

Search for lepton-flavor-violating tau decays into a lepton and a vector meson at the Belle experiment

Belle実験におけるタウ粒子がレプトンフレーバーを破って
レプトンとベクトル中間子に崩壊する事象の探索

Noritsugu Tsuzuki

*High Energy Physics Laboratory
Graduate School of Science, Nagoya University*

February 20, 2023

Abstract

We are searching for the perfect physics model that provides every elementary particle phenomenon. Although many experimental results agree to the Standard Model (SM) within their uncertainties, some candidates of deviations from the SM raise many New Physics (NP) models. These NP models can be verified by predictions and experimental searches of other deviations. Charged lepton-flavor violation (CLFV) is clear proof because CLFV events have much larger branching fractions in some NP models than in the SM. Searches of various CLFV modes reveal which NP models remain.

This thesis reports the search for the CLFV in the τ decays into a light charged lepton ($\ell = e^\pm$ or $\mu^\pm$) and a vector meson ($V^0 = \rho^0, \phi, \omega, K^{*0}$ or $\bar{K}^{*0}$). These decay modes are sensitive to leptoquark models, which can explain the reported deviation candidates in some of B meson semileptonic decays (B -anomalies). A simple NP model with only one leptoquark predicts a large enhancement of the branching fractions of $\tau \rightarrow \mu\phi$ that consists of second and third generation particles. The reason is that the leptoquark should interact with these particles to explain the B -anomalies. In addition, several other NP models also predict branching fractions of $\mathcal{O}(10^{-10})$ – $\mathcal{O}(10^{-8})$ for $\tau \rightarrow \ell V^0$. To improve the sensitivity of $\tau \rightarrow \ell V^0$ searches, lots of τ particles should be measured in a clean experimental environment.

This analysis uses the 980 fb^{-1} of data collected with the Belle detector, corresponding to $9.0 \times 10^8 \tau^+\tau^-$ pairs. Most of the events in this data set have only one e^+e^- collision event with the center-of-mass energy around 10.58 GeV at the KEKB collider. The $\tau \rightarrow \ell V^0$ events are reconstructed from three charged tracks (and two gammas for the $\ell\omega$ modes). Such neutrino-less τ decays have a unique kinematical feature; the ℓV^0 invariant mass ($M_{\ell V^0}$) is close to the τ mass and the difference of the ℓV^0 energy from the beam energy (ΔE) is close to zero. The blind analysis approach is taken in this thesis, i.e., the signal-like events in the data are blinded in the $M_{\ell V^0}$ – ΔE plane until finalizing the event selection and background estimation. The author reconsiders the event selection criteria from the previous $\tau \rightarrow \ell V^0$ searches by introducing new variables and a multi-variate analysis using Boosted Decision Tree.

The signal efficiencies after the event selection are estimated to be 6.0% on average using the Monte Carlo simulation samples of $\tau \rightarrow \ell V^0$ events in the Belle detector. The distribution of background events is interpolated from the sideband region in the $M_{\ell V^0}$ – ΔE plane outside of the blind region, and the number of background events is estimated to be less than one for all of the $\tau \rightarrow \ell V^0$ modes. The largest background source is $ee \rightarrow q\bar{q}$ ($q = u, d, s$ and c), and the second largest is the $ee \rightarrow \tau^+\tau^-$ with $\tau^\pm \rightarrow \pi^\pm \pi^+ \pi^- \nu$ decay. Some two-photon events of $ee \rightarrow ee q\bar{q}$ and $ee \rightarrow ee \ell\ell$ also contribute to the background.

No significant excess of signal events is observed, and thus 90% confidence level upper limits are set on the $\tau \rightarrow \ell V^0$ branching fractions in the range of $(1.7\text{--}4.2) \times 10^{-8}$. These limits are relatively improved by 30% on average from the previous searches. The limit for the $\tau \rightarrow \mu\phi$ (2.3×10^{-8} , improved by 73%) sets more constraint on the single leptoquark model. The other limits for the $\tau \rightarrow \ell V^0$ also set constraints on different NP models. Although the allowed region of the single leptoquark model remains, most of the region can be searched at the Belle II experiment with 50 ab^{-1} data set.

Contents

1	Introduction	4
2	Physics motivation	6
3	Belle experiment	9
3.1	KEKB accelerator	9
3.2	Belle detector	10
3.2.1	Silicon Vertex Detector (SVD)	10
3.2.2	Central Drift Chamber (CDC)	12
3.2.3	Aerogel Cherenkov Counter (ACC)	13
3.2.4	Time-of-flight counter (TOF)	13
3.2.5	Electromagnetic Calorimeter (ECL)	14
3.2.6	K_L and muon detector (KLM)	14
3.2.7	Trigger	15
3.2.8	Data Acquisition (DAQ) system	17
3.2.9	Particle identification (PID)	17
3.3	Data set	19
4	Reconstruction and event selection	26
4.1	Reconstruction	26
4.2	Low multiplicity event suppression	30
4.3	Cut-base selection	30
4.4	Selection with Boosted Decision Tree	32
4.4.1	Control samples of background events	32
4.4.2	BDT training and outputs	35
5	Signal efficiency and background estimation	43
5.1	Correction for the MC events	43
5.2	Signal and background estimation	44
5.2.1	Signal region	44
5.2.2	Number of background events estimation	45
5.2.3	Validation of number of background estimation	48
5.2.4	Comparison with old background estimation method	48
5.3	Systematic uncertainty of the expected number of signal events	49
6	Results	67
6.1	LFV decay branching fraction upper limits	67
6.2	Discussion	68
6.2.1	Improvements from the previous paper	68
6.2.2	Prospect for the future Belle II experiment	69
7	Conclusion	76

A	Validation of N_{BG} estimation method using background MC	84
B	BDT training with MC control samples	92

Chapter 1

Introduction

The elementary particle physics aims to find all fundamental elements and describe their behaviors. The Standard Model (SM) can explain most of the experimental results with six quarks, six leptons, four gauge bosons, and one Higgs boson. On the other hand, the SM does not have solutions to some astrophysical observations, such as dark matter and matter-antimatter asymmetry. Many New Physics (NP) models have been discussed to explain these observations or some gaps between the SM predictions and experimental results. We can verify the NP models by measuring physical quantities for which the models predict deviations from the SM. A deviation can arise even in the SM particle decay if an NP particle is in the intermediate state.

The SM introduces a lepton-flavor conservation law to explain observed lepton interactions. Fermions (quarks and leptons) are categorized into three generations. The lepton-flavor is assigned to each lepton generation: e and ν_e have electron number one, μ and ν_μ have muon number one, τ and ν_τ have tau number one, and their antiparticles have minus one of each. These numbers are conserved before and after the interaction of elementary particles, e.g., $\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$. However, the observation of neutrino oscillations [1–3] tells us that the lepton-flavor conservation can be broken. It is possible to include the neutrino oscillations by extending the SM. Even in that case, the other processes with lepton-flavor violation are quite suppressed so that they are unobservable by current experiments.

In addition, the SM expects lepton-flavor universality (LFU); the coupling constants that each lepton generation has will be the same value. The gaps between the SM predictions and experimental results, muon magnetic anomaly [4, 5] and B -anomalies ($B \rightarrow D^{(*)} \tau \nu$ [6–13], $b \rightarrow s \ell \ell$ [14–17], etc.), imply the LFU violation. The standard deviations (σ) of the gaps are from two to four. These anomalies often motivate the lepton-flavor violation in NP models.

Charged lepton-flavor violation (CLFV) event is clear proof of NP. This paper focuses on the search for τ decays into a charged lepton ($\ell = e^\pm$ or $\mu^\pm$) and a neutral vector meson ($V^0 = \rho^0, \phi, \omega, K^{*0}$ or $\bar{K}^{*0}$). The Belle experiment is the most powerful in the $\tau \rightarrow \ell V^0$ searches thanks to its integrated luminosity 980 fb^{-1} (corresponds to 9.0×10^8 $\tau^+ \tau^-$ pairs) and clean experimental environment. The data are taken at and around the center-of-mass energies of $\Upsilon(1-4S)$ resonance. Those CLFV searches have been reported using 854 fb^{-1} of Belle data [18], and 90% confidence level upper limits have been set on the branching fractions in the range of $(1.2-8.4) \times 10^{-8}$.¹ These branching fractions are below some NP predicted upper limits. Because the event selection criteria in ref. [18] have room for improvement, I optimize it with new variables and a multi-variate analysis.

Chapter 2 is about the physics motivation of $\tau \rightarrow \ell V^0$ searches. Chapter 3 describes the KEKB accelerator, Belle detector, and Belle data set. Chapter 4 is about the event reconstruction and selection criteria. I estimate the signal efficiencies, the numbers of background events, and their uncertainties in Chapter 5. Then, I discuss the results of all $\tau \rightarrow \ell V^0$ analyses in Chapter 6, and finally conclude this thesis in Chapter 7.

¹In common high-energy physics usage, this credibility level has been reported as “confidence level,” which is a frequentist-statistics term.

Chapter 2

Physics motivation

The SM has a lepton-flavor conservation law and prohibits CLFV processes. The CLFV processes via the neutrino oscillations (e.g., figure 2.1) can occur in the SM extended model with non-zero neutrino mass. Even in that case, their branching fractions are extremely small [19] ($< \mathcal{O}(10^{-40})$) because they are proportional to the square of the ratio of the neutrino mass differences to the $W^\pm$ boson mass. In other words, observation of any CLFV event is clear evidence of NP.

An important motivation for the CLFV search is the recent experimental observations about the LFU violation. The weak interaction bosons (Z and $W^\pm$) have the same coupling constant for each lepton-generation. Whereas most of the experimental results agree with the predictions of the SM with LFU, some of B meson semileptonic decay processes show deviations (B -anomalies [6–17]). That implies something NP particle that causes LFU violation especially in these processes. Many NP models, which explain these LFU violations, predict enhancement of the CLFV branching fractions.

The CLFV events are searched for various processes: $\mu \rightarrow e\gamma$ [22], $\mu \rightarrow eee$ [23], μ -to- e conversion in nuclei [24], τ decay modes, boson decays [25–28], etc. No CLFV event has been observed yet, and the most stringent upper limit of the CLFV branching fraction is set for $\mu \rightarrow e\gamma$ as 4.2×10^{-13} . The CLFV searches in boson decays are sensitive if the boson violates the lepton-flavor conservation; the upper limits of the CLFV branching fractions are $\mathcal{O}(10^{-7})$ – $\mathcal{O}(10^{-6})$ for the Z decays and $\mathcal{O}(10^{-5})$ – $\mathcal{O}(10^{-3})$ for the Higgs decays. Although the upper limits of τ CLFV decay modes are $\mathcal{O}(10^{-8})$, they are sensitive for some NP models whose new particles couple to heavy or third generation particles. The B -anomalies motivate τ CLFV searches because they are observed in processes that include second and third generation quarks and leptons.

