/c^2 [77] for the theoretical calculation. The efficiency variation is treated as below.

- Use the results in Ref. [45] for the measurement by BaBar, and Ref. [43] for the measurement with the semileptonic tag by Belle.
- Extrapolate the two measured point at $\tan\beta/m_{H^\pm} = 0$ and 0.5 $(\text{GeV}/c^2)^{-1}$ to the other points up to $\tan\beta/m_{H^\pm} = 1.0$ $(\text{GeV}/c^2)^{-1}$ in Ref. [42] for the measurement of Belle with the hadronic tag and the leptonic τ decays.

- Assume no efficiency variation in the measurement by LHCb [49] and our measurement.

Figure 6.7 (top) shows the estimated constraint on the plane for $\tan\beta$ and $m_{H^\pm}$ using the results at Belle. The C.L. for exclusion is calculated from χ^2 divided by the number of degrees of freedom. Compared to the previous constraint, the 3-σ excluded region around $\tan\beta/m_{H^\pm} = 0$ $(\mathrm{GeV}/c^2)^{-1}$ (the SM point) slightly decreases. This is because our measurement is close to the SM prediction with less than 1σ. The smallest χ^2 value corresponds to about 80% C.L. exclusion at $\tan\beta/m_{H^\pm} = 0$ $(\mathrm{GeV}/c)^{-1}$, and the second minimum of χ^2 is found around $\tan\beta/m_{H^\pm} = 0.5$ $(\mathrm{GeV}/c)^{-1}$ at 84% C.L. exclusion.

Figure 6.7 (bottom) is combination of all the experimental results of $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$. The entire region has been excluded at 99.0% C.L.

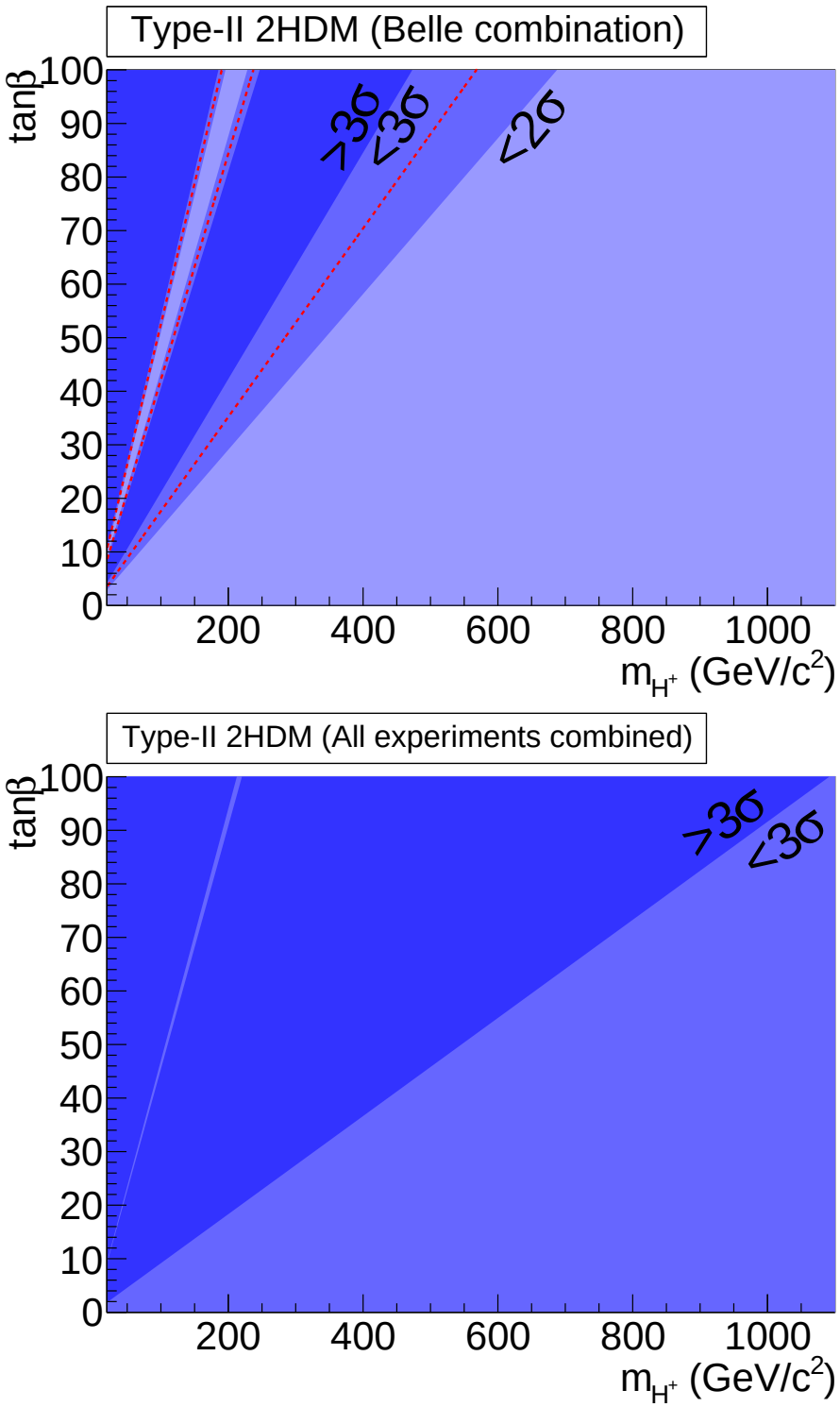


Figure 6.7: Constraint on the $m_{H^\pm} - \tan\beta$ plane from the results at Belle (top) and from all the experiments (bottom). In the top panel, the dashed-red lines indicate boundary on the 3-σ excluded region without the new result presented in this thesis.

6.3.2 Leptoquark Model

As described in Sec. 1.2.3, we consider the scalar leptoquark with $m_{LQ} = 1$ TeV/c^2. Since most of the analyses do not study the efficiency variation with the leptoquark models, we assume that the efficiency does not differ from the SM over the parameter space. Figures 6.8 and 6.9 show the constraints on the Wilson coefficients of R_2 and the S_1 leptoquark models, where the constraint from the results at Belle and that from all the results are compared. The smallest χ^2 of the R_2 and the S_1 models are found at $C_{S_2} = -0.09 \pm 0.67i$ and at $C_{S_2} = -1.7$, respectively. The R_2 model favors a significantly large imaginary part, that may indicate a large CP violation in the decay $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$.

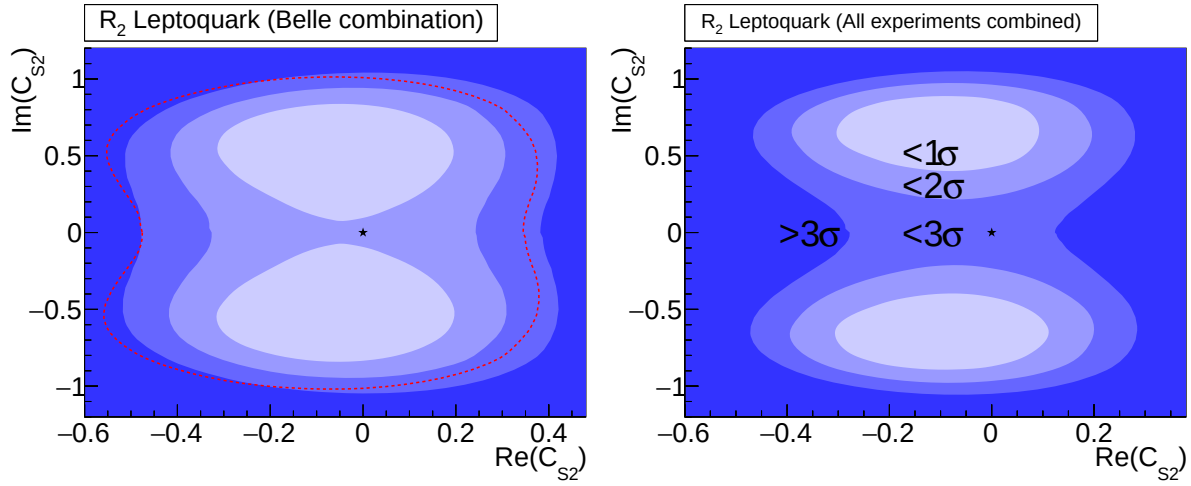


Figure 6.8: Constraint on the real part and the imaginary part of C_{S_2} for the R_2 leptoquark model. The left and the right panels include the results from Belle and the results from all the experiments, respectively. The stars show the SM point, where $\mathrm{Re}(C_{S_2}) = \mathrm{Im}(C_{S_2}) = 0$. The dashed-red line in the left panel indicates the boundary on the 3-σ excluded region without the new result presented in this thesis.

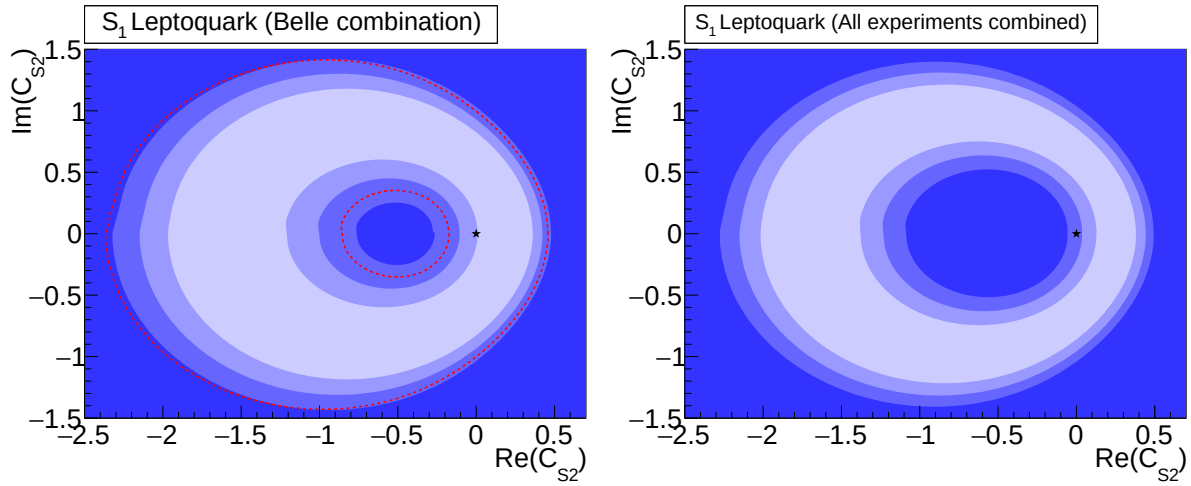


Figure 6.9: Constraint on C_{S_2} for the S_1 leptoquark model. Meaning of each panel, marker and line is the same as Fig. 6.8.

6.4 Prospect for the Belle II Experiment

6.4.1 Detector Upgrade for Belle II

The Belle experiment is upgraded to the Belle II experiment with the SuperKEKB accelerator [78]. The SuperKEKB aims at the instantaneous luminosity of 8×10^{35} cm^{2-}s^{-1}, which is approximately 40 times higher than that of KEKB. Integrated luminosity of 50 ab^{-1} will be accumulated.

Not only the accelerator, but also the sub-detectors are upgraded. The four-layer SVD system is replaced by the six-layer vertex detector (VXD) system. The inner-two layers are composed of the DEPFET pixel detector with an approximate pixel size of 50×50 μm^2 [79]. The inner radius is 14 mm, which is much smaller than 20 mm for the Belle SVD. The closer distance to the IP improves the impact parameter resolution and the efficiency for low-momentum tracks. Outer four layers are the upgraded SVD system [80], where the outer radius is expanded from 88 mm to 140 mm. The larger SVD volume provides a better K^0_S reconstruction efficiency. The PID devices, the ACC and the TOF, are replaced by the time-of-propagation (TOP) counter [81, 82, 83, 84] for the barrel region and the aerogel ring-imaging Cherenkov (ARICH) counter [85] for the forward-endcap region. By exploiting ring images of the Cherenkov light, these new PID devices achieve better the PID efficiency for $K^{\pm}$ and $\pi^{\pm}$ than the PID system of Belle.

6.4.2 Time-Of-Propagation Counter

I have been leading the development for the TOP counter, which plays an important role in the PID for the barrel region. As shown in Fig. 6.10, the TOP counter consists of a $20 \times 450 \times 2710$ mm^2 quartz bar and an array of the 16×2 square-shaped micro-channel-plate (MCP) PMTs at the end of the bar. In total, 16 TOP modules compose a cylindrical structure on the barrel region of the Belle II detector. Using the TOP counter, we target to improve the PID efficiency for $K^{\pm}$ and $\pi^{\pm}$ by about 5% per track with $\mathcal{O}(1)$ GeV/c while reducing the fake rate by a factor of about 5.

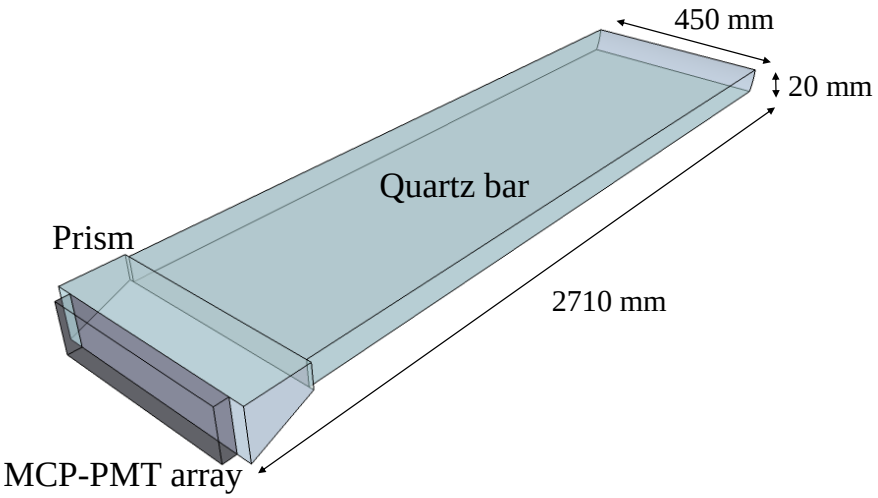


Figure 6.10: Drawing of the TOP counter.

The quartz bar is the Cherenkov radiator for high-momentum charged particles. Since the bar surface is polished accurately with a roughness of less than 5 Å, the photons propagate inside the quartz bar with conserving information of the Cherenkov angle θ_c. These photons are detected by the MCP-PMTs. Compared to the TOF counter, the TOP counter has an advantage that the time of photon propagation is used in addition to the time of flight. The difference of the photon arrival time is about 100 ps for 3 GeV/c $K^{\pm}$ and $\pi^{\pm}$, and the number of photons reaching the bar end is around 100. To achieve the target performance of the TOP counter, each photon needs to be detected with a time resolution less than 50 ps and a good efficiency. The MCP-PMT, having a single-photon time resolution of typically 30 ps, is a unique solution for photosensors of the TOP counter, and it is a key component of the TOP performance.