The τ mass, which is $1.78 \text{ GeV}/c^2$ and the heaviest of the leptons, allows τ hadronic decays. The searches for CLFV in τ decays are reported for 48 modes. Figure 2.2 shows the 90% confidence level upper limits of these searches [29]. The Belle experiment provides the best upper limits for many modes. Representative τ CLFV modes are the $\tau \rightarrow \ell\gamma$, which have large enhancement in the supersymmetric models [30–32], the model with Majorana neutrinos [33],

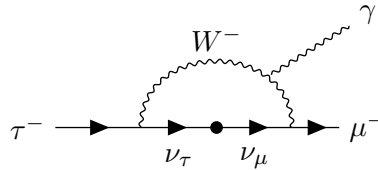


Figure 2.1: Feynman diagram for $\tau \rightarrow \mu\gamma$ CLFV process in the SM. The branching fraction of this process is extremely small ($\mathcal{O}(10^{-54})$). Figure is generated using the TikZ-Feynman package [20, 21]

etc. In terms of experimental sensitivity, the $\tau \rightarrow \mu\mu\mu$ mode is promising because we can strongly suppress its background events even in the future Belle II experiment. To verify various NP models, we should search various CLFV modes [34].

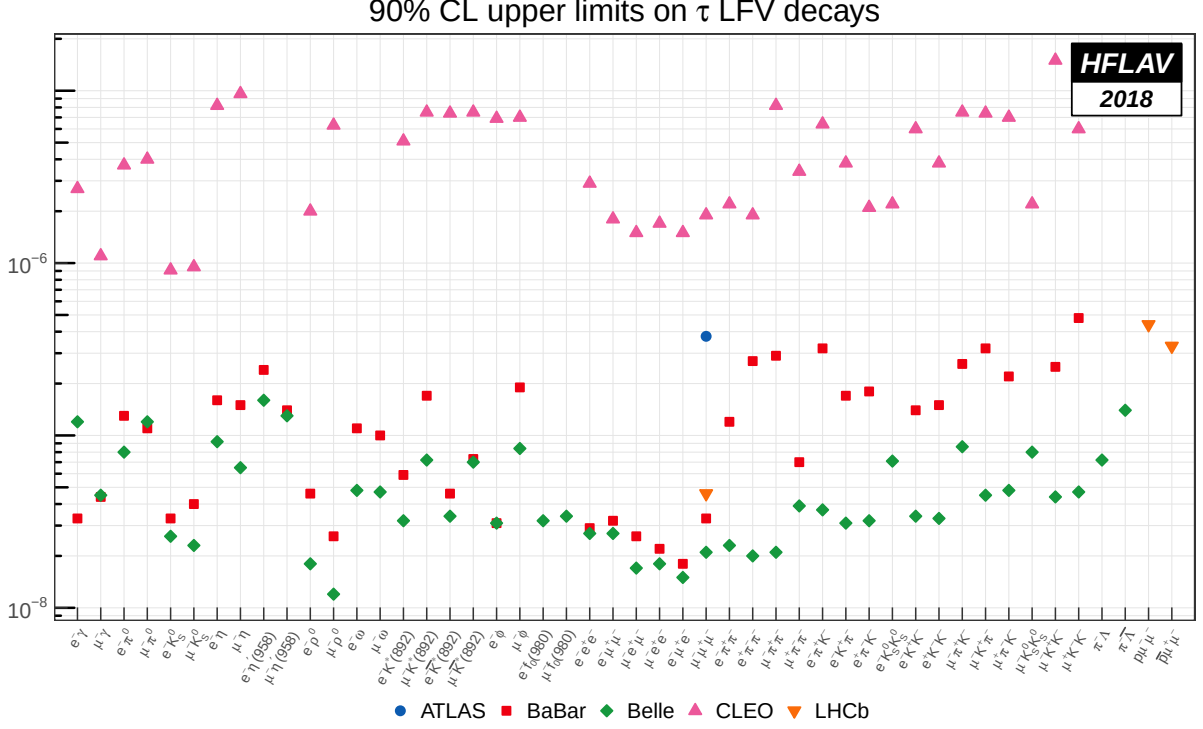


Figure 2.2: Tau lepton-flavor-violating decay branching fraction upper limits. Figure from ref. [29]. The $\tau \rightarrow \ell\gamma$ searches are updated with the full 980 fb⁻¹ of Belle data [35].

The $\tau \rightarrow \ell V^0$ searches have an advantage in the verification of leptoquark models [36–41]¹. The leptoquark is an NP particle that couples to a lepton-quark pair. These leptoquark models are motivated by the B -anomalies. If we assume a simple model with only one leptoquark to explain the B -anomalies, the leptoquark should couple to both the second and third generation fermions at least. Then, the $\tau \rightarrow \mu\phi$ decay is induced via the leptoquark (figure 2.3), and the prediction of its branching fraction is $\mathcal{O}(10^{-10})$ – $\mathcal{O}(10^{-8})$ [36]. Figure 2.4 shows the predictions of τ CLFV branching fractions with the single leptoquark model. We may observe the $\tau \rightarrow \mu\phi$ events in the Belle data set, even if others ($\tau \rightarrow \ell\gamma$, $\tau \rightarrow \mu\mu\mu$, etc.) could not be observed.

Some other NP models also predict branching fractions of $\mathcal{O}(10^{-10})$ – $\mathcal{O}(10^{-8})$ for $\tau \rightarrow \ell V^0$: e.g., the littlest Higgs model with T parity [44], type-III seesaw model [45], unparticle physics [46], and SM extended by heavy singlet Dirac neutrinos [47]. Part of parameter region in these models are also verifiable in this paper.

The sensitivities of $\tau \rightarrow \ell V^0$ searches are dominated by the number of τ particles measured in a clean experimental environment. I use the Belle experiment data-set (980 fb⁻¹) provided by e^+e^- collisions in the KEKB accelerator. The upper limit of $\tau \rightarrow \ell V^0$ branching fraction comes down in an inverse proportion to the integrated luminosity or to the square-root of the integrated luminosity, depending on the number of background events. The upper limit is also in an inverse proportion to the signal efficiency. The integrated luminosity increases by 15% from the previous paper [18], it results in 9% improvement on the average upper limit if I use the same event selection. For further improvement, refinement of the event selection is necessary. First, I loosen some event selection in the previous paper to improve the signal efficiency. Then

¹One of the B anomalies which motivated the models described in those references is the $R(K^{(*)})$ anomaly reported by the LHCb experiment [42], but it disappeared in their updated analysis [43].

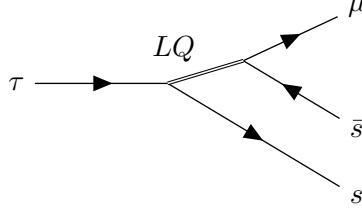


Figure 2.3: Feynman diagram for $\tau \rightarrow \mu\phi$ via a leptoquark.
Figure is generated using the TikZ-Feynman package [20, 21].

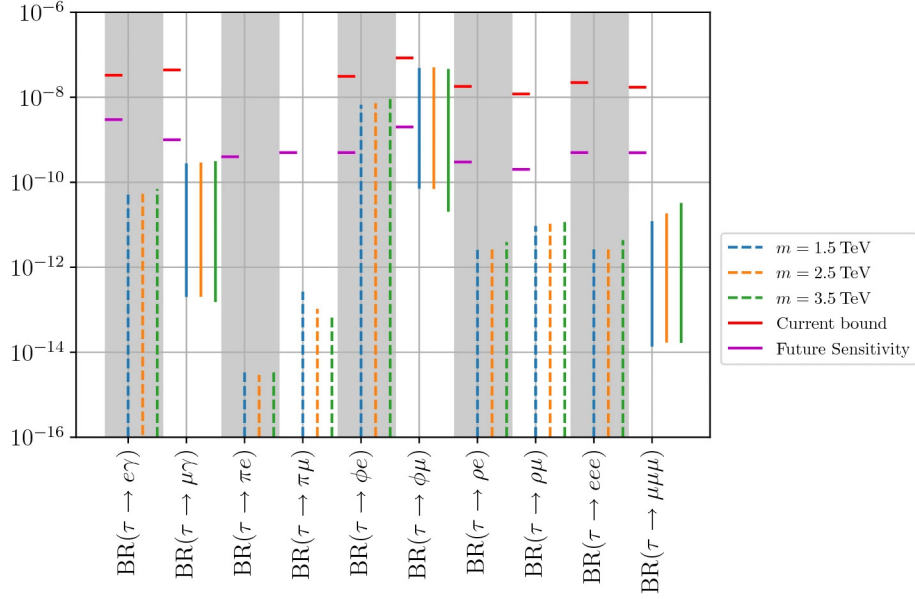


Figure 2.4: Lepton flavour violating τ decay modes expected to be searched for at the Belle II experiment. The blue, orange and green lines respectively denote the 90% range for leptoquark masses of 1.5, 2.5 and 3.5 TeV while the horizontal red (purple) lines denote the current (future) bound at 90% CL. The dashed lines correspond to predictions of observables depending only on couplings that are compatible with 0 and their top edges correspond to 90% upper limits. Figure from ref. [36].

I introduce some variables and a multi-variate analysis to the event selection for the background suppression.

Chapter 3

Belle experiment

This thesis uses the e^+e^- collision data in the KEKB-Belle experiment. This chapter shows the design of the KEKB accelerator and Belle detector at first. The method of the particle identification in the Belle detector is described in the next section. I also describe the data set for this analysis and the simulation data set of the signal and background events.

3.1 KEKB accelerator

The KEKB accelerator [48, 49] is designed for a B -factory experiment, which produces and analyzes b -quarks. The target center-of-mass (c.m.) energy is 10.58 GeV, corresponding to the mass of the $\Upsilon(4S)$ resonance, and $\Upsilon(4S)$ decays into a B meson and an anti- B meson. Because the cross section of $e^+e^- \rightarrow \tau^+\tau^-$ is 0.9 nb and similar to that of $e^+e^- \rightarrow B\bar{B}$ (1.1 nb) at the c.m. energy, a B -factory is also a τ -factory. KEKB also provides collisions with the c.m. energy between 9.4 to 11.2 GeV (see section 3.3 for each integrated luminosity).

The advantages of KEKB are high luminosity and clean e^+e^- collisions. The recorded peak luminosity is $2.11 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ and we can use the 980.4 fb^{-1} data in total. Most of the data record only one e^+e^- collision per event.

The High Energy Ring (HER) is for the electron beam with energy of 8.0 GeV, and the Low Energy Ring (LER) is for the positron beam with energy of 3.5 GeV. They are located at KEK in Japan utilizing the existing TRISTAN tunnel. Figure 3.1 shows the layout of KEKB. The electron beam is produced from thermoelectrons in an electron gun, shaped into bunches, and accelerated to 8.0 GeV in the linac before injecting to the HER. The positron beam is produced at the middle of the linac by the collision between a tungsten target and the electron beam with energy of 3.7 GeV, and accelerated to 3.5 GeV before injecting to the LER. Both beams can be injected to the rings while the e^+e^- collision is ongoing at the opposite side (Tsukuba experimental hall).

The injected beam bunches are stored in the HER and LER. The maximum beam currents were 1.4 A in HER and 2.0 A in LER during the actual operation. Accelerator Resonantly coupled with Energy Storage (ARES in figure 3.1) is a RF cavity system to stabilize such high current beam operation. Superconducting RF cavity (Super RF in figure 3.1) system also suppress the instability of the high current beam. Wiggler magnets suppress the beam emittance and equalize the radiation damping time between the two rings. The crossing angle at the interaction point (IP) is 22 mrad. The Crab cavity changes the angle of a beam bunch and leads to head-on collision at the IP to increase luminosity. Near the IP region, the e^+ beam pipe is aligned with the axis of the Belle detector solenoid to reduce magnetic bending for the lower momentum beam. Both beams are squeezed at the IP by a pair of superconducting final-focus quadrupole magnets (QCS). Final state particles from a e^+e^- collision are recorded by the Belle detector surrounding the IP.

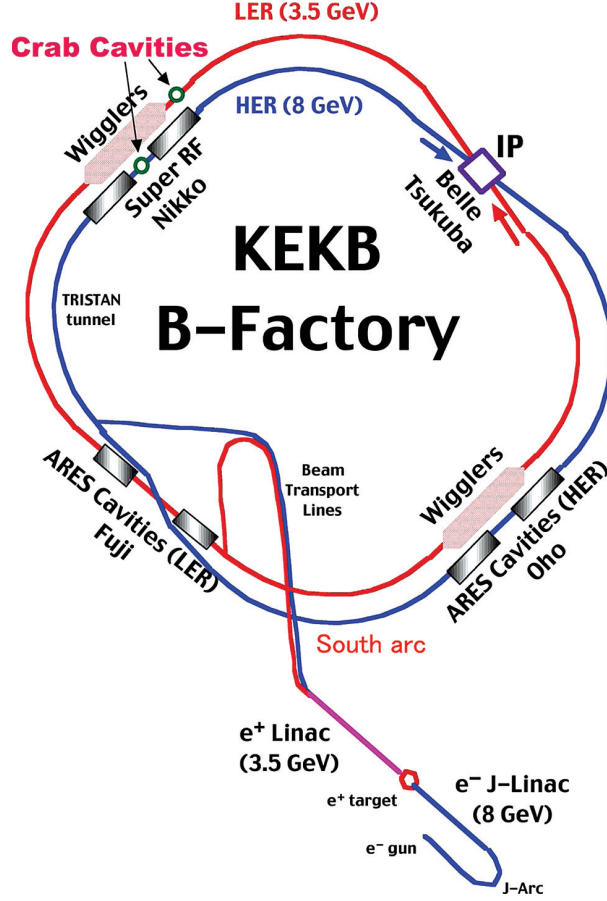


Figure 3.1: Schematic layout of KEKB. The four experimental halls are called Tsukuba, Nikko, Fuji and Oho. Figure from ref. [50]

3.2 Belle detector

The Belle detector is a spectrometer covering a large solid angle (figure 3.2). This detector consists of a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), time-of-flight scintillation counters (TOF) and an electromagnetic calorimeter composed of 8736 CsI(Tl) crystals (ECL). These devices are located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux return located outside of the coil is instrumented to detect K_L^0 mesons and identify muons (KLM). Charged particles are tracked by SVD and CDC in 1.5 T, and their momenta are calculated from the track curvature. Charged particle species (electron, muon, pion, kaon and proton) are identified by CDC, TOF, ACC, ECL and KLM. Energy of Photon is measured by ECL. The Belle detector is described in detail elsewhere [49, 51].

EFC is a pair of BGO (Bismuth Germanate Oxide) crystal arrays on the QCS cryostats. Because I do not use this EFC information, I omit the explanation in this thesis.

The Belle detector is separated into forward, barrel, and backward regions. The forward (backward) region is the base of the cylinder in the direction of the electron (positron) beam heading, and the barrel region is the side of the cylinder.

3.2.1 Silicon Vertex Detector (SVD)

SVD is for a precision measurement of the decay vertices near the IP. A charged particle excites electrons in silicon and the excited electrons are amplified by an electric field. We detect the electric current as a SVD hit. The tracks reconstructed in CDC (section 3.2.2) are extrapolated

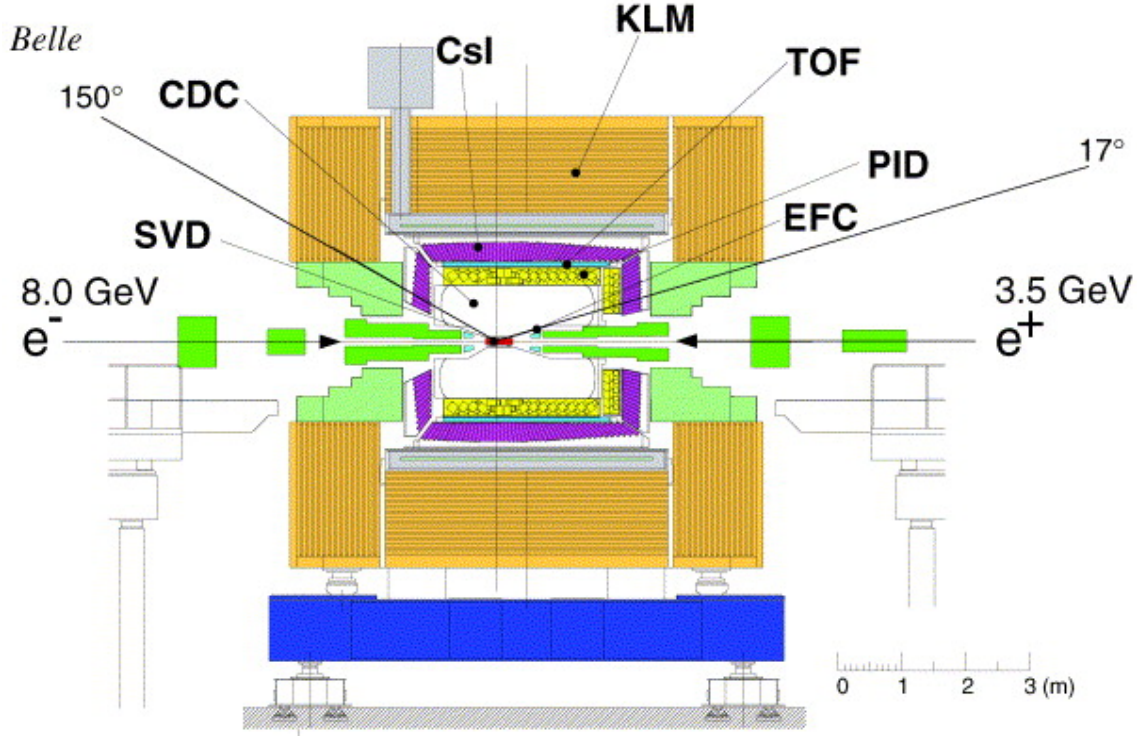


Figure 3.2: Side view of the Belle detector. Figure from ref. [51].

to the SVD hits that are the innermost parts of the Belle detector. Figure 3.3 shows the structure of SVD. Each ladder is composed of several double-sided silicon strip detectors (DSSD) with hit efficiency of about 95%. The first (innermost) layer of the SVD located just outside of the beam pipe to improve the vertex resolution. The readout electronics is placed outside of the tracking volume to reduce the multiple-Coulomb scattering. The support structure to avoid bending is also low in mass.

SVD was upgraded and replaced at the middle of the Belle experiment to improve the vertex resolution and the radiation tolerance of the readout electronics. The SVD detector coverage in θ is also extended, where θ is the polar angle with respect to the direction opposite to the e^+ beam in the laboratory frame. Table 3.1 is the comparison of SVD1 (the first period) and SVD2 (the last period). The impact parameter resolutions of the SVD1 and SVD2 are shown in figure 3.4.

Table 3.1: Comparison between the SVD1 and SVD2.