We have developed a square-shaped MCP-PMT shown in Fig. 6.11 (left) in collaboration with Hamamatsu Photonics K.K [86, 87, 88, 89]. The square shape achieves a large effective area for the photon detection in an array. Figure 6.11 (right) shows the cross section of the MCP-PMT. As electron dynodes, it contains two 400-μm-thick MCPs, which are made of lead glass with 10-μm-diameter pores. About 1000 V is applied between the surface of the MCP. Electrons are multiplied up to $\mathcal{O}(10^6)$ by the strong electric field in the pore. The short electron path of about 5 mm between the photocathode to the anode realizes an excellent transit-time spread. It also provides a tolerance to the magnetic field perpendicular to the window. This is an essential property because the TOP modules are immersed in the 1.5 T magnetic field provided by the superconducting solenoid of the Belle II detector.

Performance of the MCP-PMT in 1.5 T

For the TOP counter operation, it is important to well understand performance of the MCP-PMTs. To conduct the performance evaluation for all the 512 MCP-PMTs in a magnetic field before assembling them to the TOP module, I have constructed the measurement system using a 1.5 T resistive dipole magnet as shown in Fig. 6.12. The mea-

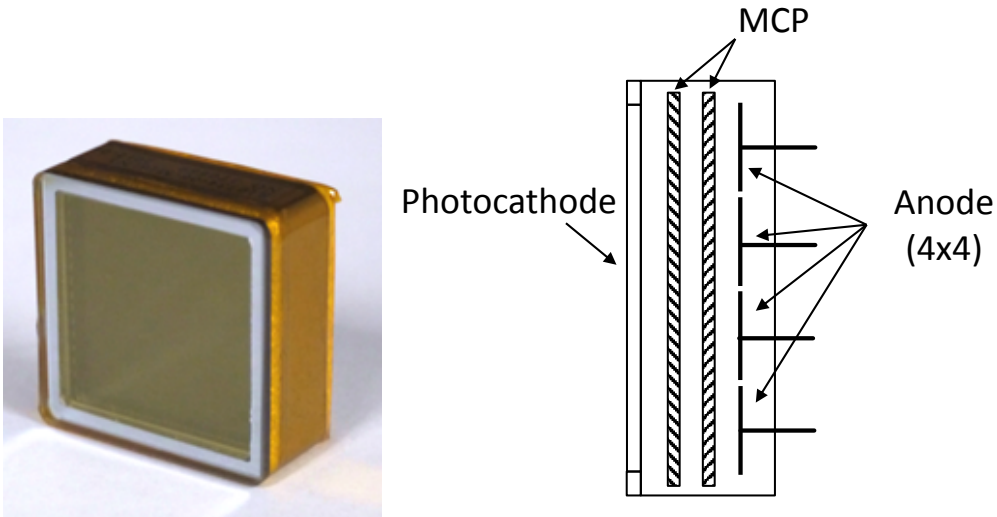


Figure 6.11: (left) Photograph of the MCP-PMT. (right) Cross section of the MCP-PMT. Slant lines on two MCPs represent pores with a bias angle of 13 degrees.

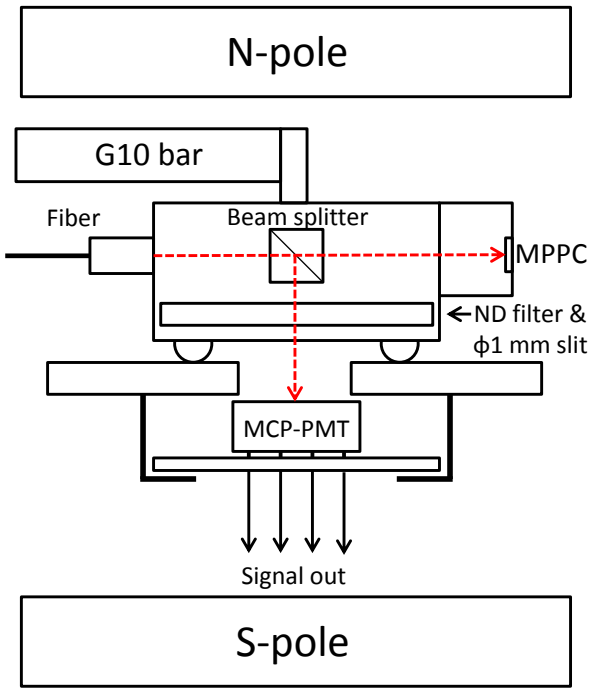


Figure 6.12: Schematic drawing of the jig for the MCP-PMT measurement in a magnetic field [90]. Dashed red lines show the path of the laser light.

surement jig in the dipole magnet consists of two parts. The upper part introduces light from a laser (Hamamatsu Picosecond Light Pulser PLP-02-040; a pulse width is about 20 ps and a wavelength is 400 nm) onto the MCP-PMT. The light intensity is reduced to be single-photon level by neutral-density (ND) filters. The aperture is collimated to be ϕ 1 mm by a slit. The bottom part hold an MCP-PMT in the magnetic field. These two parts contact each other by four teflon balls on the teflon plate. To cope with the strong magnetic field, the jig is made of polyacetal with plastic or brass screws. Through a 2-m-long G-10 bar, the laser part is connected to a motorized stage movable for orthogonal two directions, located outside the magnetic field. Owing to the low-friction contact of the laser part with the MCP-PMT holder by teflon, the light irradiation position to the MCP-PMT can be adjusted smoothly by controlling the stage. The multi-pixel photon counter (MPPC) monitors the laser intensity for measurement of a single-photon detection efficiency.

Figure 6.13 (left) shows the relative gain in 1.5 T with respect to that in 0 T for 488 MCP-PMTs. There are two types of the MCP-PMTs; one is called conventional MCP-PMT with lead-glass MCPs, and the other is called ALD (atomic layer deposition) MCP-PMT with ALD-coated MCPs. The ALD MCP-PMT is developed to improve the MCP-PMT lifetime; it is described in the next section. Although the gain drops by 90% at the maximum in 1.5 T, all the MCP-PMTs are confirmed to have a gain greater than 5×10^5. This gain is sufficient to detect single photon signals without losing the efficiency. Figure 6.13 (right) shows the result of the time resolution measurement of single photon detection for 115 MCP-PMTs in 1.5 T. All the MCP-PMT channels show the time resolution better than 50 ps, that satisfies the requirement. For the efficiency, about 5% drop is observed in 1.5 T relative to 0 T [91]. Degradation of the PID performance due

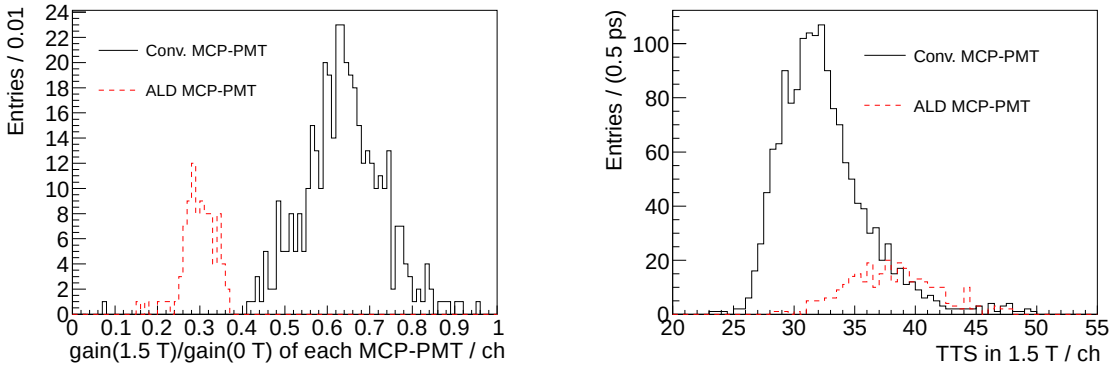


Figure 6.13: Results from the gain (left) and the time resolution (right) measurements in 1.5 T [92]. One entry indicates a measurement of one channel.

to this efficiency drop is negligible.

MCP-PMT lifetime

It is known that the quantum efficiency (QE) of the MCP-PMT degrades during operation [87, 89]. This is induced by the residual gas absorbed on the surface of the MCPs. The gas molecules are released from the surface by electron collisions. The released gas molecules reach the photocathode through paths in a gap between the MCPs and the tube, and deteriorate the photocathode. Since the total amount of the released residual gas increases depending on the total number of output electrons, the MCP-PMT lifetime is determined as a function of the integrated output charge. In the Belle II detector, the largest source of the Cherenkov photons is the high energy background photons originating from the beam. The high energy photons create electron-positron pairs in the quartz bar, and these electrons and positrons emit Cherenkov light. In our previous study and development, we have successfully improved the MCP-PMT lifetime from $\mathcal{O}(0.1)$ C/cm^2 to around 1 C/cm^2 [89]. However, in the present background estimation, the integrated output charge until 50 ab^{-1} data are accumulated is estimated to be 5 to 8 C/cm^2 for operation at a gain of 5×10^5 [93].

To study the MCP-PMT lifetime systematically, I have constructed the lifetime measurement system shown in Fig. 6.14 [92]. In the dark box, two light sources are placed. The LED with a wavelength of 400 nm illuminates up to eight MCP-PMTs at the same time and accelerate charge output. The intensity is adjusted for the output current from the MCP-PMT to be less than 2 μA/cm^2 to avoid gain saturation. It takes roughly two weeks to accumulate 1 C/cm^2. Occasionally, the LED is turned off and the laser light from a pulse laser (Advanced Laser Diode System PIL040 Picosecond Injection Laser; wavelength is 400 nm) is irradiated. The laser intensity is reduced to be single-photon level. The number of laser hits is proportional to the QE. The QE is monitored relatively to the initial QE. The variation of the laser intensity is corrected by a line-focus type PMT (Hamamatsu H1949). The line-focus PMT is turned off during LED illumination to avoid QE drop.

Figure 6.15 shows the relative QE with respect to the initial QE as a function of the integrated output charge, obtained from three conventional MCP-PMTs and three ALD MCP-PMTs. We define the lifetime by the integrated output charge at which the relative

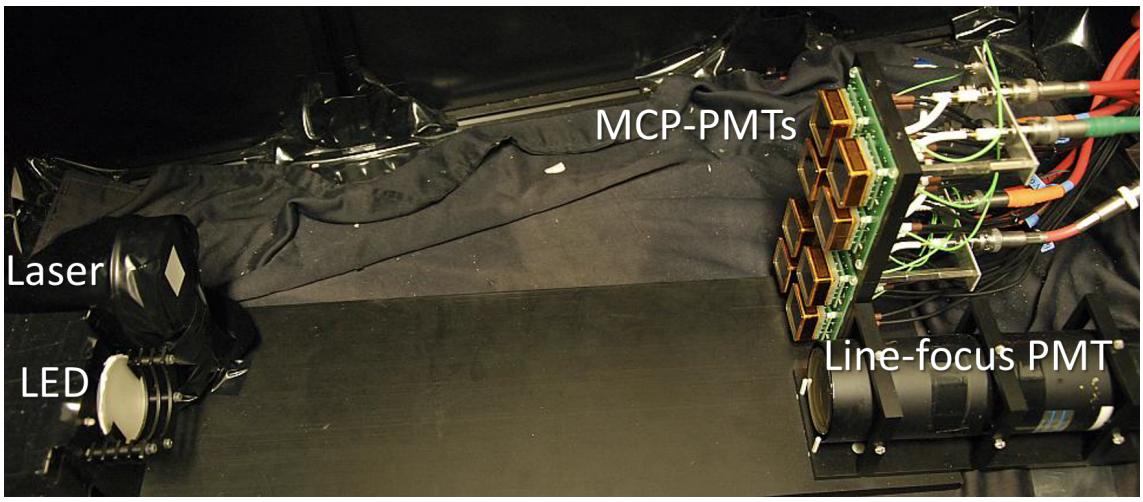


Figure 6.14: Photograph of the lifetime measurement system.

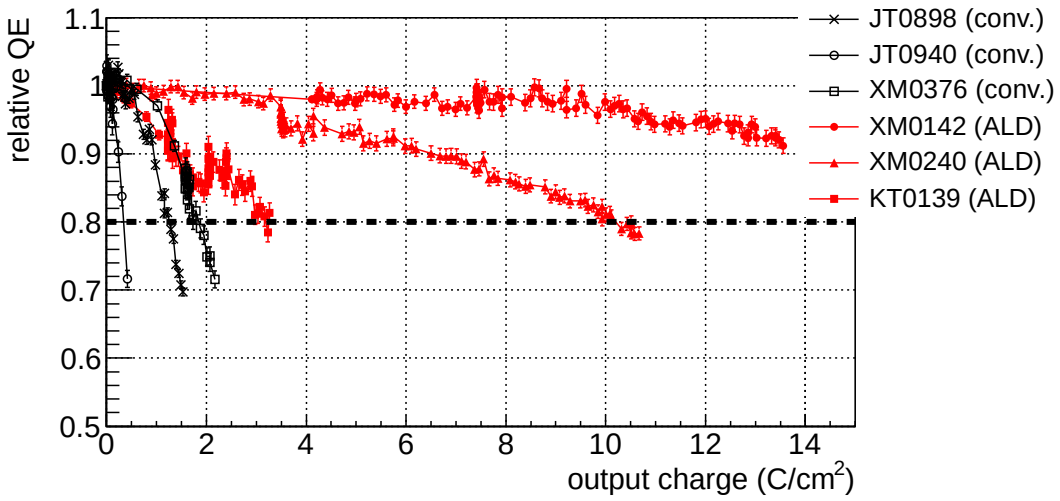


Figure 6.15: Result from the lifetime measurement for three conventional MCP-PMTs (indicated as "conv.") and three ALD MCP-PMTs (indicated as "ALD") [92]. The horizontal dashed line shows the relative QE equal to 0.8.

QE becomes 0.8. Our lifetime studies find that the lifetime of the MCP-PMT is extended by a factor of about ten compared to the conventional MCP-PMT. The dramatic lifetime improvement by the ALD technique is as expected by the previous report in Ref. [94]. This is presumably because the out gas from the MCP surface is suppressed by the atomic layer. On the other hand, the lifetime significantly varies from 2.5 to more than 14 C/cm^2. The lifetime is further improved to be 15 to 35 C/cm^2 by reducing residual gas in the production process [95].

For the TOP modules, we use 288 ALD MCP-PMTs, which are expected to have sufficient lifetime over the Belle II operation. Remaining 224 are conventional MCP-PMTs, which will be replaced during the experiment.