	SVD1	SVD2
Period	1999–2003	2003–2010
Collected luminosity (fb^{-1})	156	824
Beam pipe inner radius at the IP (mm)	20	15
Number of layers	3	4
Radii of the layers (mm)	30, 45.5, 60.5	20, 43.5, 70, 88
Detector coverage in θ ($^\circ$)	23–139	17–150
Radiation tolerance of the electronics (krad)	200	20,000
Signal-to-noise ratio	17–40	16–36

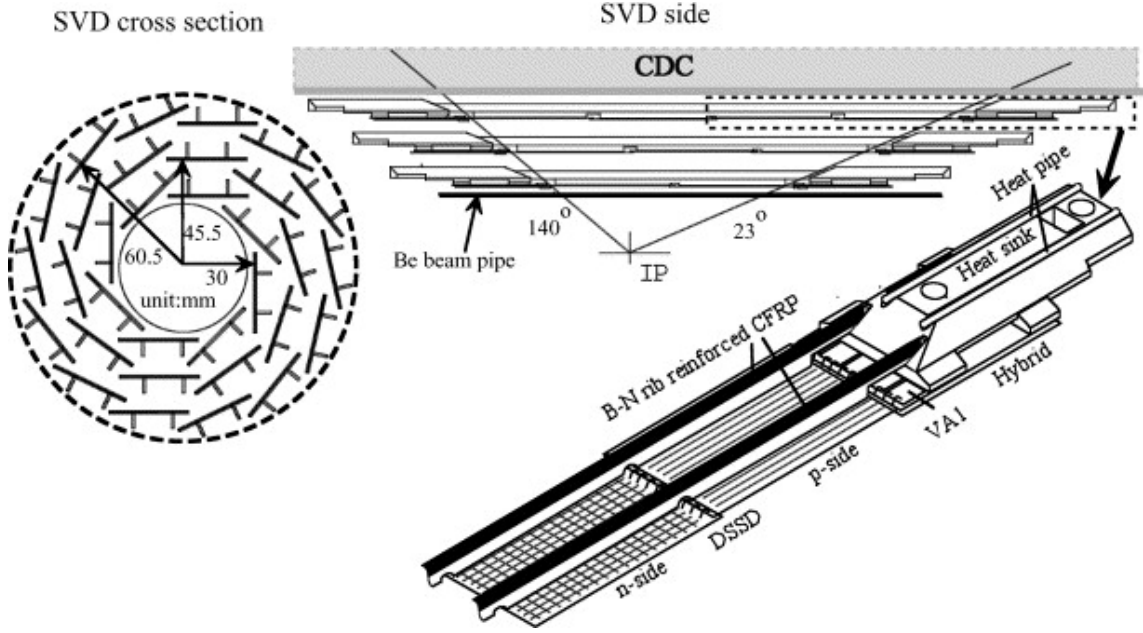


Figure 3.3: The SVD1 configuration. The support ribs are made of boron nitride (B-N) sandwiched by carbon-fiber reinforced plastic (CFRP). VA1 is a readout electronics chip for the SVD1. Figure from ref. [52].

3.2.2 Central Drift Chamber (CDC)

The CDC is a charged-particle tracking detector using gas and wires. A charged particle ionizes gas in the CDC and the ionization electrons cause an electron avalanche around an anode wire. The distance of closest approach between the charged particle and the anode wire are calculated from the drift time of the ionization electron. The track of the charged particle is obtained from the combination of many hits on the anode wires. We also get the momentum of the track from the track curvature in the 1.5 T magnetic field.

The structure of the CDC is cylindrical and asymmetric in the z direction with an angular coverage of $17^\circ \leq \theta \leq 150^\circ$. For the SVD1, the CDC contains 50 anode wire layers and three cathode strip layers. The anode wire layers consist of 32 axial layers and 18 small-angle stereo layers to obtain three-dimensional track. Each cell is almost rectangular shape and consists of one anode wire and eight field wires. The cathode strip layers are glued on the surfaces of the CDC structure; one is the innermost layer of CDC, the other two layers are between the second and third anode wire layers. Figure 3.5 shows the CDC inner structure around the cathode strip layers. The inner and outer radii are 83 and 874 mm, respectively. For the SVD2, the cathode strip layers and innermost two anode wire layers were replaced by two anode wire layers with compact small cells.

The gas in the CDC should be low mass to minimize multiple Coulomb scattering and background from synchrotron radiation, while the charged particle energy loss in the gas (dE/dx) can be measured with good resolution. We use a mixture of a He and C_2H_6 gases in equal quantities; The gas has a $\tilde{640}$ m radiation length. Figure 3.6 shows the correlation between the dE/dx and momentum for electron, pion, kaon, and proton tracks.

The transverse momentum (p_t) resolution of the CDC is $0.201\%p_t[\text{GeV}/c] \oplus 0.290\%/\beta$ (0.35% at $p_t = 1.0 \text{ GeV}/c$), where β is the velocity of the track divided by the light speed c . The resolution of the lowest 80% truncated mean of dE/dx is 7.0%, 5.5% or 5.7% for pion, electron or muon, respectively. The tracking efficiency is shown in figure 3.7.

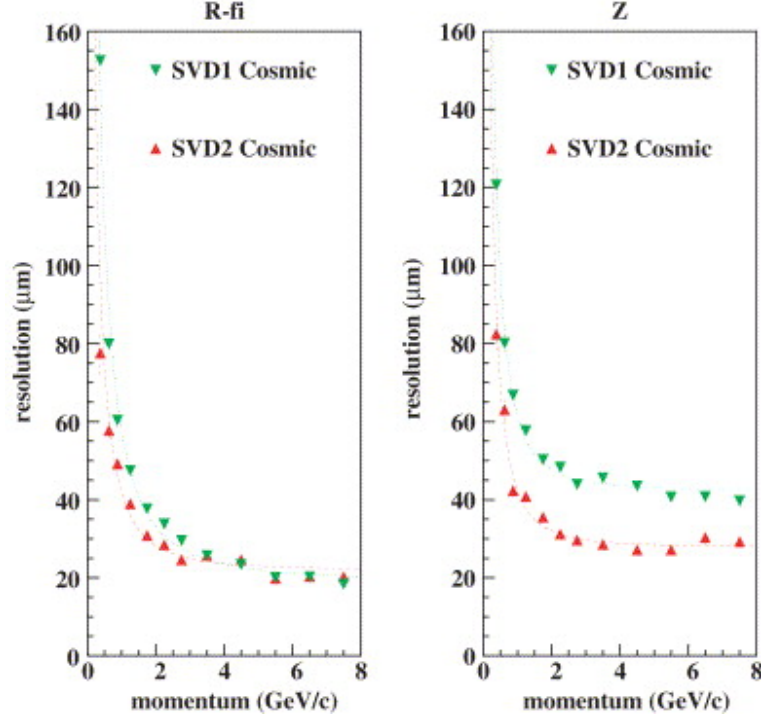


Figure 3.4: Impact parameter resolutions of the SVD1 and SVD2. Figure from ref. [53].

3.2.3 Aerogel Cherenkov Counter (ACC)

The ACC performs the particle identification by emission of Cherenkov light in a silica aerogel radiator with refractive index (n) of 1.010–1.030. Each ACC module points to the IP (figure 3.8). The velocity threshold of Cherenkov light emission is high as the refractive index is low, and thus the ACC with low refractive index covers the particle identification of high momentum particles. Because no TOF counter is installed in the forward endcap region, the ACC in the region should cover the particle identification of relatively low momentum particles.

A fine mesh-type photomultiplier tube (FM-PMT [57]), which can work in the 1.5 T magnetic field, is attached to the aerogel. The diameter and the number of FM-PMT on each aerogel depends on the aerogel refractive index, which determines the Cherenkov photon yield. The average number of photoelectrons for light velocity particles are 10–30 for each counter row in the barrel and each layer of the endcap. Figure 3.9 shows pulse heights of the ACC outputs with light speed particles (Bhabha e^+e^-) and β below the ACC threshold (Kaons).

3.2.4 Time-of-flight counter (TOF)

The TOF counter is a plastic scintillation counter in the barrel region and provides particle identification information by measuring the particle velocity (β , divided by the light speed c). The TOF module (figure 3.10) consists of two TOF counters and one Trigger Scintillation Counter (TSC), which is a thin plastic scintillation counter. A fine mesh-type photomultiplier tube (FM-PMT [59]) with two inches diameter is attached to the scintillator. The TOF system consists of 128 TOF counters and 64 TSCs and covers the polar angle range from 34° to 120° at a radius of 1.2 m from IP.

The main purpose of the TOF measurement is separation of kaon and pion tracks for particle momenta below 1.2 GeV/ c . The TOF counters have about 100 ps time resolution. Figure 3.11 shows the Kaon identification efficiency and pion fake rate.

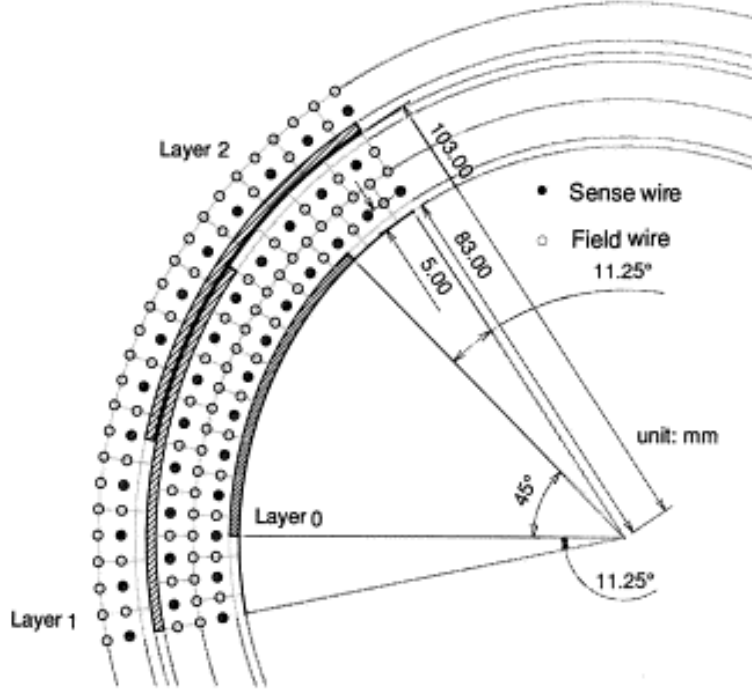


Figure 3.5: Inner cell structure and the cathode sector configuration. The shaded areas are one-eighth segments of each cathode strip layer. Figure from ref. [54].

3.2.5 Electromagnetic Calorimeter (ECL)

The ECL measures the energy of a photon and identifies electrons by the energy loss in the CsI(Tl) crystal (dE/dx). Figure 3.12 shows the structure of the ECL. The CsI(Tl) crystal nearly points to the IP and the ECL covers the polar angle range from 17.0° to 150.0° . Small gaps between the endcap and barrel calorimeters are necessary for the cable pathways and supporting structures of the inner detectors. These gaps are approximately 3% of the total acceptance.

The size of each CsI(Tl) crystal is designed to absorb 80% of the photon energy when the photon injects at the center of the crystal. The typical dimensions are $5.5 \text{ cm} \times 5.5 \text{ cm}$ (inner face) and $6.5 \text{ cm} \times 6.5 \text{ cm}$ (outer face). The 30 cm length is 16.2 times longer than the CsI(Tl) radiation length. Each crystal is wrapped in a single layer of $200 \mu\text{m}$ thick Gore-Tex teflon as a diffuse reflector to collect scintillation light at the photo-detector.

Two silicon photodiodes with a $10 \text{ mm} \times 20 \text{ mm}$ active area are glued on the outer face of the crystal via acrylic plate of 1 mm thickness. The output signals from the two photodiodes on the same crystal are summed up. The environment for the ECL is a low-humidity (5%) and lower than 30°C temperature.

The energy resolution is studied in the photon beam test (figure 3.13). From the KEKB collision data of $e^+e^- \rightarrow \gamma\gamma$ and Bhabha events, the ECL energy resolution for high energy photons and electrons ($E = 4\text{--}7 \text{ GeV}$) is 2.2%. The mass resolution of π^0 , reconstructed from two photons with $E_\gamma > 20 \text{ MeV}$, is $4.9 \text{ MeV}/c^2$.

3.2.6 K_L and muon detector (KLM)

The KLM discriminates a muon from the other charged particles and detect a K_L . The KLM consists of alternating layers of glass-electrode resistive plate counters (RPCs) and 4.7 cm thick iron plates. A hadron produces a shower of ionizing particles in the iron and charged particles in the shower are detected by the RPCs. A muon track penetrates the KLM with smaller deflection on average than hadrons. The KLM covers an angular range from 20° to 155° . Electrons and photons stop in the ECL, and tracks with momentum of $500 \text{ MeV}/c$ or less also do not reach

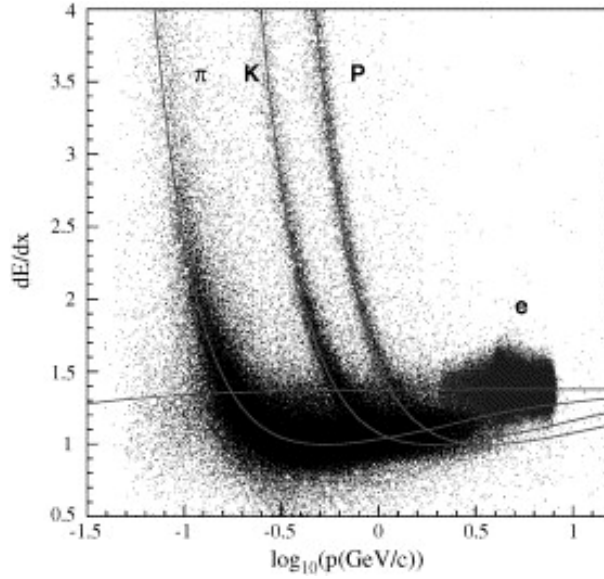


Figure 3.6: The mean of the lowest 80% of dE/dx for each track vs. momentum for electron, pion, kaon, and proton (observed in collision data). Figure from ref. [55].

KLM due to the magnetic field inside the Belle detector.

The RPC consists of two parallel plate glass-electrodes with high bulk resistivity and a gas-filled gap between the glass-electrodes. The barrel (endcap) RPCs are operated at 4.3 (4.2) kV/mm in the gas-filled gap. An ionizing particle create a streamer in the gap and the streamer makes a local discharge of the plate. The discharge induces a signal on external pickup-strips and the signal location and time are recorded. The discharge is quenched with the high resistivity of the plates and the gas. The bulk resistivity of the glass-electrode, about $5 \times 10^{12} \Omega\text{cm}$, limits the rate capability to about 0.2 Hz/cm². This is acceptable because the typical RPC rate is 0.03 Hz/cm² in the KLM. The glass-electrode in the barrel (endcap) KLM is 2.4 (2.0) mm thick and carbon-coated to achieve the surface resistivity of 10^6 – 10^7 (10^7 – 10^8) Ω/square . The gas is a non-flammable mixture of 62% HFC-134a (CH_2FCF_3), 30% argon and 8% butane-silver (C_4H_{10}). Butane-silver is a mixture of approximately 70% n-butane and 30% iso-butane.

Figure 3.14 shows the cross-section of a KLM superlayer module, which consists of two RPCs inserted between the orthogonal two copper pickup-strips. Each RPC is electrically insulated with a double layer of 0.125 mm thick mylar. The thickness of a superlayer module is less than 3.7 cm including aluminum enclosure. The module provides three-dimensional space point.

The barrel (endcap) KLM has 15 (14) superlayers and 14 layers of iron plates. This is 3.9 interaction length of material for a particle traveling perpendicular to the detector plane. ECL also provides another 0.8 interaction length of material to convert K_L . Plumbing for the cooling of the superconducting solenoid is in the upper backward region. This amounts to less than 2% of the solid angle of the barrel coverage.

The efficiency of single RPC is 90–95% and that of superlayer module is 98% or more. The module can work even one of the RPCs has trouble. The average efficiency of KLM is better than 97%. The spatial resolution of the module is 1.2 cm.

The K_L detection is not necessary in this thesis. The muon identification performance is in section 3.2.9, including the other detectors.

3.2.7 Trigger

The trigger system chooses the events to be recorded within 2.2 μs from a e^+e^- collision. The CDC, TOF/TSC, ECL and KLM send trigger signals to the global decision logic (GDL). The GDL categorizes events, synchronizes the events with the beam crossing time, and sends the

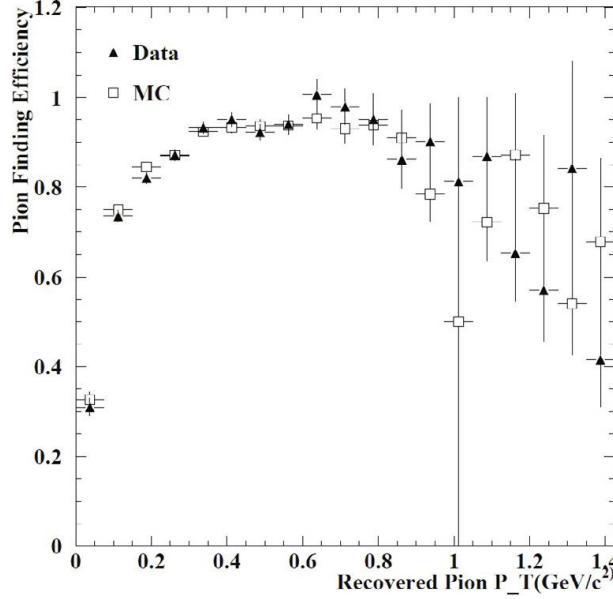


Figure 3.7: Reconstruction efficiency for charged tracks as a function of the particle’s transverse momentum for MC (open square) and data (filled triangle). Figure from ref. [56]

final trigger to the DAQ system in section 3.2.8.

The collision event rate of our interesting physics is about 91 Hz at luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. Bhabha and $ee \rightarrow \gamma\gamma$ events are partially (1%) recorded for the detector calibrations and luminosity monitor, their trigger rate is about 5 Hz at luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The total trigger rate is typically about 200 Hz, including background events from the e^+ and e^- beams. The background event rate depends on the beam current and condition. The DAQ efficiency is stable up to 500 Hz of trigger rate.

Each sub-detector provides the following information for the trigger:

- The CDC trigger provides the number of tracks from the IP and opening angle of tracks. The number of tracks is counted in the r - φ plane (only axial layers) and r - z plane (cathode layers and part of wire layers), individually.
- The TOF/TSC trigger provides the event timing with about 5 ns of time jitter, track multiplicity and topology. The TSC hit is necessary to remove background TOF hits from photon conversion inside the TOF.
- The ECL trigger provides the total ECL energy and the number of ECL clusters. It also provides the event timing with RMS timing resolution less than 20 ns if the TSC have no hit.
- The KLM trigger finds muon candidates. The trigger efficiency is about 98% except for the chimney hole of the superconducting solenoid and the sector gap in the forward endcap.

Figure 3.15 shows the flows of the trigger signals from sub-detectors to the GDL. Although the EFC is used for the trigger, I omit the explanation in this thesis. The major trigger criteria are:

- three or more tracks in the r - φ plane and at least one track in the r - z plane,
- high total ECL energy excluding Bhabha and cosmic trigger,
- four or more ECL clusters excluding cosmic trigger.

The loss of $\tau \rightarrow \ell V^0$ events after the selection in the chapter 4 are less than 1% by this trigger.

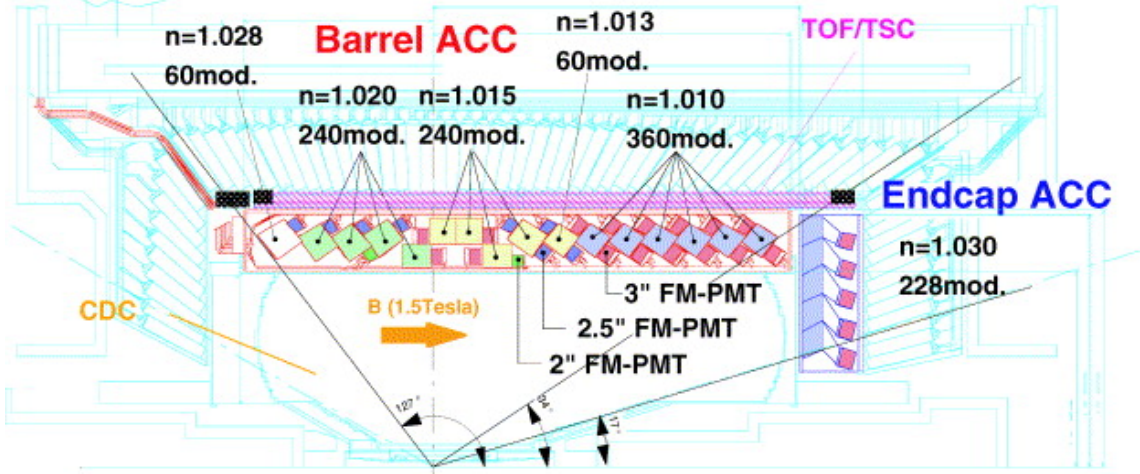


Figure 3.8: The arrangement of ACC at the central part of the Belle detector. Figure from ref. [51].

3.2.8 Data Acquisition (DAQ) system

The DAQ system collects the signals from the sub-detectors and sends combined data sets to the storage system. The readout system for the SVD records the analog pulse-height information by the fast analog-to-digital converter (ADC) [65, 66]. For the other detectors, a timing and charge of the signal is recorded by the unified readout system that is based on the charge-to-time (Q-to-T) converter and time-to-digital converter (TDC) [67]. The numbers of channels are about 82,000 for SVD, 45,000 for KLM, 10,000 for CDC, 9,000 for ECL, 1,800 for ACC, and 320 for TOF. A special time expansion module is used in the TOF readout for precise timing information.

An event builder on the online computer farm combines each sub-detector information and data sets are sent to the data storage system (figure 3.16). The trigger (TRG) and sequence controller (SEQ) information are also treated like sub-detectors. A few events are sent to the data quality monitor (DQM) before writing to the data storage. Each event size is 30–40 kbyte. The overall error rate of the DAQ flow is less than 0.05%.