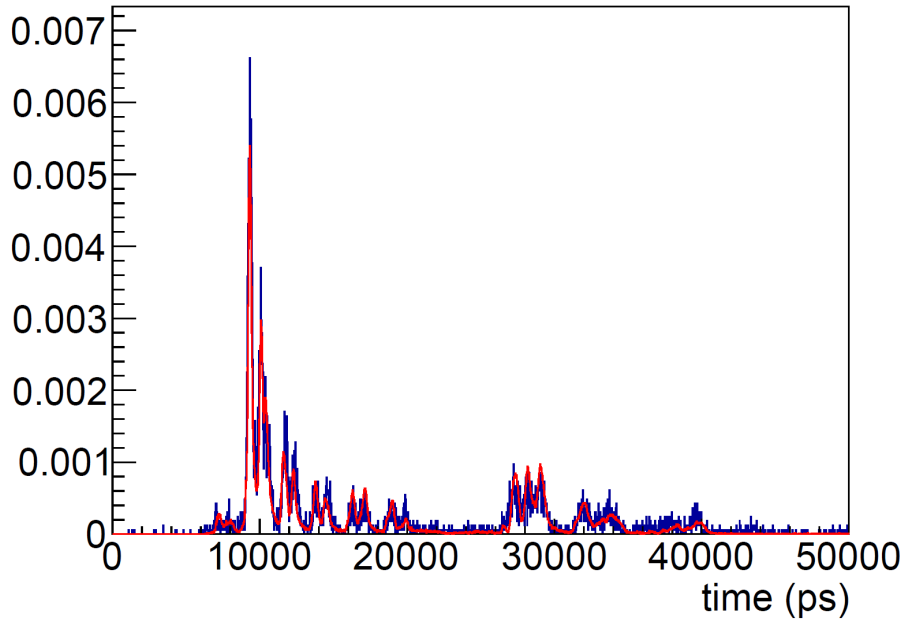


Figure 6.16: Distribution of the photon detection timing taken at the beam test (blue histogram), where the distribution expected by the MC simulation is overlaid (red line) [92]. Both distributions are normalized to unit area.

TOP counter performance with test beam

To evaluate performance of the TOP counter, a beam test is conducted using the LEPS (Laser Electron Positron Facility at SPring-8) [96]. At LEPS, 355-nm laser is injected to the storage ring of the 8 GeV electron beam. Photons with energy up to 2.4 GeV are obtained by the backward Compton scattering and converted to electron-positron pairs in a lead target. The electrons and positrons are swept by a 0.7 T dipole magnet. By a pair of the trigger counters with a size of 5×5 mm^2, 2 GeV positrons are selected. Further details of the beam test setup is described in Refs. [97, 98].

Figure 6.16 shows the comparison of distributions of the photon detection timing obtained from the data and the MC simulation in one MCP-PMT channel. The agreement of the peak position, timing and tail shape demonstrates good understanding of the optics as well as characteristics of the MCP-PMT. Interpretation of the experimental results are discussed in Refs. [97, 98] in detail.

6.4.3 $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ Study at Belle II

The improved hadronic tag algorithm for Belle II has been developed [99]. The total number of B decay chains used in the full reconstruction, which increases the tag efficiency by about 30 to 50%. In addition, the excellent tracking efficiency by the VXD and the PID efficiency by the PID devices contribute to further improvement of the tag efficiency by a factor of two. In total, the new tagging algorithm provides about three times better tagging efficiency than the current algorithm at Belle. Precision measurements of $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ using about 100 times larger signal sample will be possible at Belle II.[1] With this

[1] From the 50 times larger integrated luminosity and the three times better tagging efficiency, we expect about 150 times larger signal sample. However, it is uncertain how accurate the current MC simulation

statistics, the systematic uncertainty will dominate the total uncertainty.

To discuss the prospect for Belle II, we divide the systematics sources of our measurement into three components and consider possible improvements.

Category 1. Reducible component

Statistics of the data available for calibration increases proportional to the integrated luminosity. Also, the MC sample is produced so that the MC statistical errors are sufficiently small with respect to the statistics of data. We therefore assume that the systematic uncertainties arising from the hadronic B composition (calibrated part using the data), the fake D^* yield calibration factor and the MC statistics for the PDF construction are improved proportional to the square-root of the integrated luminosity.

Category 2. Semi-reducible component

The other uncertainty arising from the branching fractions of the D meson and the background B decay modes can be improved in the future measurements. However, their improvement will be limited by the systematic uncertainty of measurements. We therefore assume that the improvement will be at most by a factor of 2. This component includes the branching fractions of the D meson, the $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ decay and the hadronic B decays which have been already measured experimentally.

Category 3. Irreducible component

The uncertainty arising from unmeasured background modes may not be improved so much because there are a huge number of exclusive modes. Some of them are difficult to be experimentally reconstructed due to multiple neutral light mesons. To estimate the systematic uncertainties conservatively, we consider that these uncertainties will not be improved. This component includes the unmeasured parts of the hadronic B composition, $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ and all the other sources.

The estimated uncertainties based on the above assumption are summarized in Table 6.2. Figure 6.17 shows the comparison of the current allowed regions on the type-II 2HDM

is. As discussed in Chapter 3, we have observed about 30% discrepancy of the tag efficiency between the data and the Belle MC simulation. We conservatively assume that the statistical increase is a factor of 100.

Table 6.2: Estimates of the experimental uncertainties at Belle II. The value in the left (right) side of the arrow represents the current uncertainty (predicted uncertainty for Belle II).

	$\delta R(D^*)$	$\delta P_\tau(D^*)$
	Belle $\to$ Belle II	Belle $\to$ Belle II
Statistical	$\pm 0.035 \to \pm 0.0035$	$\pm 0.51 \to \pm 0.051$
Syst. category 1	$^{+0.016}_{-0.010} \to ^{+0.0016}_{-0.0010}$	$^{+0.18}_{-0.12} \to ^{+0.018}_{-0.012}$
Syst. category 2	$\pm 0.009 \to \pm 0.0045$	$\pm 0.025 \to \pm 0.013$
Syst. category 3	$\pm 0.021 \to \pm 0.0210$	$^{+0.11}_{-0.10} \to ^{+0.110}_{-0.100}$
Total	± 0.022	$^{+0.12}_{-0.11}$

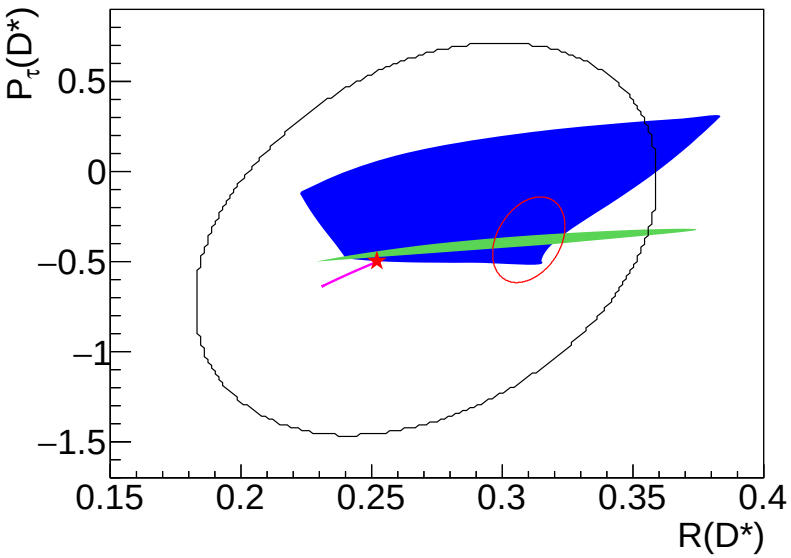


Figure 6.17: Predicted constraints at Belle II. The black and the red ellipses are the result presented in this thesis and the expected precision at Belle II, respectively. The center values of $R(D^*)$ and $P_\tau(D^*)$ are set at the current average of the $R(D^*)$ measurements at Belle. The star is the SM prediction. The magenta line, the blue region and the green region show the current allowed regions for the type-II 2HDM, the R_2 leptoquark and the S_2 leptoquark models, respectively. All the regions are indicated by 95% C.L.

and two leptoquark models with the result presented in this thesis and the prospects for Belle II. The allowed region on the $R(D^*) - P_\tau(D^*)$ plane is estimated based the constraints from the measurements of Belle, discussed in Secs. 6.3.1 and 6.3.2. For $R(D^*)$, improvements from other measurements such as analyses using the $\tau^-\ell^-\bar{\nu}_\ell\nu_\tau$ decays and the semileptonic tag are taken into account. We find that there is a region on which the $P_\tau(D^*)$ measurement provides unique constraints.

Chapter 7

Conclusion

Using the data sample containing $(7.72 \pm 0.11) \times 10^8$ $B\bar{B}$ pairs accumulated with the Belle detector, I conduct the measurement of $R(D^*)$ and $P_\tau(D^*)$ in the decay $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ with the τ decays $\tau^- \to \pi^- \nu_\tau$ and $\rho^- \nu_\tau$.

A challenge of this analysis is the measurement of $\cos\theta_{\rm hel}$, which is defined in the rest frame of τ. Due to one unconstrained degree of freedom, the τ momentum is not fully determined. In this study, I have established a method to obtain $\cos\theta_{\rm hel}$ using the symmetry of the kinematic cone around the τ daughter in the rest frame of W^*. Another challenge is the estimation of the background from various hadronic B decays. I determine the composition of the seven major modes by reconstructing B mesons from the calibration data sample.

I finally obtain the results

$$\begin{aligned} R(D^*) &= 0.270 \pm 0.035({\rm stat.})^{+0.028}_{-0.025}({\rm syst.}), \\ P_\tau(D^*) &= -0.38 \pm 0.51({\rm stat.})^{+0.21}_{-0.16}({\rm syst.}). \end{aligned}$$

This result is consistent with the SM prediction within 0.4σ. This is a new $R(D^*)$ measurement independent of the previous studies, and the first measurement using only hadronic τ decays. The precision of my result is comparable to the previous measurements using leptonic τ decays. Including my $R(D^*)$ result, the new world-average $R(D^*)$ is estimated to be 0.310 ± 0.018, and the precision is improved by 5% from the previous world average. The current discrepancy from the SM prediction is 3.2σ, which is slightly smaller than the previous discrepancy. The $P_\tau(D^*)$ region greater than $+0.5$ is excluded at 90% C.L., which is the first constraint on $P_\tau(D^*)$.

In the Belle II experiment, 50 times larger data sample than that in Belle and the improved hadronic tag algorithm will be available. Therefore measurements of kinematics in $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ as well as the conventional $R(D^*)$ measurement will provide opportunities to precise SM tests and NP searches. I estimate the precision of $R(D^*)$ and $P_\tau(D^*)$ measurements and find that there is parameter space on which $P_\tau(D^*)$ provides unique constraints. At Belle II, using the $P_\tau(D^*)$ measurement method I have established in this study, more precise NP searches using $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ will be performed.

Acknowledgements

First of all, I would like to express my great gratitude to Prof. Toru Iijima for giving me an opportunity to be involved in the research in high energy physics. Thanks to his encouragement and advice, I learned various skills in the high energy physics, which are not only related to this thesis but are also important in many other aspects, including technique for hardware, software and computing as well as deep understanding of physics.

In this research, many helps from and discussions with Dr. Yutaro Sato were significantly important. Without his technical advice and detailed suggestions, I was not able to complete this challenging data analysis. Also, members of the Belle physics analysis group in Nagoya, Dr. Alessandro Gaz, Dr. Kenji Inami, Dr. Yuji Kato and a former Nagoya member Dr. Kiyoshi Hayasaka helped me in many situations.

This research was achieved owing to support from many collaborators in the Belle experiment. I especially appreciate timely reviewing and detailed feedback by my internal referees Dr. Phillip Urquijo, Prof. Maria Różańska and Dr. Karim Trabelsi; assistance in situations such as the analysis, the public presentations, the journal publication and so on by Mr. Karol Adamczyk, Dr. Florian Bernlochner, Dr. Robin Glattauer, Dr. Koji Hara, Prof. Peter Križan, Prof. Youngjoon Kwon, Prof. Kenkichi Miyabayashi, Prof. Mikihiko Nakao, Dr. Shohei Nishida, Prof. Leo Piilonen, Prof. Yoshiyuki Sakai and Dr. Christoph Schwanda; helpful suggestions and corrections for the journal publication by Prof. Simon I. Eidelman, Dr. Bryan Fulsom, Dr. Akimasa Ishikawa, Prof. Yasushi Watanabe, Prof. Olivier Schneider and Dr. Makoto Takizawa. Also, Prof. Abner Soffer gave me substantial discussions for the analysis during his occasional visits to Nagoya.

I received generous support from Dr. Minoru Tanaka, Dr. Ryotaro Watanabe and Dr. Yasuhito Sakaki, who research theories in the particle physics. Their suggestions were exceedingly helpful for theoretical interpretation of my result.

In my research for the MCP-PMT development, discussions and helps with Dr. Kodai Matsuoka and all the other colleagues in the Nagoya-TOP group were essential for me to complete the MCP-PMT tests. I occasionally discussed the test results with Mr. Toshikazu Hakamata and the staff members of Electron Tube Division, Hamamatsu Photonics K.K. Their analysis and suggestion based on professional knowledge and experiences were very important to understand our results more deeply. For the MCP-PMT tests at the 1.5 T magnet in KEK, Dr. Ichiro Adachi supported me in many points for setting-up the test system, its operation and administration procedures for us to be permitted to carry out experiments at KEK North Counter Hall. Also, many kind supports from the LEPS group were essential to complete the TOP counter beam test successfully.

Dr. Makoto Tomoto and all the other N-lab members also gave me ideas from different viewpoints. Especially, I obtained a number of important ideas from Dr. Yasuyuki Horii based on his experiences for physics analysis at Belle in the beginning stage of my research.

Finally, I acknowledge our secretaries Ms. Mieko Miyake and Ms. Kaori Yonemura for much support regarding to administration and so on. Their support was significantly important for me to concentrate on and proceed my research smoothly.

This work was supported by a Grant-in-Aid for JSPS Fellows (No. 25.3096).

Appendix A

Calculation of $\cos\theta_{\rm hel}$ distribution

A.1 $\tau^- \to \pi^-\nu_\tau$

Since the pion is a pseudo-scalar meson and the neutrino is always left-handed, two helicity amplitudes in Fig. A.1 are allowed in the rest frame of τ:

$$M_+ = Ad^{1/2}_{1/2,1/2} = A\cos\frac{\theta_{\rm hel}}{2}, \tag{A.1}$$

$$M_- = Ad^{1/2}_{-1/2,1/2} = -A\sin\frac{\theta_{\rm hel}}{2}, \tag{A.2}$$

where $M_\pm$ and A denote the helicity amplitudes and their normalization constants, respectively. The function $d^S_{S_z,S'_z}$ is the Wigner's small d-matrix, where S, S_z and S'_z are the total spin, the z component of τ and the $\pi-\nu$ system, respectively. Using the τ polarization $P_\tau(D^*)$, the differential decay rate is represented by

$$\frac{d\Gamma}{d(\cos\theta_{\rm hel})} = \frac{1+P_\tau(D^*)}{2}|M_+|^2 + \frac{1-P_\tau(D^*)}{2}|M_-|^2. \tag{A.3}$$

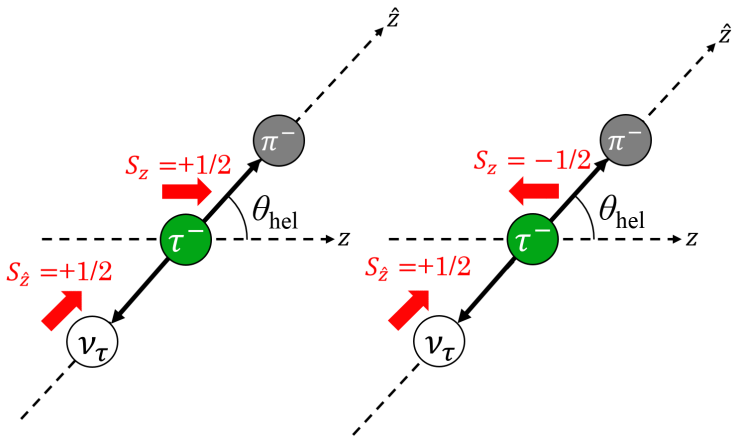


Figure A.1: Possible helicity states in the $\tau^- \to \pi^-\nu_\tau$ mode. The red arrows and the numbers show the spin directions and values, respectively.