Because the triggered raw-data set still contains much background events with tracks not from IP, an event selection is applied with an offline fast-track-finding algorithm before the event reconstruction. The events triggered by the CDC are required to have at least one good track that satisfy $p_t > 300$ MeV/c, $|dr| < 1.0$ cm, and $|dz| < 4.0$ cm. The background event reduction rate is 78% with the tau-pair event efficiency of 97.6%. From 2001, part of this good track selection was performed by an online fast-track-finding algorithm to suppress the raw data size.

The DAQ dead-time is depends on the trigger rate: about 5.5% at trigger rate 300 Hz (figure 3.17). In addition, the continuous injection scheme implemented from 2002 needs an injection veto time of 2 ms. The injection rate is 20 Hz and the dead-time due to the injection veto is 4%.

3.2.9 Particle identification (PID)

Particle identification (PID) is provided by CDC (de/dx), ACC (the number of photo-electrons, N_{pe}), TOF, ECL (cluster-track matching χ^2 , ratio between the measured energy in ECL and momentum, and shower shape) and KLM (shower shape). We calculate a likelihood for each particle and combine them by multiplication. If the detector is not used or does not cover the region, its likelihood is ignored. I use likelihood ratios for electron identification ($\mathcal{P}(e)$) [69] and muon identification ($\mathcal{P}(\mu)$) [70]. For hadron identification, I use a binary likelihood ratio $\mathcal{P}(i|j) = \mathcal{L}_i/(\mathcal{L}_i + \mathcal{L}_j)$, where $\mathcal{L}_{i(j)}$ is the likelihood of particle i (j) [55] and i (j) is π , K or p .

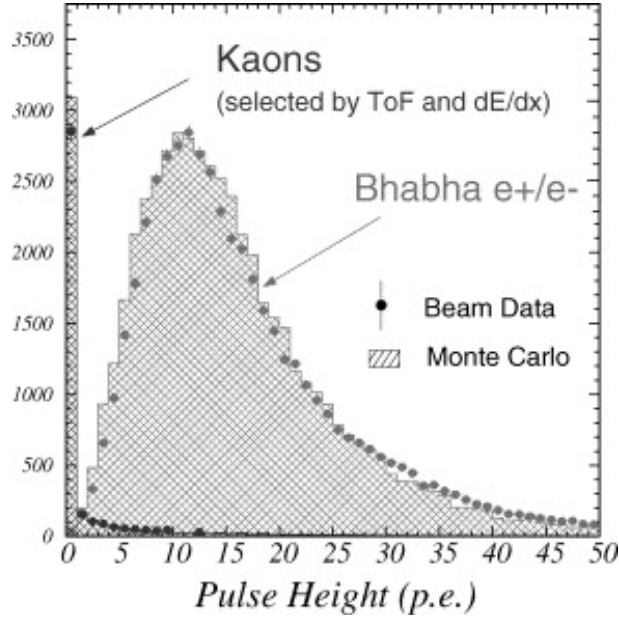


Figure 3.9: Pulse-height spectra (in units of photoelectrons) observed by the barrel ACC for electrons and kaons. Kaon candidates are obtained by dE/dx and TOF measurements. The Monte Carlo expectations are superimposed. Figure from ref. [58].

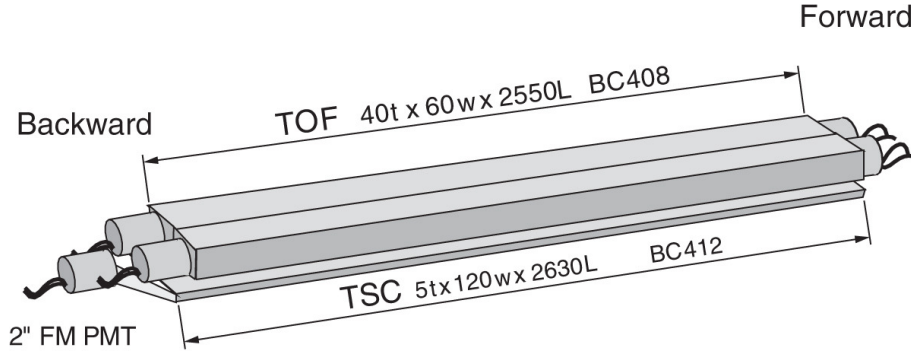


Figure 3.10: TOF/TSC module consists of two TOF counters and one TSC. BC408 and BC412 are the materials of the plastic scintillators. Figure from ref. [60].

The kaon identification is provided by CDC, ACC and TOF. The CDC and TOF likelihoods are defined assuming Gaussian distribution for dE/dx and TOF distribution. ACC likelihood is determined from the N_{pe} and its probability density function (PDF) for each particle. The kaon efficiency and pion fake rate are shown in figure 3.18.

The electron identification is provided by CDC, ACC and ECL. Three variables (cluster-track matching χ^2 , E/p and shower shape), which are combinations of CDC and ECL, are used for the electron identification. The cluster-track matching χ^2 is based on the distance between the ECL cluster and track extrapolated to the ECL region, their uncertainties are also considered. The E/p is the ratio between the measured energy in ECL and the measured momentum by CDC. Because an electron produces a electromagnetic shower in ECL, it has large E/p . The shower shape is the ratio of energy deposit in 3×3 CsI(Tl) crystals and 5×5 crystals around the peak energy deposit. This is parameterization of lateral shape of the shower. The PDFs of these three variables are defined by fitting a function to the actual distribution. The CDC dE/dx likelihood and ACC N_{pe} likelihood are also used for the electron identification. Figure 3.19 shows the electron efficiency and figure 3.20 shows the pion fake rate with the $\mathcal{P}(e)$ more than 0.5.

The muon identification is provided by two variables from KLM. A range difference (ΔR) is

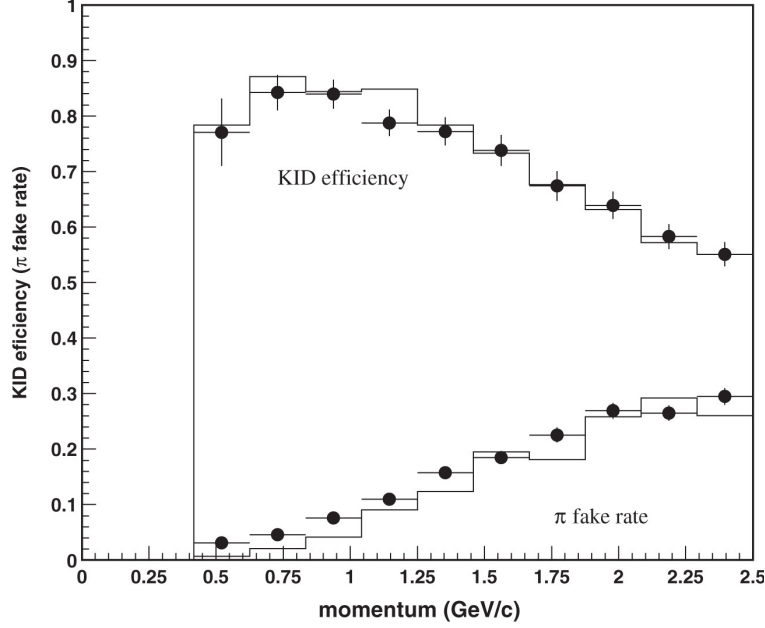


Figure 3.11: K identification (KID) efficiency and π fake rate as a function of momentum. The circles are data; the histogram is MC. Figure from ref. [60].

the difference between the track length measured by KLM and that of expected for each particle. A hit finding χ^2 (χ_r^2) is reduced χ^2 of track fitting within KLM volume. These PDFs are defined for a muon, pion and kaon because an electron stops at the ECL. The muon efficiency and pion fake rate with the $\mathcal{P}(e)$ more than 0.1 (0.9) are shown in figure 3.21 and figure 3.22, respectively.

3.3 Data set

Of the 980 fb^{-1} data set, 703 fb^{-1} was collected at the $\Upsilon(4S)$ resonance, 121 fb^{-1} at the $\Upsilon(5S)$, 89 fb^{-1} at an energy 60 MeV below the $\Upsilon(4S)$, 28 fb^{-1} of energy scans above the $\Upsilon(4S)$, and the remainder at and near the $\Upsilon(1-3S)$. Compared to the previous paper, the following data sets have been added: 78 fb^{-1} at and near the $\Upsilon(5S)$, 38 fb^{-1} at and near the $\Upsilon(1-3S)$, and 10 fb^{-1} at an energy 60 MeV below the $\Upsilon(4S)$. Experiment numbers are assigned to the Belle data, which is incremented in a few months. The SVD1 period is from 7 to 27 in the experiment number, and the SVD2 period is from 31 to 73.

Signal events of $\tau \rightarrow \ell V^0$ are generated by a dedicated Monte Carlo (MC) simulation with KKMC and TAUOLA [71], where $\tau^+\tau^-$ pairs are initially produced and one of the τ 's decays into ℓV^0 and the other decays generically. The number of generated signal MC events are 1.1×10^6 events at the $\Upsilon(4S)$ resonance, 0.4×10^6 events at the $\Upsilon(5S)$, 0.1×10^6 events at each of the $\Upsilon(1-3S)$ and 0.1×10^6 events at the energy 60 MeV below the $\Upsilon(4S)$. We assume an uniform CLFV decay angle in the τ rest frame. No specific NP model is assumed in the CLFV decay process, and the spin direction of V^0 is set randomly and independently of the spin of the mother τ .

For background MC simulations, $e^+e^- \rightarrow q\bar{q}$ ($q = u, d, s, c$; 5720 fb^{-1}), $e^+e^- \rightarrow \tau^+\tau^-$ (8300 fb^{-1}), Bhabha (236 fb^{-1}) and two-photon processes ($\mu\mu$, $eeee$, $ee\mu\mu$ and $eehh$; $6640\text{--}8300 \text{ fb}^{-1}$) are generated by EvtGen [72], KKMC [71], BHLUMI [73], and AAFH [74], respectively. The detector responses are simulated by GEANT3 [75]. The signal and background MC samples are generated with the accelerator and detector conditions at the experiment number.

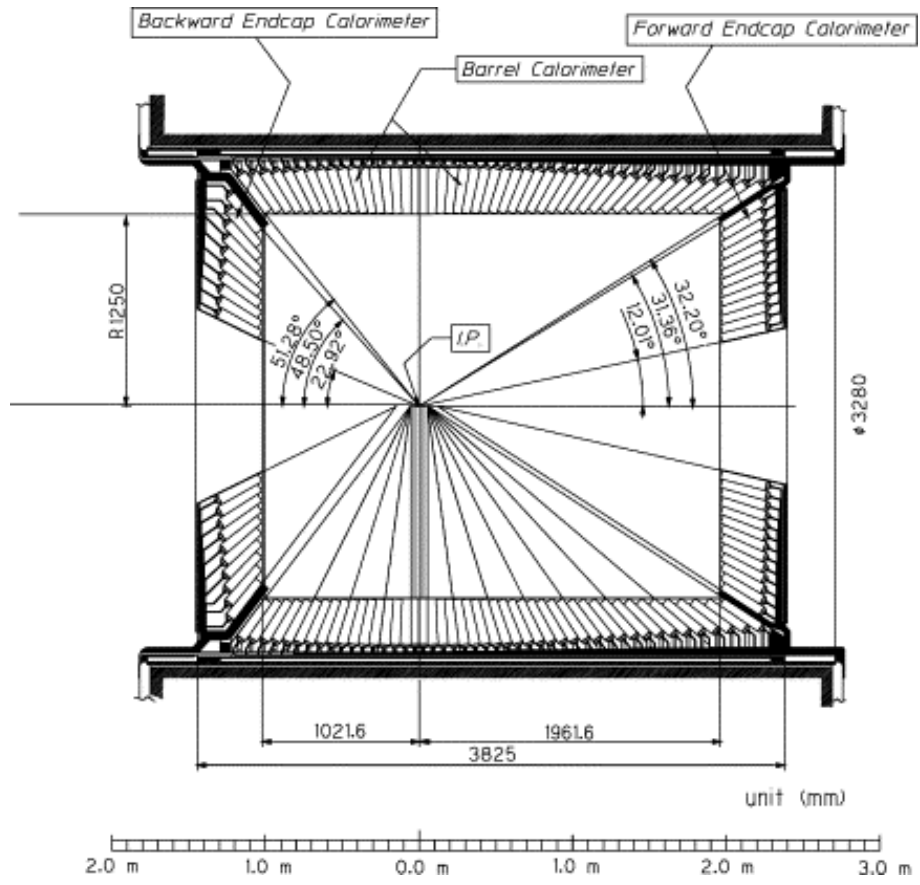


Figure 3.12: Overall structure of Belle electromagnetic calorimeter (side view). Figure from ref. [61].

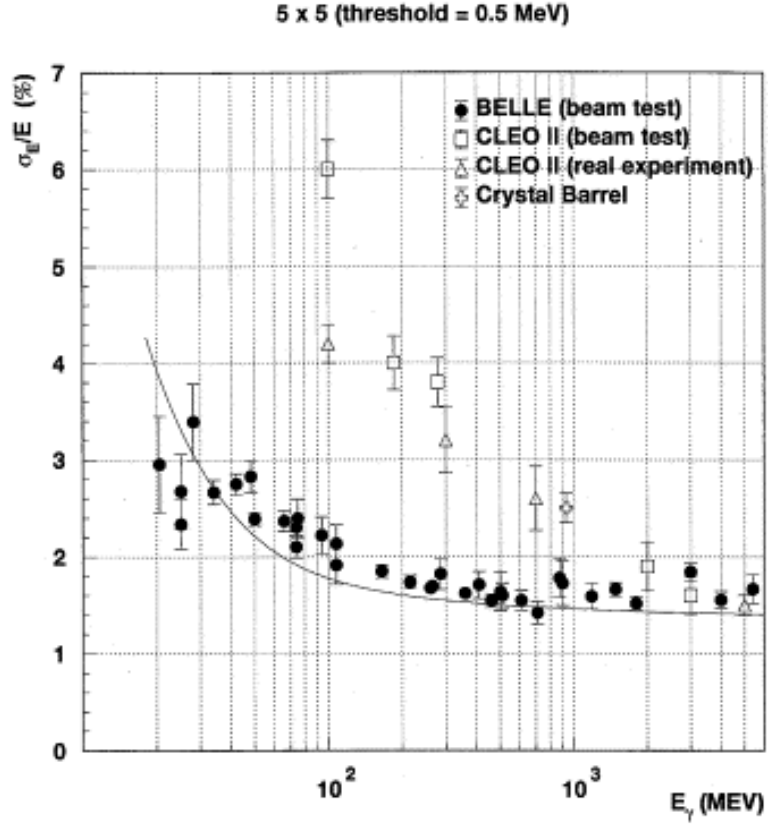


Figure 3.13: The energy resolution of the 5×5 matrix crystals with a threshold of 0.5 MeV in the photon beam test. The solid line is from the Monte-Carlo simulation. CLEO II beam test results were obtained for a 3×3 matrix. Figure from ref. [62].

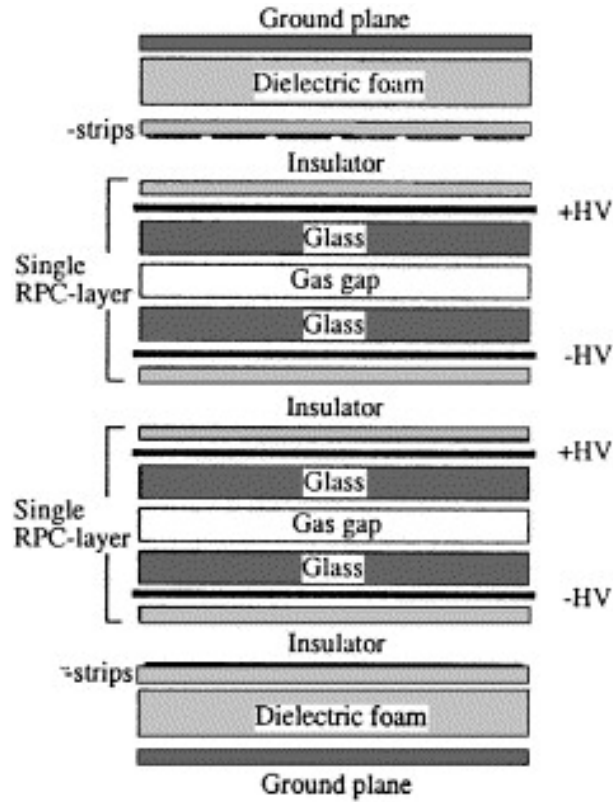


Figure 3.14: Cross-section of a KLM superlayer. Figure from ref. [63].

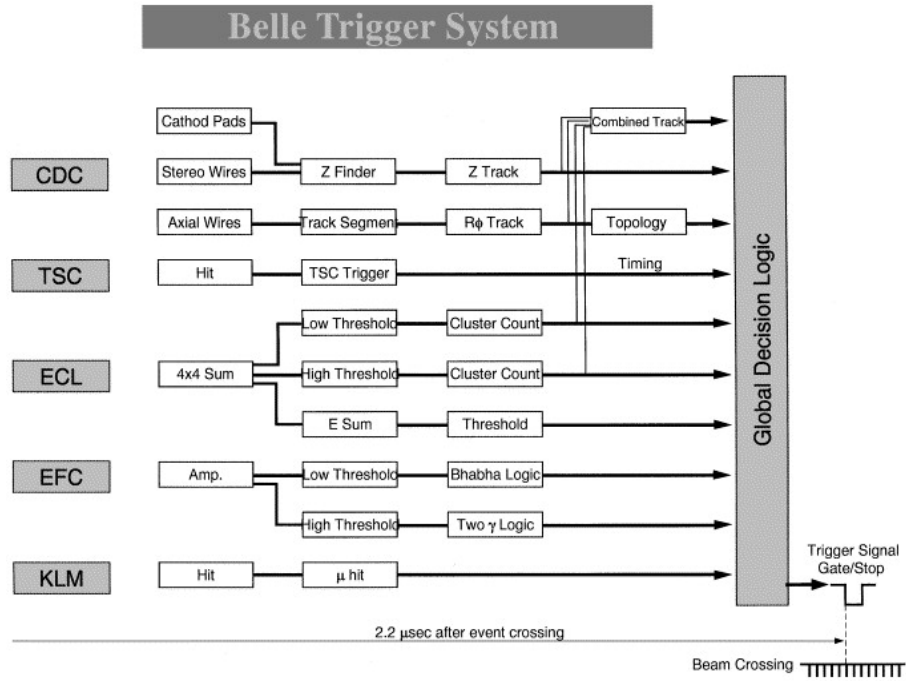


Figure 3.15: The hardware trigger system for the Belle detector. Figure from ref. [64].

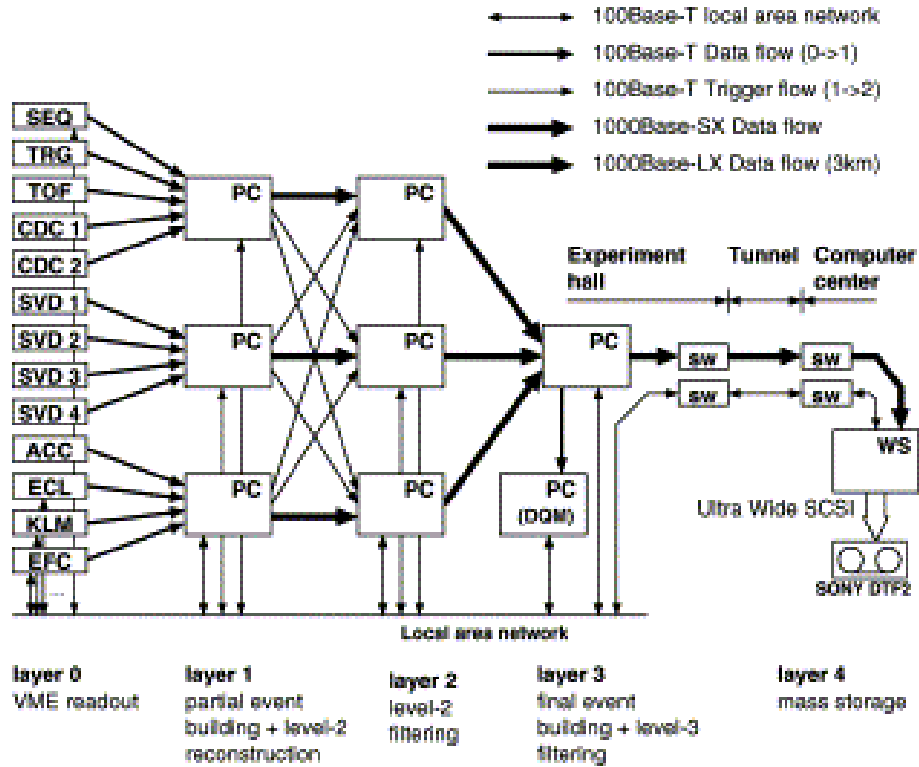


Figure 3.16: Structure of an event building farm and the DAQ data flow. Figure from ref. [68].