By substituting Eqs. A.1 and A.2 for M_+ and M_-, respectively, and using $|A|^2 = \Gamma$,

$$\frac{1}{\Gamma}\frac{d\Gamma}{d(\cos\theta_{\rm hel})} = \frac{1+P_\tau(D^*)}{2}\cos^2\frac{\theta_{\rm hel}}{2} + \frac{1-P_\tau(D^*)}{2}\sin^2\frac{\theta_{\rm hel}}{2}, \quad \text{(A.4)}$$

$$= \frac{1}{4}[(1+P_\tau(D^*))(1+\cos\theta_{\rm hel}) + (1-P_\tau(D^*))(1-\cos\theta_{\rm hel})], \quad \text{(A.5)}$$

$$= \frac{1}{2}[1+P_\tau(D^*)\cos\theta_{\rm hel}]. \quad \text{(A.6)}$$

A.2 $\tau^- \to \rho^-\nu_\tau$

The difference from the $\tau^- \to \pi^-\nu_\tau$ is that the ρ meson can take two spin states: -1 and 0, as shown in Fig. A.2. The state $S = +1$ is prohibited as the neutrino must be always left-handed. The four helicity amplitudes are therefore possible:

$$(\lambda_\rho, \lambda_\tau) = (0, +1/2) \to M_+^0 = A_0 d^{1/2}_{1/2,1/2} = A_0\cos\frac{\theta_{\rm hel}}{2}, \quad \text{(A.7)}$$

$$(\lambda_\rho, \lambda_\tau) = (-1, +1/2) \to M_+^{-1} = A_{-1} d^{1/2}_{1/2,-1/2} = -A_{-1}\sin\frac{\theta_{\rm hel}}{2}, \quad \text{(A.8)}$$

$$(\lambda_\rho, \lambda_\tau) = (0, -1/2) \to M_-^0 = A_0 d^{1/2}_{-1/2,1/2} = -A_0\sin\frac{\theta_{\rm hel}}{2}, \quad \text{(A.9)}$$

$$(\lambda_\rho, \lambda_\tau) = (-1, -1/2) \to M_-^{-1} = A_{-1} d^{1/2}_{-1/2,-1/2} = A_{-1}\cos\frac{\theta_{\rm hel}}{2}, \quad \text{(A.10)}$$

where the $A_i (i = -01, 0)$ is the normalization factor. The differential decay rate is then calculated as

$$\frac{d\Gamma}{d(\cos\theta_{\rm hel})} = \frac{1+P_\tau(D^*)}{2}(|M_+^0|^2 + |M_+^{-1}|^2) + \frac{1-P_\tau(D^*)}{2}(|M_-^0|^2 + |M_-^{-1}|^2), \quad \text{(A.11)}$$

$$= \frac{1+P_\tau(D^*)}{2}\left(|A_0|^2\cos^2\frac{\theta_{\rm hel}}{2} + |A_{-1}|^2\sin^2\frac{\theta_{\rm hel}}{2}\right) + \frac{1-P_\tau(D^*)}{2}\left(|A_0|^2\sin^2\frac{\theta_{\rm hel}}{2} + |A_{-1}|^2\cos^2\frac{\theta_{\rm hel}}{2}\right), \quad \text{(A.12)}$$

$$= \frac{1}{4}(1+P_\tau(D^*))(|A_0|^2 + |A_0|^2\cos\theta_{\rm hel} + |A_{-1}|^2 - |A_{-1}|^2\cos\theta_{\rm hel}) + \frac{1}{4}(1-P_\tau(D^*))(|A_0|^2 - |A_0|^2\cos\theta_{\rm hel} + |A_{-1}|^2 + |A_{-1}|^2\cos\theta_{\rm hel}), \quad \text{(A.13)}$$

$$= \frac{1}{2}(|A_0|^2 + |A_{-1}|^2 + P_\tau(D^*)|A_0|^2\cos\theta_{\rm hel} - P_\tau(D^*)|A_{-1}|^2\cos\theta_{\rm hel}). \quad \text{(A.14)}$$

The total decay rate Γ is equal to $|A_0|^2 + |A_{-1}|^2$, and then

$$\frac{1}{\Gamma}\frac{d\Gamma}{d(\cos\theta_{\rm hel})} = \frac{1}{2}\left[1 + P_\tau(D^*)\frac{|A_0|^2 - |A_{-1}|^2}{|A_0|^2 + |A_{-1}|^2}\cos\theta_{\rm hel}\right]. \quad \text{(A.15)}$$

Using the relation $|A_0|/|A_{-1}| = \sqrt{2}m_\rho/m_\tau$ [62],

$$\frac{1}{\Gamma}\frac{d\Gamma}{d(\cos\theta_{\rm hel})} = \frac{1}{2}[1 + \alpha P_\tau(D^*)\cos\theta_{\rm hel}], \quad \text{(A.16)}$$

$$\alpha = \frac{m_\tau^2 - 2m_\rho^2}{m_\tau^2 + 2m_\rho^2}. \quad \text{(A.17)}$$

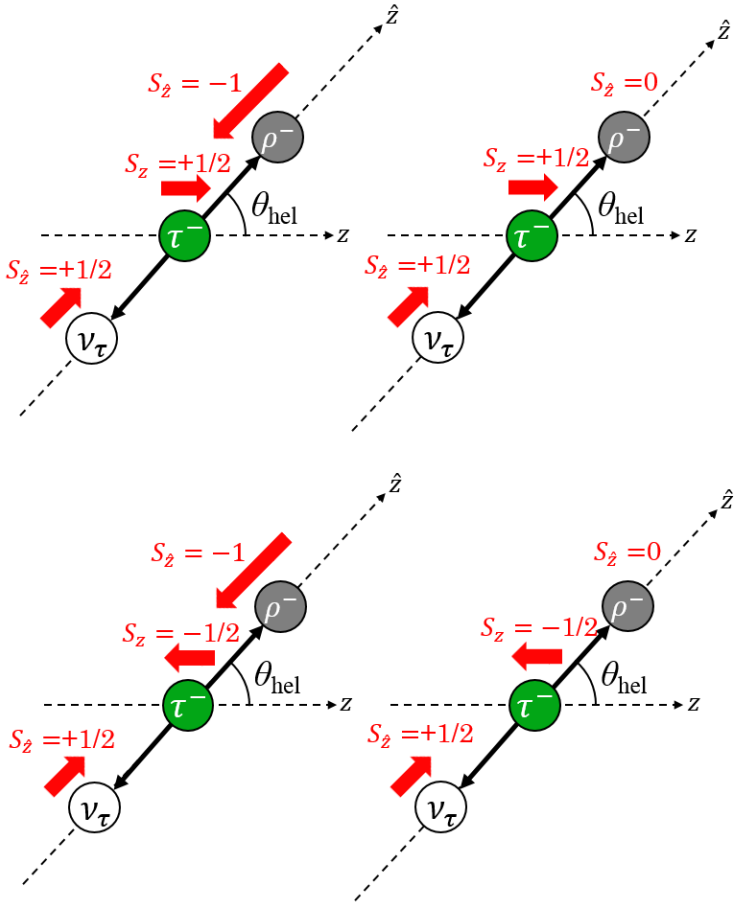


Figure A.2: Possible helicity states in the $\tau^- \to \rho^- \nu_\tau$ mode. The red arrows and the numbers show the spin directions and values, respectively.

(This discussion applies to any two-body τ decays with a vector meson.)

Appendix B

Semileptonic Decay Model Correction

Since the updated FF parameters for $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$ and the improved decay model for $\bar{B} \to D^{**} \ell^- \bar{\nu}_\ell$ are available, we correct the MC samples by reweighting the decay kinematics.

Table B.1 is the comparison of the FF parameters used for $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$ MC production and the updated parameters. These parameters are usually discussed using two variables: the cosine of the angle between the ℓ^- momentum with respect to the direction opposite to the $\bar{B}$ momentum in the rest frame of W^* ($\cos\theta_\ell$) and w defined in Eq. 1.23. We calculate the weight factors using corresponding two variables: q^2 and p^*_ℓ, where p^*_ℓ is the charged lepton momentum in the rest frame of $\bar{B}$. Figure B.1 shows the comparison of the q^2 and the p^*_ℓ distributions with the old and the updated FF parameters. We obtain

Table B.1: HQET FF parameters used for the $\bar{B} \to D^* \ell^- \bar{\nu}_\ell$ MC production and the updated parameters in Eqs. 1.31 to 1.33.

Parameter	Old value	Updated value
ρ^2	1.3	1.207 ± 0.026
R_1	1.18	1.406 ± 0.033
R_2	0.71	0.853 ± 0.020

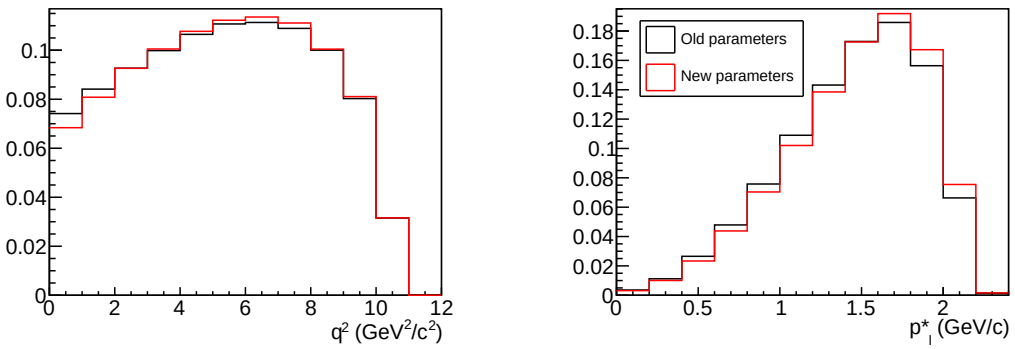


Figure B.1: Comparison of the q^2 and the p^*_ℓ distributions with the old and the new FF parameters.

the weight factors as ratios of the kinematic distributions.

For the decays $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$, the MC sample has been generated by the ISGW model [67], and it is corrected to match the LLSW FFs [68]. In the case of the $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ correction, we directly use the distributions of $\cos\theta_\ell$ and w to calculate the weight factors.

The weight factors have errors arising from the uncertainties in the FF parameters. These uncertainties are taken into account as the systematic uncertainty in Secs. 5.4 and 5.5.

Appendix C

Constraint for τ Momentum

The decay chain $\bar{B} \to D^* \tau^- \bar{\nu}_\tau, \tau^- \to \pi^- \nu_\tau$ is divided into three processes: (1) $\bar{B} \to W^{*-} D^*$, (2) $W^{*-} \to \tau^- \bar{\nu}_\tau$ and (3) $\tau^- \to \pi^- \nu_\tau$. The process (1) has $4 \times 3 = 12$ degrees of freedom (d.o.f.) in total. Since we can fully reconstruct the four-momentum of D^* and obtain the four-momentum of $\bar{B}$ using the hadronic tag, eight d.o.f are determined. Additional four d.o.f are constrained by the four-momentum conservation. All the 12 d.o.f. are therefore constrained.

Next, the right side of the process (2) has $4 \times 2 = 8$ d.o.f and the W^{*-} momentum is fully determined by the process (1). Out of the eight d.o.f, four are constrained by the four-momentum conservation and two by the masses of τ^- and $\bar{\nu}_\tau$. Because we cannot detect either of τ^- or $\bar{\nu}_\tau$, $8-4-2 = 2$ d.o.f. remain. The final process (3) has two d.o.f in the left side and $4 \times 2 = 8$ d.o.f in the right side. Out of the 10 d.o.f, four are constrained by the momentum conservation and the other four are determined by detecting π^-. One more d.o.f is fixed by the neutrino mass. However, we do not have constraints on the three-momentum of ν_τ. Therefore $10 - 4 - 4 - 1 = 1$ d.o.f. cannot be constrained.

Appendix D

Event Selection Optimization

The signal selection requirements are optimized as below.

If we reconstruct events without any event selection, the SNR is extremely low. We first adopt the rough event selection as follows.

- $O_{\rm NB} > 0.01$, $M_{\rm bc} > 5.27$ GeV/c^2 and $|\Delta E| < 50$ MeV for $B_{\rm tag}$.
- $|M_D - m_D| < 3\sigma_D$.
- $|\Delta M - \Delta m| < 3\sigma_{D^*}$.
- $|M_{\pi\pi^0} - m_\rho| < 150$ MeV/c^2 for ρ meson candidates, where $m_\rho = 775.3$ MeV/c^2 is the ρ meson reported in Ref. [10].
- $q^2 > 4$ GeV$^2/c^2$.
- $p_d^* > 0.5(0)$ GeV/c for $\tau^- \to \pi^-\nu_\tau(\rho^-\nu_\tau)$, where p_d^* is the three-momentum of the τ-daughter meson (π or ρ) in the CM frame.
- $0.5 < M^2_{\rm miss} < 7$ GeV$^2/c^4$.
- $|\cos\theta_{\rm hel}| < 1$.

In the signal selection optimization, we need to avoid distortion of the $\cos\theta_{\rm hel}$ acceptance. Therefore,

- we do not apply any requirement on $M^2_{\rm miss}$ and p_d^*, which have strong correlation to $\cos\theta_{\rm hel}$.
- To avoid change of the selection efficiency on q^2 due to modification of the q^2 shape by NP, we adopt only the minimum requirement on q^2 ($q^2 > 4$ GeV$^2/c^2$).
- In the $\tau^- \to \pi^-\nu_\tau$ sample, the $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ background makes a significant peak at $\cos\theta_{\rm hel} \sim 1$. We reject events with $\cos\theta_{\rm hel} > 0.8$ for this sample.

Eventually, we optimize the selection requirements for candidate selection of ρ, D, D^* and $B_{\rm tag}$. The selection conditions are then optimized based on FOM in Eq. 3.13.

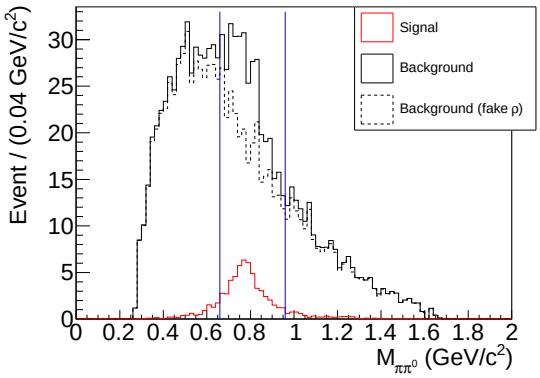


Figure D.1: Distribution of $M_{\pi\pi^0}$. The red and the black histograms are the signal and the background distributions. In the background histogram, the component above (below) the dashed line shows correctly-reconstructed ρ candidates (candidates from random combination of $\pi^\pm$ and π^0). The region within the blue lines shows the selected $m_{\pi\pi^0}$ region.

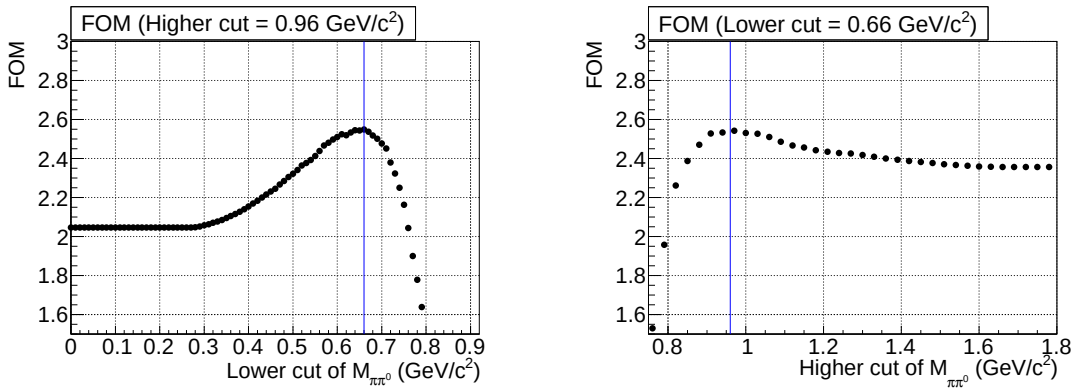


Figure D.2: FOM as a function of the lower (upper) requirement on $M_{\pi\pi^0}$ for the left (right) panel. Here, the upper (lower) requirement is fixed at 960 MeV/c^2 (660 MeV/c^2). The blue lines show the selected signal region.

D.1 ρ Selection

Figure D.1 shows the $M_{\pi\pi^0}$ distribution. The correctly-reconstructed ρ candidates in the background mainly originate from the decay $\bar{B} \to D^* \tau^- \bar{\nu}_\tau$ with a ρ meson in the different τ decays from $\tau^- \to \rho^- \bar{\nu}_\tau$ and the decay $\bar{B} \to D^* D_s^{*-}$ with a ρ meson in the D_s decay. Figure D.2 shows the FOM as a function of the upper and lower $M_{\pi\pi^0}$ requirements. The defined requirement 660 to 960 MeV/c^2 is indicated in the figure.

D.2 $D^{(*)}$ Selection

Figure D.3 shows the FOM for M_D as a function of the width of the signal regions. The defined signal regions are indicated.

Since, as shown in Fig. 3.7, the SNR of the ΔM distribution is much different among the four D^* decay, we optimize the ΔM signal regions separately.

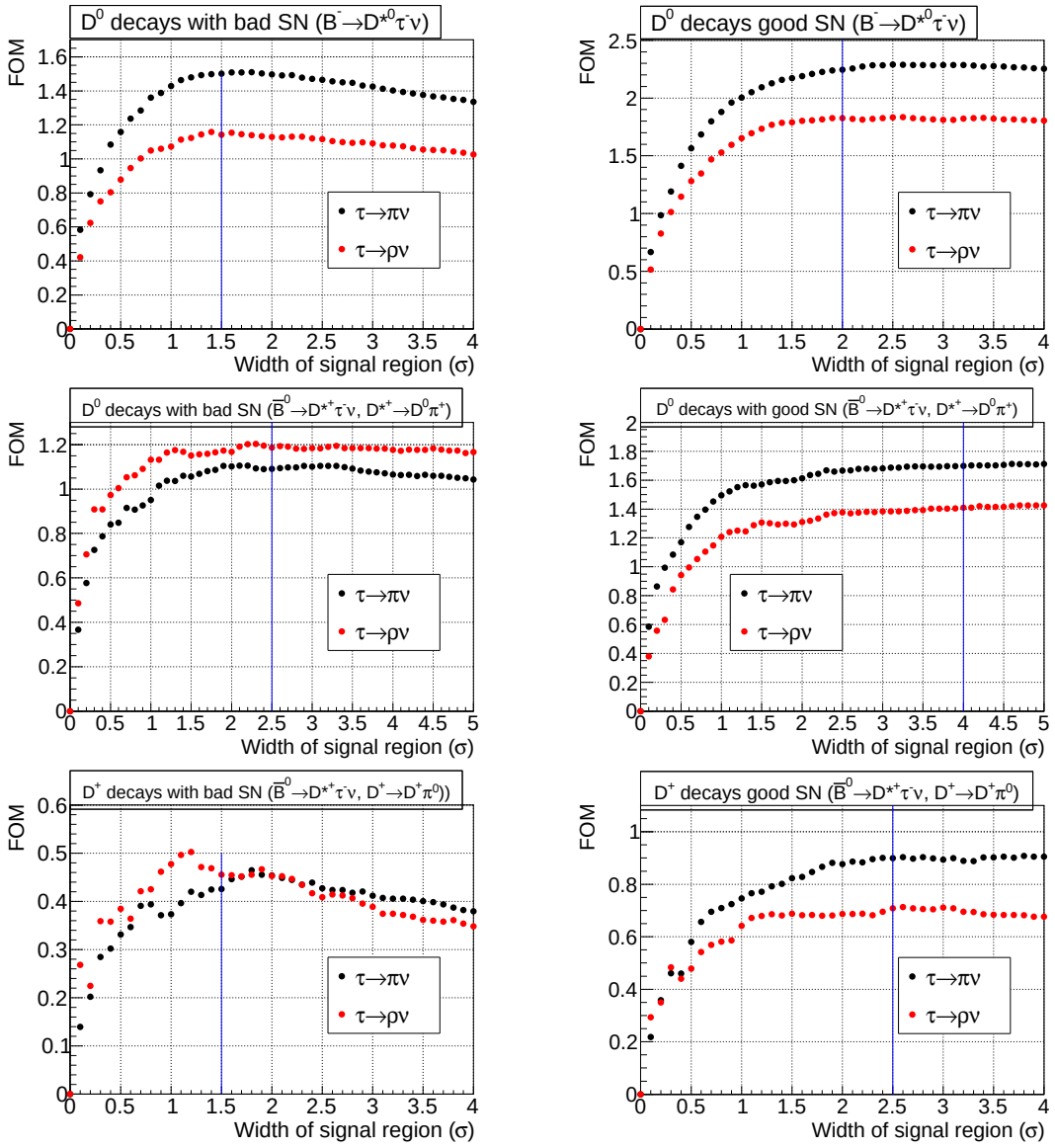


Figure D.3: FOM as a function of the width of the M_D signal region. The left panel includes the low-SNR modes while the right is the high-SNR modes. The blue lines show the selected signal region.

For $D^{*+} \to D^+\pi^0$ and $D^{*0} \to D^0\pi^0$, we use the energy-asymmetry of a photon pair of the slow π^0 in the laboratory frame. Figure D.4 (left) shows comparison of the A_γ distributions between correctly-reconstructed and fake D^* candidates ($D^{*+} \to D^+\pi^0$ and $D^{*0} \to D^0\pi^0$ are combined). The value of A_γ tends to be lower for the correctly-reconstructed D^* candidates. Figure D.4 (right) shows the FOM as a function of the $A_/gamma$ requirement. Based on this result, we select candidates satisfying $A_\gamma < 0.6$.

Figure D.5 shows the FOM as a function of the ΔM requirement. The defined requirement are indicated.

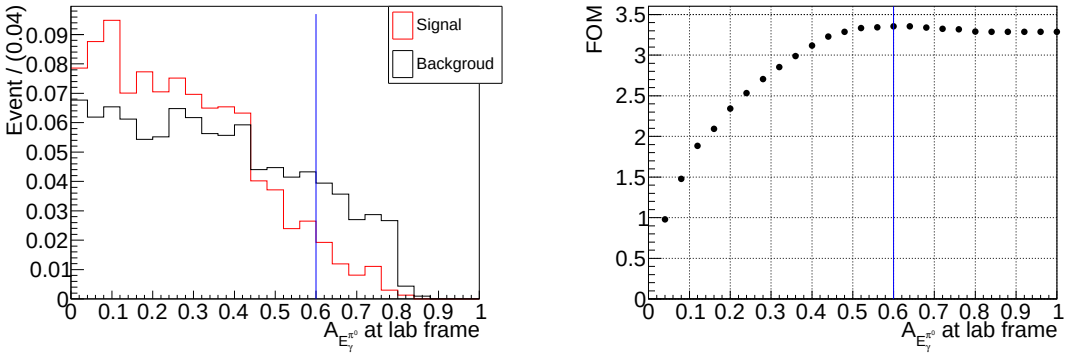


Figure D.4: (left) Comparison of the A_γ distribution. The red and black histogram are signals and backgrounds, respectively. Both histograms are normalized to be unity. (right) FOM as a function of the threshold of A_γ. The blue lines show the selected threshold.

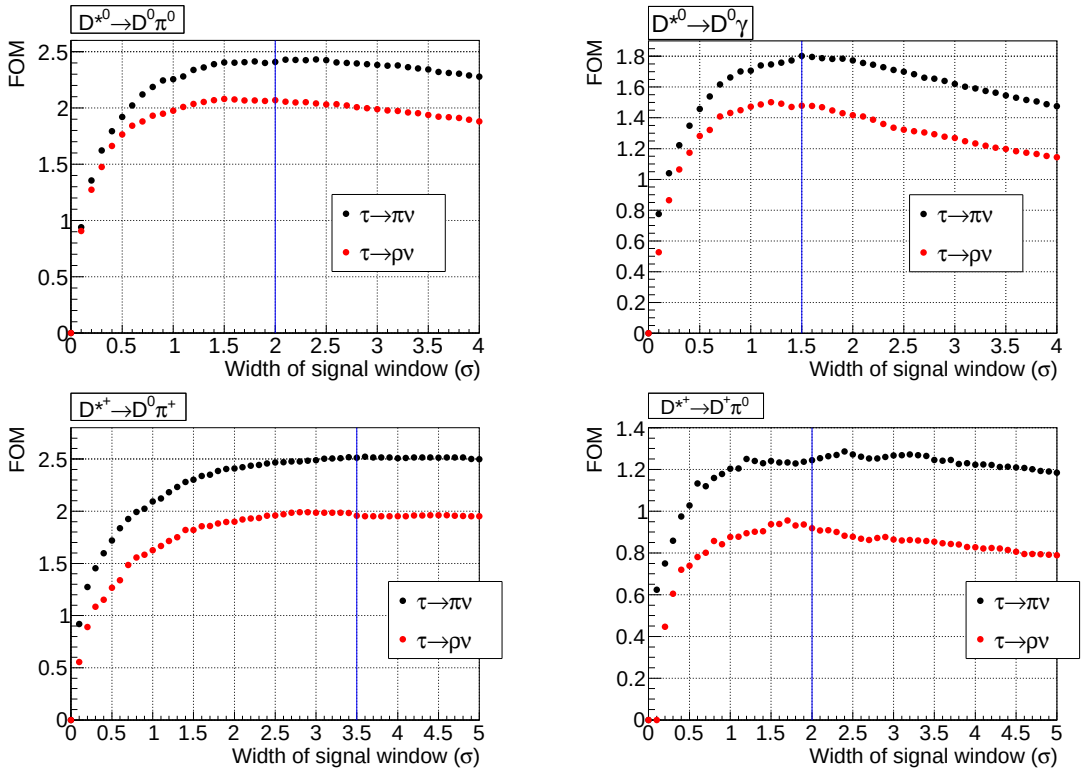


Figure D.5: FOM as a function of the width of the ΔM signal region. The blue lines show the selected signal region.

D.3 $B_{\rm tag}$ Selection

Figures D.6 and D.7 show the $O_{\rm NB}$ ($M_{\rm bc}$) distribution for the signal and the background events and the corresponding FOM as a function of the $O_{\rm NB}$ ($M_{\rm bc}$) requirement. The defined requirements are indicated.

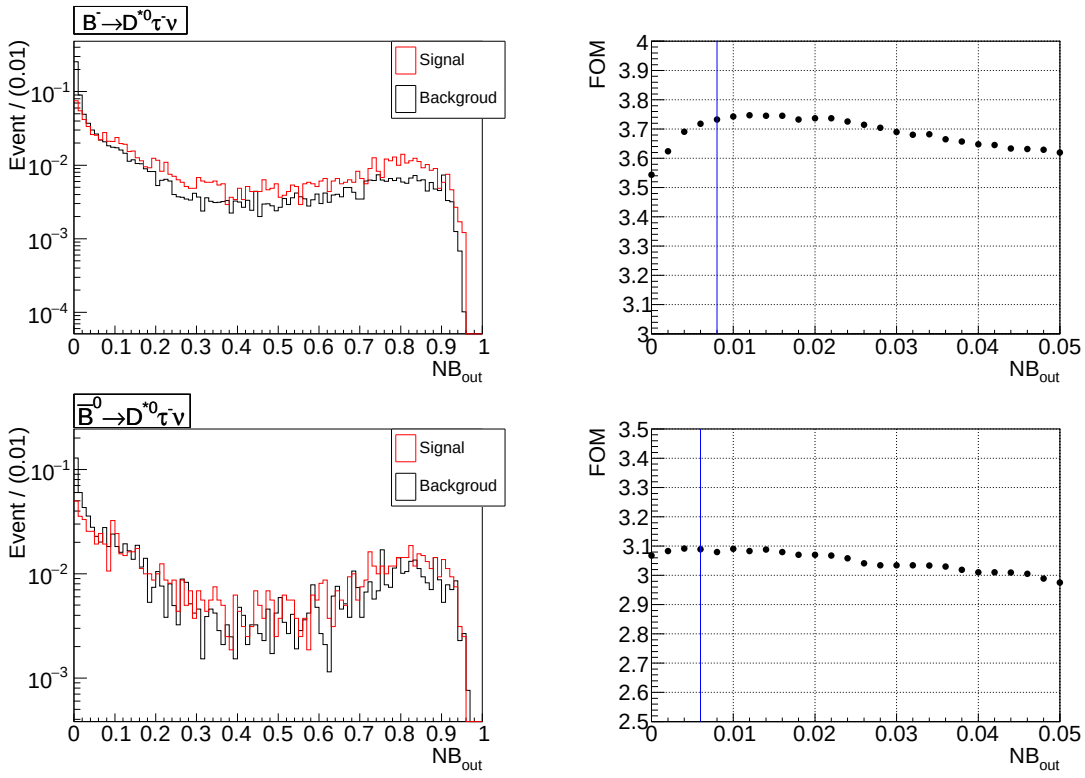


Figure D.6: (left) Distribution of O_{NB} for the signal events (red) and the background events (black). (right) FOM as a function of the O_{NB} threshold. The blue lines show the selected threshold.

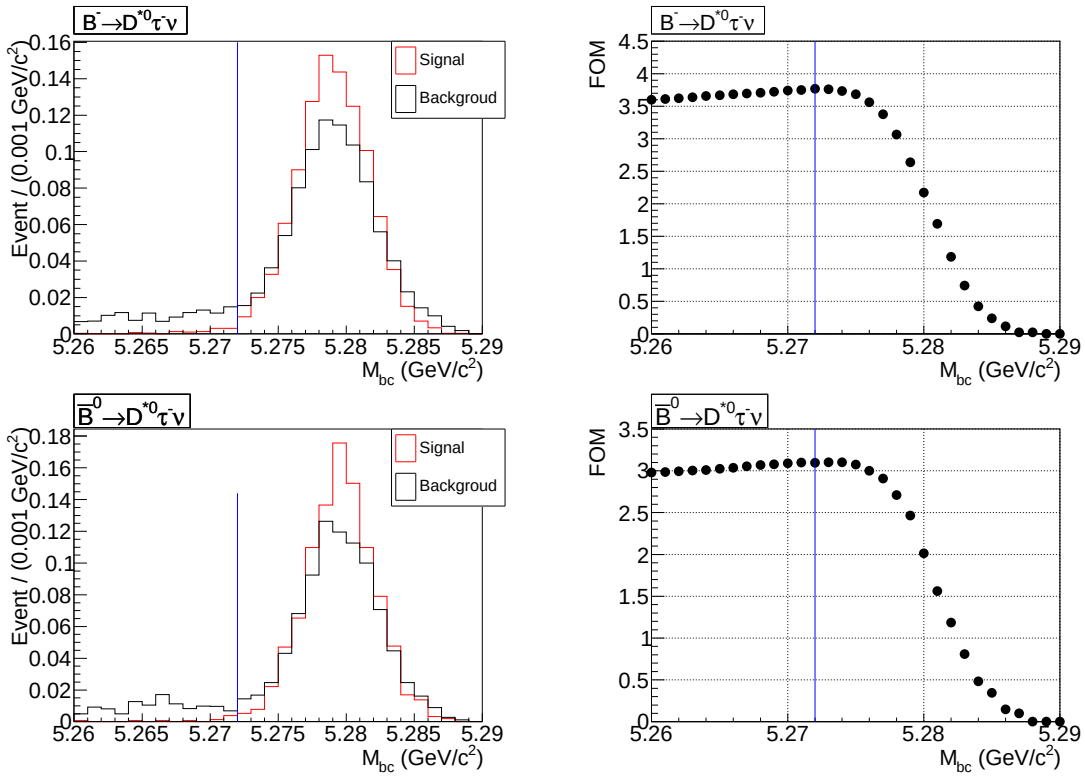


Figure D.7: (left) Distribution of $M_{\rm bc}$ for the signal events (red) and the background events (black). (right) FOM as a functions of the $M_{\rm bc}$ threshold. The blue lines show the selected threshold.

Appendix E

Measurement of the Hadronic B Yields

As discussed in Sec. 3.5.4, the composition of the multi-π/η mode is calibrated using the calibration data sample. The yield is extracted using $M_{\rm bc}^{\rm sig}$, which has a narrow peak around the B meson mass, with the requirement $|\Delta E^{\rm sig}| < 100$ MeV. However, the samples of $\bar{B} \to D^*\pi^-\pi^-\pi^+$, $D^*\pi^-\pi^-\pi^+\pi^0$ and $D^*\pi^-\pi^0\pi^0$ show a peaking-background component. Figure E.1 shows the $M_{\rm bc}^{\rm sig}$ distribution of the $\bar{B} \to D^*\pi^-\pi^-\pi^+\pi^0$ sample in the MC simulation. For such samples, we use $\Delta E^{\rm sig}$ and require $M_{\rm bc}^{\rm sig} > 5.27$ GeV/c^2 for the yield extraction.

A model function is constructed as below.

$M_{\rm bc}^{\rm sig}$
: The signal component is modeled by the Crystal Ball function, and the background component by another empirically-parameterized function by the Argus Collaboration [100]. Since the $M_{\rm bc}^{\rm sig}$ resolution is mostly determined by the beam energy resolution and independent of the B decay mode, the resolution parameter in the Crystal Ball function is fixed from the tag-side $M_{\rm bc}$ distribution.

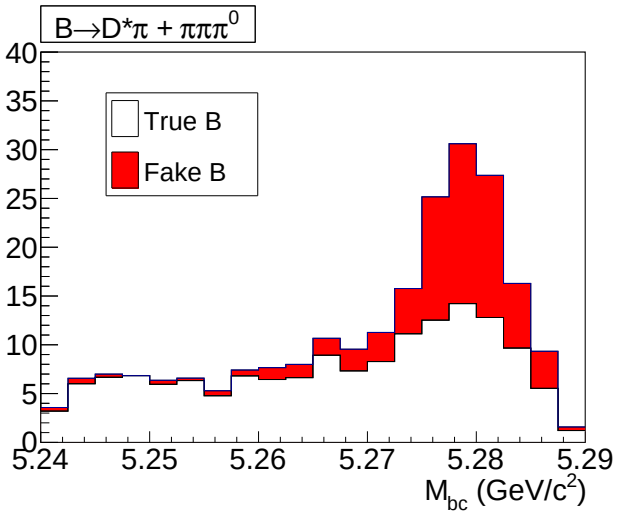


Figure E.1: Distribution of $M_{\rm bc}^{\rm sig}$ for the $B^- \to D^*\pi^-\pi^-\pi^+\pi^0$ sample in the MC simulation.

$\Delta E^{\rm sig}$ The signal component is modeled by the Crystal Ball function, and the background component by a quadratic function.

The fit results to all the seven calibration samples for the B^- and the $\bar{B}^0$ modes are shown in Figs. E.2 and E.3.

- We take a ratio of the yield in the data to that in the MC sample. The ratio is used as a calibration factor for the hadronic B background composition. The uncertainty in the calibration factor is estimated based on the statistical errors of the yields.

- The sample of $B^- \to D^*\pi^-\pi^-\pi^+$ shows a negative yield in the data, while several events are found in the MC sample. We therefore assume that there is no event in the data and assign a 68% C.L. upper limit ($C_{68\%}$) on the yield of the data. The upper limit is treated as a statistical uncertainty of the yield; namely, the yield is considered as $0.0^{+C_{68\%}}_{-0.0}$.

- In the sample of $\bar{B}^0 \to D^*\pi^-\pi^0$, no event is observed in the data, while several events are expected from the MC simulation. According to the MC simulation, this mode is almost free from the background. We therefore assume that there is no background and assign a 68% C.L. upper limit in the yield based on a Poisson distribution. It is estimated to be 1.1 events. The upper limit is treated as a statistical uncertainty; namely, the yield is considered as $0.0^{+1.1}_{-0.0}$.

- In the sample of $\bar{B}^0 \to D^*\pi^-\eta$, only two events are observed in the data. This number of events is less than the number of free parameters in the model function and therefore not sufficient to perform a fit. According to the MC simulation, this mode is almost free from the background. We therefore assume that these two events, existing around the nominal B meson mass, are the signal events and assign a 68% C.L. upper limit in the yield based on a Poisson distribution. The upper limit is estimated to be 3.5 events. The positive statistical uncertainty is therefore estimated to be $3.5 - 2.0 = 1.5$. For the negative uncertainty, considering that the minimum value of the yield is zero, we assign 2.0. The yield is eventually obtained as $2.0^{+1.5}_{-2.0}$ events.

In Table 3.15, in addition to the $\bar{B} \to D^*\pi^-\eta$, a few modes show significantly smaller yield calibration factors than 1. This is because, in the MC simulation, branching fractions for some of the non-resonant multi-pion modes such as $\bar{B} \to D^*\pi^-\pi^-\pi^+$ are overestimated compared to the values in Ref. [10]. These overestimation is corrected by our calibration.

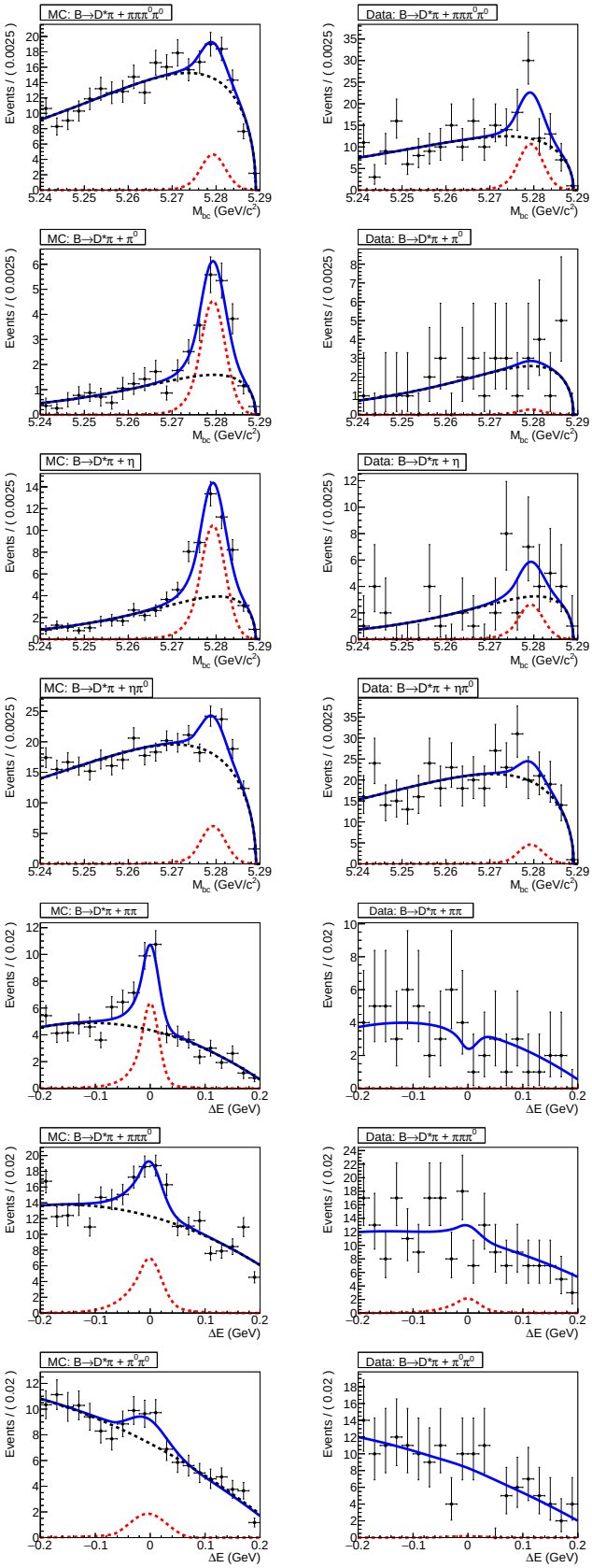


Figure E.2: Fit results to the signal-side $M_{\rm bc}^{\rm sig}$ or $\Delta E^{\rm sig}$ distributions of the seven control samples for the charged B sample. The left and the right panels show distributions of the MC and the data, respectively. The black points show data distribution, and the solid blue, dashed red and the dashed black lines show the sum of the fitted function, signal component and the background component, respectively. The numbers shown by the red texts are the obtained yield ratios.

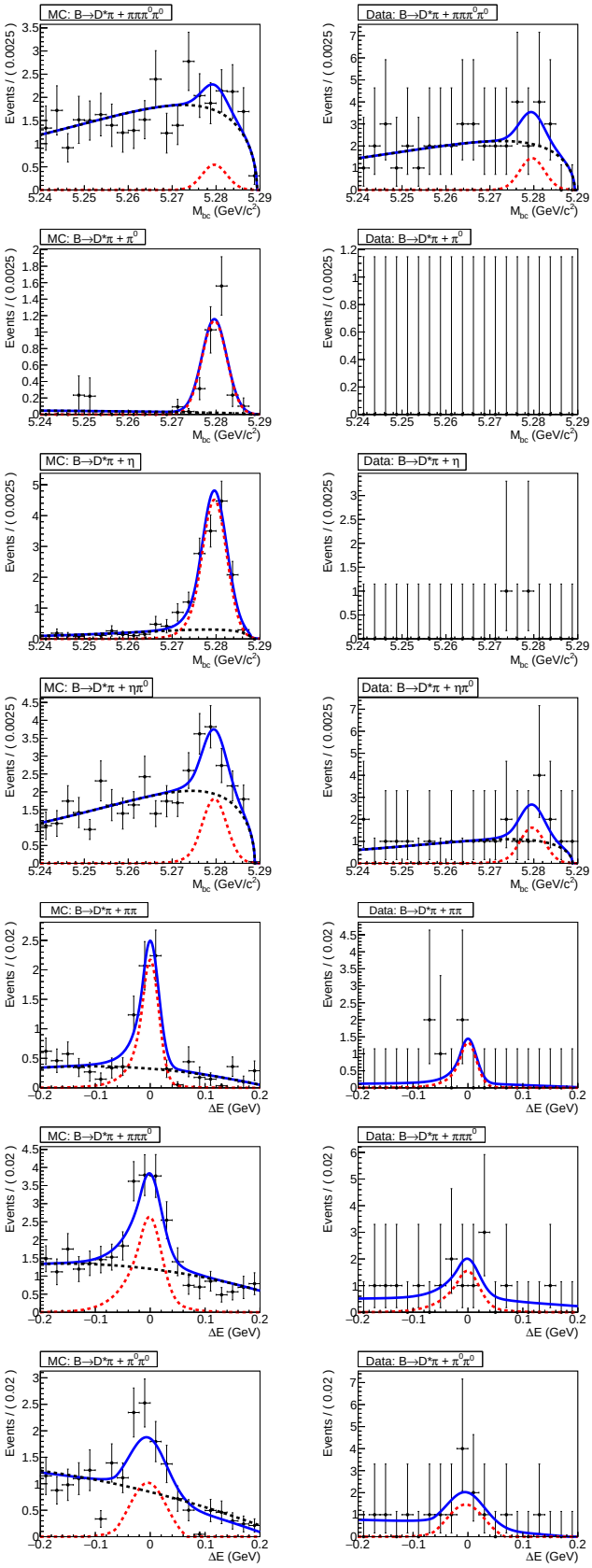


Figure E.3: Fit results to the signal-side $M_{\rm bc}^{\rm sig}$ or $\Delta E^{\rm sig}$ distributions of the seven control samples for the neutral B sample. The meaning of each colored line, points and texts is the same as Fig. E.2.

Appendix F

E_{ECL} Comparison for the Fake D^* Component

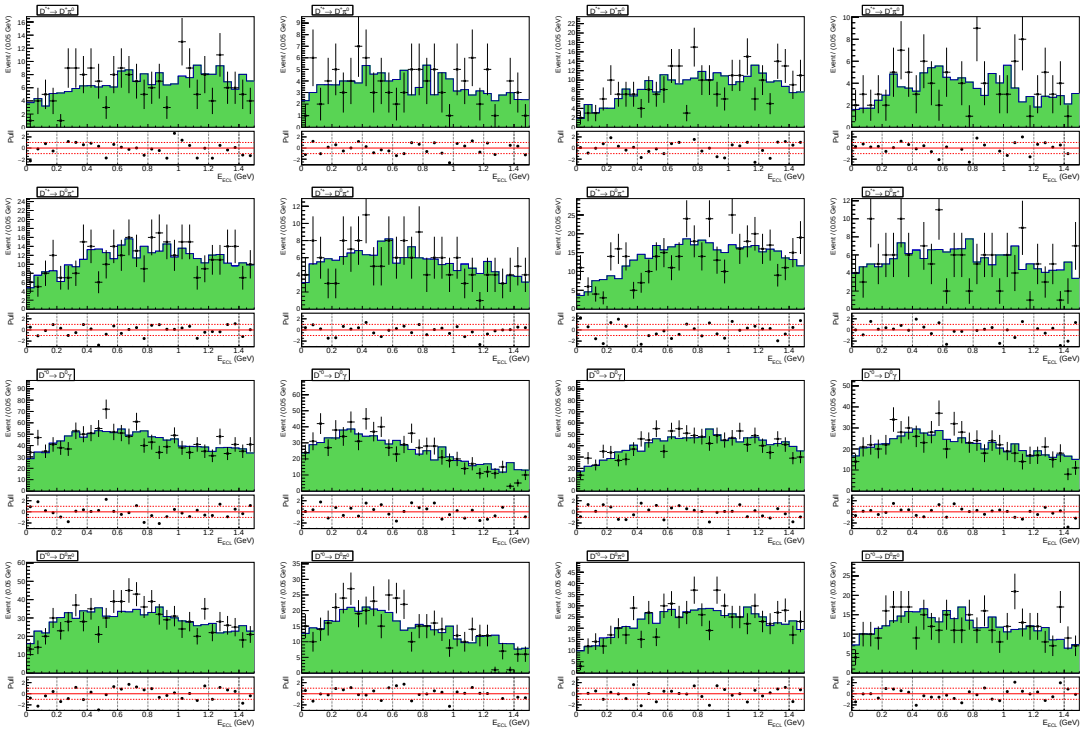


Figure F.1: Comparison of the E_{ECL} distribution between the data and the MC simulation in the Δm sideband region. For every 16 plot, the solid green and blue histograms show the fake D^* component and the other component, respectively, in the MC simulation. The black dots show the data distribution. The bottom panel indicate the E_{ECL} distribution between the data and the MC simulation.

For the E_{ECL} comparison in Sec. 3.5.3, the figure in each sample is shown in Fig. F.1.

Appendix G

Mode-by-mode Efficiency Variation

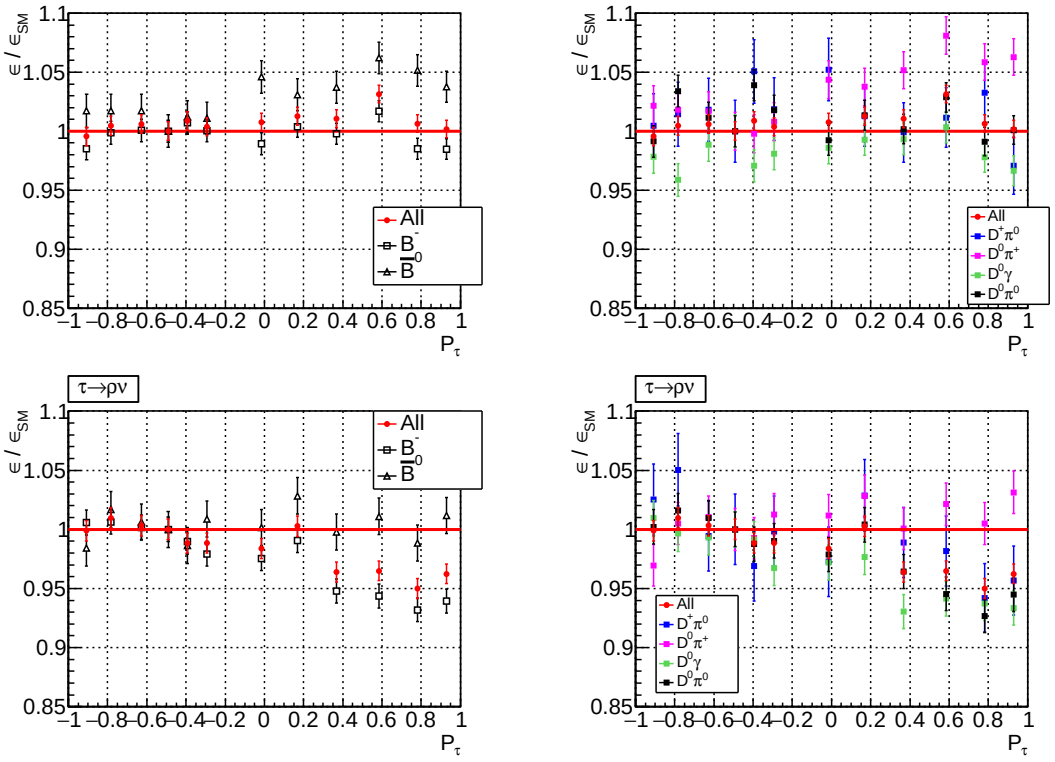


Figure G.1: Reconstruction efficiency for the signal component of the $\tau^- \to \pi^- \nu_\tau$ mode (top) and the $\tau^- \to \rho^- \nu_\tau$ mode (bottom), relative to that in the SM case as a function of $P_\tau(D^*)$.

Figure G.1 is the comparison of the relative efficiency for each B and D^* sub-sample. At the higher $P_\tau(D^*)$ region, the $D^{*+} \to D^0\pi^+$ sample shows a significantly different efficiency. We may need to employ the different efficiency correction if the measured $P_\tau(D^*)$ is larger than 0.2.

Figures G.2 and G.3 are the comparison of the $P_\tau(D^*)$ correction function for each B and D^* sub-sample. The slope and the intercept of the $P_\tau(D^*)$ correction functions are compared in Figs. G.4 and G.5. As we do not observe any significant difference, we use the common correction functions for all the samples.

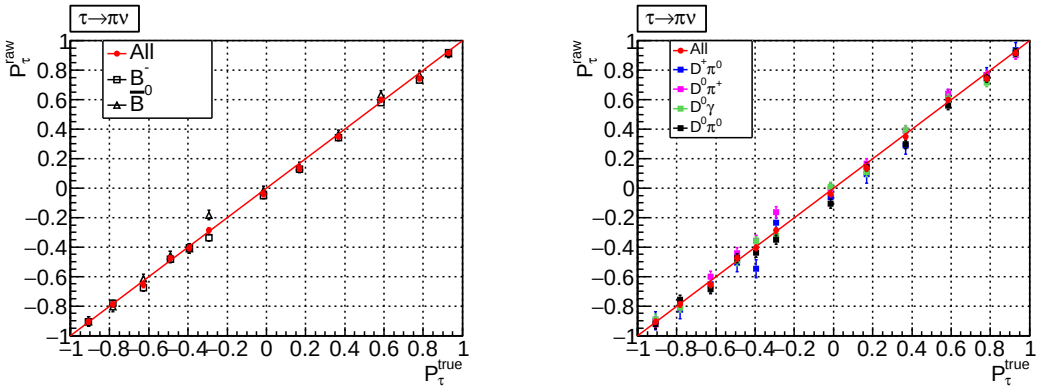


Figure G.2: $P_\tau(D^*)$ correction functions for the signal components of the $\tau^- \to \pi^-\nu_\tau$ sample. Samples of the different B charges (left) or the different D^* decay (right) are shown.

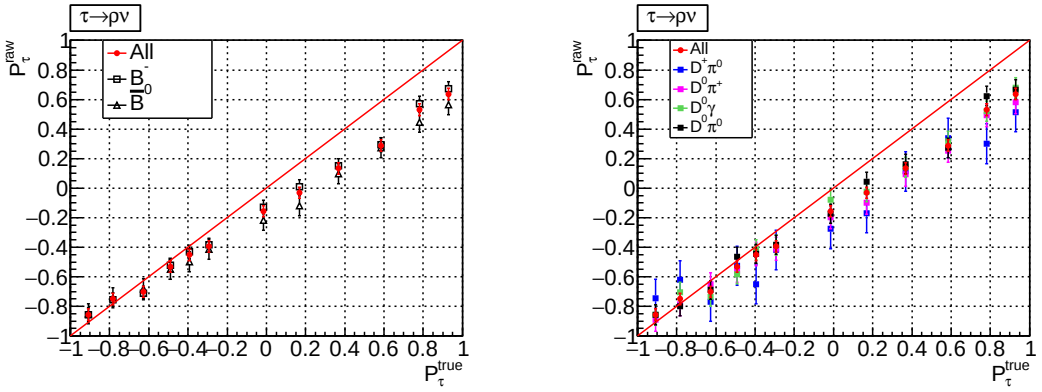


Figure G.3: $P_\tau(D^*)$ correction functions for the signal components of the $\tau^- \to \rho^-\nu_\tau$ sample. Meaning of each marker is the same as Fig. G.2.

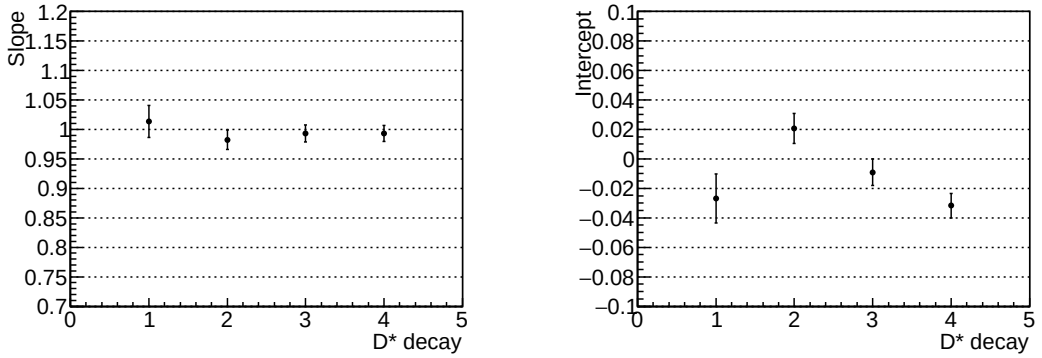


Figure G.4: Comparison of the slope (left) and the intercept (right) of the correlation functions for $\tau^- \to \pi^-\nu_\tau$ sample with each D^* decay. The D^* decay axes mean: $1 = D^{*+} \to D^+\pi^0$, $2 = D^{*+} \to D^0\pi^+$, $3 = D^{*0} \to D^0\gamma$ and $4 = D^{*0} \to D^0\pi^0$.

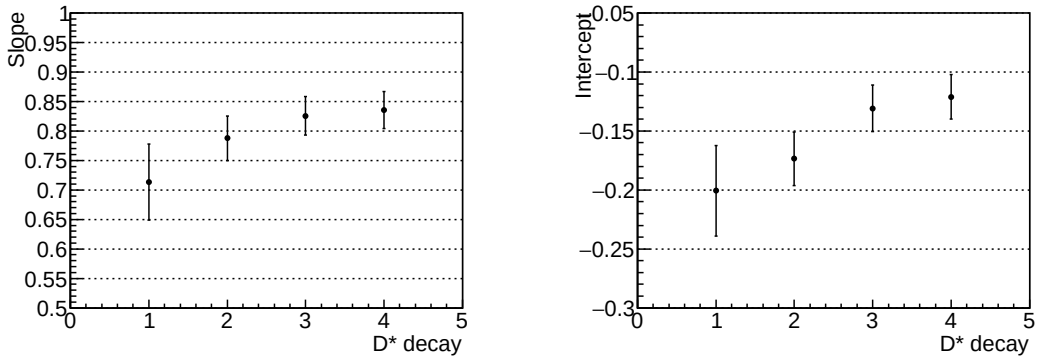


Figure G.5: Comparison of the slope (left) and the intercept (right) of the correlation functions for the $\tau^- \to \rho^- \nu_\tau$ sample. The meaning of the D^* decay axes is described in Fig. G.4.

Bibliography

[1] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons* Phys. Rev. Lett. **13**, 321 (1964).

[2] P.W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13**, 508 (1964).

[3] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716**, 1 (2012).

[4] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716**, 30 (2012).

[5] L. Evans and P. Bryant, *LHC Machine*, JINST **3**, S08001 (2008).

[6] Y. Fukuda *et al.* (Super-Kamiokande Collaboration), *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. **81**, 1562 (1998).

[7] Q.R. Ahmad *et al.* (SNO Collaboration), *Measurement of the Rate of* $\nu_e + d \rightarrow p + p + e^-$ *Interactions Produced by* ^8B *Solar Neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87**, 071301 (2001).

[8] T.S. van Albada, J.N. Bahcall, K. Begeman and R. Sancisi, *Distribution of dark matter in the spiral galaxy NGC 3198*, Astro. J. **295**, 305 (1985).

[9] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49**, 652 (1973).

[10] C. Patrignani *et al.* (Particle Data Group), *Review of Particle Physics*, Chin. Phys. C **40**, 100001 (2016).

[11] M. Tanaka and R. Watanabe, *New physics in the weak interaction of* $\bar{B} \rightarrow D^{(*)}\tau\bar{\nu}$, Phys. Rev. D **87**, 034028 (2013).

[12] K. Hagiwara, A.D. Martin and M.F. Wade, *Exclusive semileptonic B-meson decays*, Nucl. Phys. B **327**, 569 (1989).

[13] S. Fajfer, J.F. Kamenik and I. Nišandžić, $B \rightarrow D^*\tau\bar{\nu}$ *sensitivity to new physics*, Phys. Rev. D **85**, 094025 (2012).

[14] I. Caprini, L. Lellouch and M. Neubert, *Dispersive bounds on the shape of* $\bar{B} \rightarrow D^{(*)}\ell\bar{\nu}$ *form factors*, Nucl. Phys. B **530**, 153 (1998).

[15] J.A. Bailey *et al.* (Fermilab Lattice and MILC Collaborations), *Update of $|V_{cb}|$ from the $\bar{B} \to D^* \ell \bar{\nu}$ form factor at zero recoil with three-flavor lattice QCD*, Phys. Rev. D **89**, 114504 (2014).

[16] M. Antonelli *et al.*, *Flavor physics in the quark sector*, Phys. Rep. **494**, 197 (2010).

[17] Y. Amhis *et al.* (Heavy Flavor Averaging Group), *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, arXiv:1412.7515 (2014) and online update at http://www.slac.stanford.edu/xorg/hfag/.

[18] J.F. Gunion, H.E. Haber, G.L. Kane and S. Dawson, *The Higgs Hunter's Guide*, Front. Phys. **80**, 1 (2000).

[19] Reviewed by S.P. Martin, *A supersymmetry Primer*, arXiv:hep-ph/9709356v7 (2016).

[20] For example, G. Bertone, D. Hooper and J. Silk, *Particle Dark Matter: Evidence, Candidates and Constraints*, Phys. Rept. **405**, 279 (2005).

[21] ATLAS Collaboration, *Search for charged Higgs bosons in the τ + jets final state using 14.7 fb^{-1} of pp collision data recorded at $\sqrt{s}$ = 13 TeV with the ATLAS experiment*, ATLAS-CONF-2016-088 (2016).

[22] CMS Collaboration, *Search for charged Higgs bosons with the $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay channel in the fully hadronic final state at $\sqrt{s}$ = 13 TeV*, CMS-PAS-HIG-16-031 (2016).

[23] M. Tanaka, *Charged Higgs effects on exclusive semi-tauonic B decays*, Z. Phys. C **67**, 321 (1995).

[24] K. Kiers and A. Soni, *Improving constraints on $tan\beta/m_H$ using $B \to D\tau\bar{\nu}$*, Phys. Rev. D **56**, 5786 (1997).

[25] H. Itoh, S. Komine and Y. Okada, *Tauonic B decays in the Minimal Supersymmetric Standard Model*, Prog. Theor. Phys. **114**, 179 (2005).

[26] M. Tanaka and R. Watanabe, *Tau longitudinal polarization in $\bar{B} \to D\tau\bar{\nu}$ and its role in the search for the charged Higgs boson*, Phys. Rev. D **82**, 034027 (2010).

[27] ALEPH, DELPHI, L3 and OPAL Collaborations and The LEP working group for Higgs boson searches, *Search for Charged Higgs bosons: Combined Results Using LEP Data*, CERN-PH-EP-2012-369 (2013).

[28] M. Misiak *et al.*, *Updated Next-to-Next-to-Leading-Order QCD Predictions for the Weak Radiative B-Meson Decays*, Phys. Rev. Lett. **114**, 221801 (2015).

[29] T. Enomoto and R. Watanabe, *Flavor constraints on the Two Higgs Doublet Models of Z_2 symmetric and aligned types*, J. High Energy Phys. **2016**, 2 (2016).

[30] B. Schrempp and F. Schrempp, *Light leptoquarks*, Phys. Lett. B **153**, 101 (1985).

[31] P. Biancofiore, P. Colangelo and F. De Fazio, *Anomalous enhancement observed in $B \to D^{(*)}\tau\bar{\nu}_{\tau}$ decays*, Phys. Rev. D **87**, 074010 (2013).

[32] I. Doršner, S. Fajfer, N. Košnik and I. Nišandžić, *Minimally flavored colored scalar in $\bar{B} \to D^{(*)}\tau\bar{\nu}$ and the mass matrices constraints*, J. High Energy Phys. **1311**, 084 (2013).

[33] Y. Sakaki, R. Watanabe, M. Tanaka and A. Tayduganov, *Testing leptoquark models in $\bar{B} \to D^{(*)}\tau\bar{\nu}$*, Phys. Rev. D **88**, 094012 (2013).

[34] M. Duraisamy, P. Sharma and A. Datta, *Azimuthal $B \to D^*\tau^-\bar{\nu}_\tau$ angular distribution with tensor operators*, Phys. Rev. D **90**, 074013 (2014).

[35] Y. Sakaki, M. Tanaka, A. Tayduganov and R. Watanabe, *Probing new physics with q^2 distributions in $\bar{B} \to D^{(*)}\tau\bar{\nu}$*, Phys. Rev. D **91**, 114028 (2015).

[36] M. Freytsis, Z. Ligeti and J.T. Ruderman, *Flavor models for $\bar{B} \to D^{(*)}\tau\bar{\nu}$*, Phys. Rev. D **92**, 054018 (2015).

[37] CMS Collaboration, *Search for the third-generation scalar leptoquarks and heavy right-handed neutrinos in $\tau_l\tau_h jj$ final states in pp collisions at 13 TeV*, CMS-PAS-EXO-16-023 (2016).

[38] A. Abashian *et al.* (Belle Collaboration), *The Belle detector*, Nucl. Instr. and Meth. A **479**, 117 (2002); also see the detector section in J. Brodzicka *et al.*, *Physics achievements from the Belle experiment*, Prog. Theor. Exp. Phys. **2012**, 04D001 (2012).

[39] B. Aubert *et al.* (BaBar Collaboration), *The BABAR detector*, Nucl. Instr. and Meth. A **479**, 1 (2002).

[40] A. Matyja *et al.* (Belle Collaboration), *Observation of $B^0 \to D^{*-}\tau^+\nu_\tau$ Decay at Belle*, Phys. Rev. Lett. **99**, 191807 (2007).

[41] A. Bozek *et al.* (Belle Collaboration), *Observation of $B^+ \to \bar{D}^{*0}\tau^+\nu_\tau$ and evidence for $B^+ \to \bar{D}^0\tau^+\nu_\tau$ at Belle*, Phys. Rev. D **82**, 072005 (2010).

[42] M. Huschle *et al.* (Belle Collaboration), *Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, Phys. Rev. D **92**, 072014 (2015).

[43] Y. Sato *et al.* (Belle Collaboration), *Measurement of the branching ratio of $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ relative to $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays with a semileptonic tagging method*, Phys. Rev. **94**, 072007 (2016).

[44] B. Aubert *et al.* (BaBar Collaboration), *Observation of the Semileptonic Decays $B \to D^*\tau^-\bar{\nu}_\tau$ and Evidence for $B \to D\tau^-\bar{\nu}_\tau$*, Phys. Rev. Lett. **100**, 021801 (2008).

[45] J.P. Lees *et al.* (BaBar Collaboration), *Evidence for an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays*, Phys. Rev. Lett. **109**, 101802 (2012); J.P. Lees *et al.* (BaBar Collaboration), *Measurement of an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays and implications for charged Higgs bosons*, Phys. Rev. D **88**, 072012 (2013).

[46] J.A. Bailey *et al.* (Fermilab Lattice and MILC Collaborations), $B \to D\ell\nu$ *form factors at nonzero recoil and* $|V_{cb}|$ *from 2 + 1-flavor lattice QCD*, Phys. Rev. D **92**, 034506 (2015).

[47] H. Na *et al.* (HPQCD Collaboration), $B \to D\ell\nu$ *form factors at nonzero recoil and extraction of* $|V_{cb}|$, Phys. Rev. D **92**, 054510 (2015).

[48] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3**, S08005 (2008).

[49] R. Aaij *et al.* (LHCb Collaboration), *Measurement of the Ratio of Branching Fractions* $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$, Phys. Rev. Lett. **115**, 111803 (2015).

[50] F.U. Bernlochner, Z. Ligeti and S. Turczyk, *A proposal to solve some puzzles in semileptonic B decays*, Phys. Rev. D **85**, 094033 (2012).

[51] S. Kurokawa and E. Kikutani, *Overview of the KEKB accelerators*, Nucl. Instr. and. Meth. A **499**, 1 (2003), and other papers included in this volume; T. Abe *et al.*, *Achievements of KEKB*, Prog. Theor. Exp. Phys. **2013**, 03A001 (2013) and references therein.

[52] K. Abe *et al.* (Belle Collaboration), *Observation of Large CP Violation in the Neutral B Meson System*, Phys. Rev. Lett. **87**, 091802 (2001).

[53] Z. Natkaniec *et al.*, *Status of the Belle silicon vertex detector*, Nucl. Instr. and Meth. A **560**, 1 (2006).

[54] O. Toker, *VIKING, a CMOS low noise monolithic 128 channel frontend for Si-strip detector readout*, Nucl. Instr. and Meth. A **340**, 572 (1994); M. Tanaka *et al.*, *A control and readout system for the BELLE silicon vertex detector*, Nucl. Instr. and Meth. A **432**, 422 (1999).

[55] H. Kichimi *et al.*, *The BELLE TOF system*, Nucl. Instr. and Meth. A **453**, 315 (2000).

[56] T. Iijima *et al.*, *Aerogel Cherenkov counter for the BELLE detector*, Nucl. Instr. and Meth. A **453**, 321 (2000).

[57] A. Abashian *et al.*, *The* K_L/μ *detector subsystem for the BELLE experiment at the KEK B-factory*, Nucl. Instr. and Meth. A **449**, 112 (2000).

[58] Section 15 in A.J. Bevan *et al.*, *The Physics of the B Factories* Eur. Phys. J. C **74**, 3026 (2014).

[59] K. Hanagaki *et al.*, *Electron identification in Belle*, Nucl. Instr. and Meth. A **485**, 490 (2002).

[60] A. Abashian *et al.*, *Muon identification in the Belle experiment at KEKB*, Nucl. Instr. and Meth. A **491**, 69 (2002).

[61] Section 5 in A.J. Bevan *et al.*, *The Physics of the B Factories*, Eur. Phys. J. C **74**, 3026 (2014).

[62] K. Hagiwara A.D. Martin and D. Zeppenfeld, *τ polarization measurements at LEP and SLC*, Phys. Lett. B **235**, 198 (1990).

[63] D.J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instr. and Meth. A **462**, 152 (2001).

[64] T. Sjöstrand *et al.*, *High-Energy-Physics Event Generation with PYTHIA 6.1*, Comput. Phys. Commun. **135**, 238 (2001).

[65] E. Barberio and Z. Wąs, *PHOTOS – a universal Monte Carlo for QED radiative corrections: version 2.0*, Comput. Phys. Commun. **79**, 291 (1994).

[66] R. Brun *et al.*, *GEANT 3*, CERN Report DD/EE/84-1 (1987).

[67] N. Isgur, D. Scora, B. Grinstein and M.B. Wise, *Semileptonic B and D decays in the quark model*, Phys. Rev. D **39**, 799 (1989); D. Scora and N. Isgur, *Semileptonic meson decays in the quark model: An update*, Phys. Rev. D **52**, 2783 (1995).

[68] A.K. Leibovich, Z. Ligeti, I.W. Stewart and M.B. Wise, *Semileptonic B decays to excited charmed mesons*, Phys. Rev. D **57**, 308 (1998).

[69] F.U. Bernlochner and Z. Ligeti, *Semileptonic $B_{(s)}$ decays to excited charmed mesons with e, μ, τ and searching for new physics with $R(D^{**})$*, Phys. Rev. D **95**, 014022 (2017).

[70] M. Feindt *et al.*, *A hierarchical NeuroBayes-based algorithm for full reconstruction of B mesons at B factories*, Nucl. Instr. and Meth. A **654**, 432 (2011).

[71] M. Feindt, *A Neural Bayesian Estimator for Conditional Probability Densities* arXiv:physics/0402093 (2004).

[72] 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).

[73] S.H. Lee *et al.* (Belle Collaboration), *Evidence for $B^0 \to \pi^0\pi^0$*, Phys. Rev. Lett. **91**, 261801 (2003).

[74] A. Sibidanov *et al.* (Belle Collaboration), *Study of exclusive $B \to X_u \ell \nu$ decays and extraction of $|V_{ub}|$ using full reconstruction tagging at the Belle experiment*, Phys. Rev. D **88**, 032005 (2013).

[75] T. Skwarnicki, *A study of the radiative CASCADE transition between the upsilon-prime and upsilon resonances*, Ph.D. thesis, DESY-F31-86-02 (1986).

[76] A. Drutskoy *et al.* (Belle Collaboration), *Observation of $B \to D^{(*)}K^-K^{0(*)}$ decays*, Phys. Lett. B **542**, 171 (2002).

[77] Z. Xing, H. Zhang and S. Zhou, *Updated values of running quark and lepton masses*, Phys. Rev. D **77**, 113016 (2008).

[78] T. Abe *et al.* (Belle II Collaboration), *Belle II Technical Design Report*, KEK Report 2010-1 (2010).

[79] For example, H.-G. Moser, On behalf of the DEPFET Collaboration, *The Belle II DEPFET pixel detector*, Nucl. Instr. and Meth. A **831**, 85 (2016).

[80] For example, K. Adamczyk *et al.* (Belle II SVD Collaboration), *Belle II scilicon vertex detector*, Nucl. Instr. and Meth. A **831**, 80 (2016).

[81] M. Akatsu *et al.*, *Time-of-propagation Cherenkov counter for particle identification*, Nucl. Instr. and Meth. A **440**, 124 (2000).

[82] T. Ohshima, *Time-of-Propagation Counter – A new Cherenkov Ring Imaging Detector*, ICFA Instr. Bull. **20**, 10 (2000).

[83] Y. Enari, *Progress report on Time-Of-Propagation counter – a new type of ring imaging Cherenkov detector*, Nucl. Instr. and Meth. A **494**, 430 (2002).

[84] K. Matsuoka, For the Belle II PID Group, *Design and performance study of the TOP counter*, Nucl. Instr. and Meth. A **732**, 357 (2013).

[85] For example, S. Iwata *et al.*, *Particle identification performance of the prototype aerogel RICH counter for the Belle II experiment*, Prog. Theor. Exp. Phys. **2016**, 033H01 (2016).

[86] M. Akatsu *et al.*, *MCP-PMT timing property for single photons*, Nucl. Instr. and Meth. A **528**, 763 (2004).

[87] N. Kishimoto *et al.*, *Lifetime of MCP-PMT*, Nucl. Instr. and Meth. A **564**, 204 (2006).

[88] K. Inami *et al.*, *Cross-talk suppressed multi-anode MCP-PMT*, Nucl. Instr. and Meth. A **592**, 247 (2008).

[89] T. Jinno *et al.*, *Lifetime-extended MCP-PMT*, Nucl. Instr. and Meth. A **629**, 111 (2011).

[90] S. Hirose, For the Belle II PID group, *Performance of the MCP-PMT for the Belle II TOP counter in a magnetic field*, Nucl. Instr. and Meth. A **766**, 163 (2014).

[91] T. Yonekura *et al.*, *Performance evaluation of the MCP-PMT for the Belle II TOP counter during mass production*, Physical Society of Japan Fall Meeting (2014); T. Yonekura, *Acceptance tests of MCP-PMT*, 19th B2GM (2014).

[92] S. Hirose *et al.*, *Development of the micro-channel plate photomultiplier for the Belle II TOP counter*, Nucl. Instr. and Meth. A **787**, 293 (2015).

[93] T. Nanut, *TOP Beam Background*, presented at 26th B2GM (2017).

[94] A. Lehmann *et al.*, *Significantly improved lifetime of micro-channel plate PMTs*, Nucl. Instr. and Meth. A **718**, 535 (2013).

[95] K. Matsuoka *et al.*, *Extension of the MCP-PMT lifetime*, presented at the 9th International Workshop on Ring Imaging Cherenkov Detectors (2016).

[96] T. Nakano *et al.*, *Multi-GeV laser-electron photon project at SPring-8*, Nucl. Phys. A **684**, 71 (2001).

[97] K. Matsuoka, For the Belle II PID Group, *Design and performance study of the TOP counter*, Nucl. Instr. and Meth. **732**, 357 (2013).

[98] K. Matsuoka, For the Belle II PID Group, *Performance study of the TOP counter with the 2 GeV/c positron beam at LEPS*, Proceedings of Science TIPP2014 (2014) 093.

[99] C. Pulvermacher, *Analysis software and full event interpretation for the Belle II experiment*, Ph.D thesis, Karlsruhe Institute of Technology, EKP-2015-00053 (2015).

[100] H. Albrecht *et al.* (Argus Collaboration), *Search for hadronic $b \to u$ decays*, Phys. Lett. B **241**, 278 (1990).