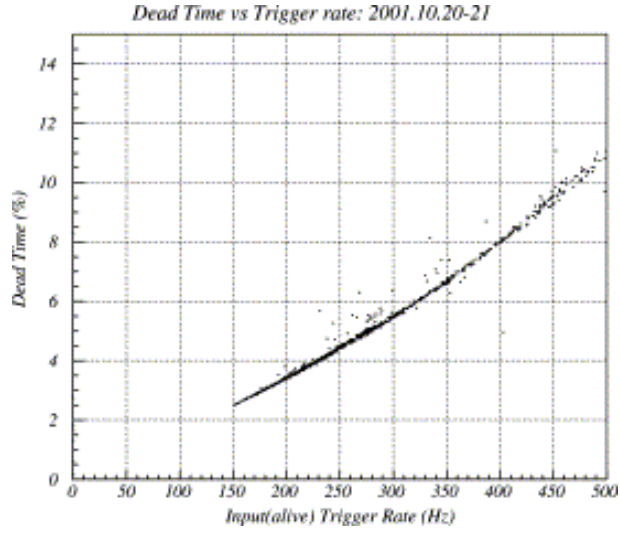


Figure 3.17: DAQ dead-time of beam run. Figure from ref. [68].

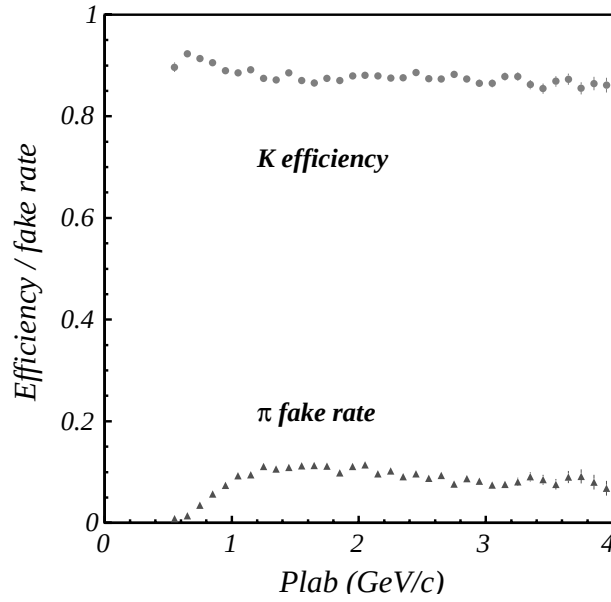


Figure 3.18: Momentum dependence of kaon efficiency and pion fake rate with $\mathcal{P}(K|\pi) > 0.6$. Upper point is kaon efficiency and lower point is pion fake rate. Figure from ref. [55].

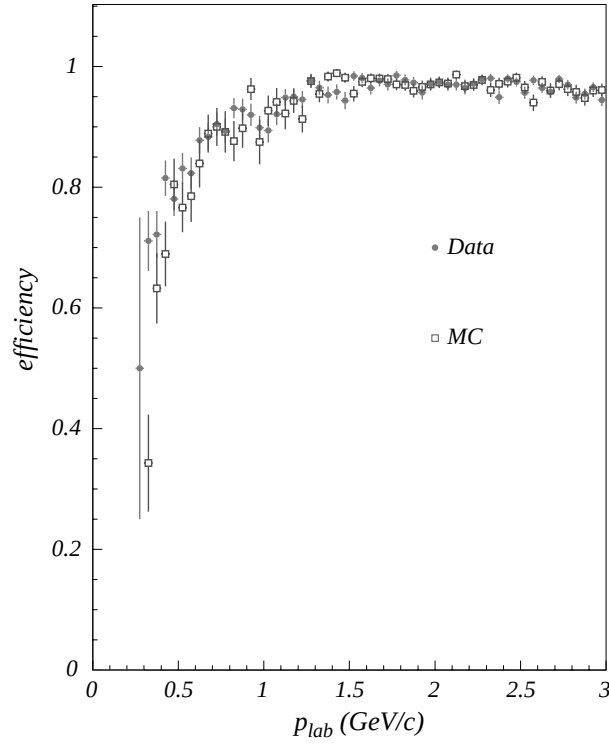


Figure 3.19: Electron efficiencies of experimental (diamond) and Monte-Carlo simulation (squares). Figure from ref. [55].

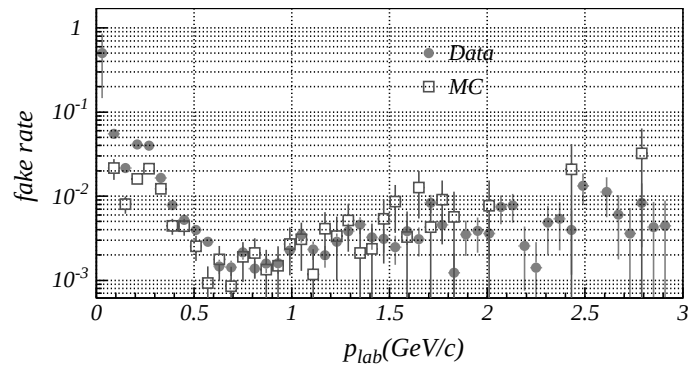


Figure 3.20: Pion fake rates of experimental (circles) and Monte-Carlo simulation (squares). Figure from ref. [55].

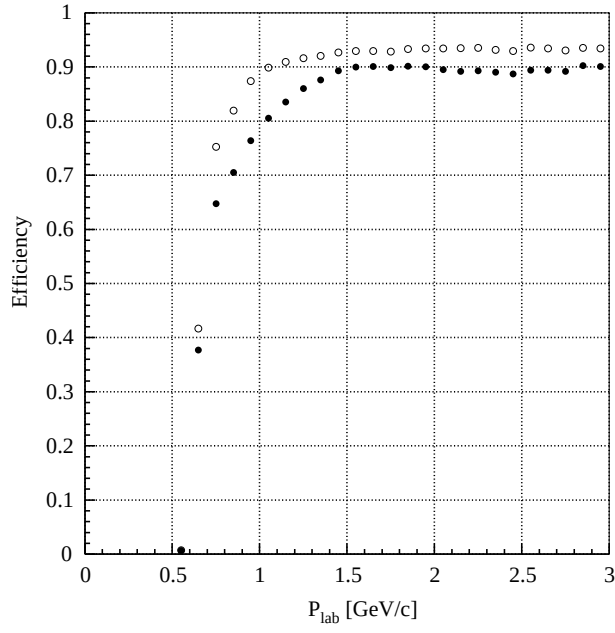


Figure 3.21: Momentum dependence of muon efficiency for loose cut (open circle) and tight cut (filled circle). Figure from ref. [55].

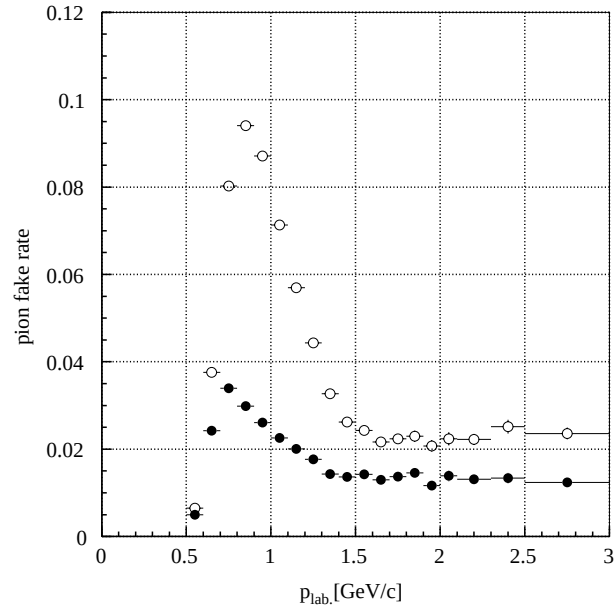


Figure 3.22: Momentum dependence of pion fake rate for loose cut (open circle) and tight cut (filled circle). Figure from ref. [55].

Chapter 4

Reconstruction and event selection

Calculated from the branching fraction upper limits and the signal efficiency in the previous paper [18], the expected number of $\tau \rightarrow \ell V^0$ events is at most five. In contrast, the number of background events is $\mathcal{O}(100\text{--}1000)$ even with basic requirements of the V^0 mass, particle identifications, etc. I aim to improve the signal efficiency, while suppressing the number of background events less than one. The $\tau \rightarrow \ell V^0$ events are reconstructed with their features in section 4.1. Part of background events are rejected in section 4.2 I establish the cut-base event selection criteria first in section 4.3. In section 4.4, I introduce a multivariate analysis for further improvement.

4.1 Reconstruction

A signal τ is reconstructed from a charged lepton (ℓ) and a neutral vector meson (V^0) in an event of $e^+e^- \rightarrow \tau^+\tau^-$. The τ^+ and τ^- go in opposite directions with a momentum of about 5.0 GeV/c, and the particles from each τ decay are well separated in the detector. I separate the event into two hemispheres in the center-of-mass (c.m.) frame by a plane perpendicular to the thrust vector ($\vec{n}_T$) [76, 77]. The thrust vector is obtained by maximizing the thrust $T = \sum_i |\vec{p}_i^{\text{c.m.}} \cdot \vec{n}_T| / \sum_i |\vec{p}_i^{\text{c.m.}}|$, where i runs over all tracks and photons, and $\vec{p}_i^{\text{c.m.}}$ is the momentum in the c.m. frame. In either of the hemispheres that contains a τ CLFV decay (called signal side), V^0 is reconstructed as follows: ρ^0 from $\pi^+\pi^-$, ϕ from K^+K^- , ω from $\pi^+\pi^-\pi^0$, K^{*0} from $K^+\pi^-$ and $\bar{K}^{*0}$ from $K^-\pi^+$. These V^0 reconstructions cover 50–100% of the V^0 decay branching fractions. In the other hemisphere (called tag side), the other τ (τ_{tag}) is reconstructed from $\ell^\pm \nu \bar{\nu}$, $\pi^\pm \nu$, $\pi^\pm \pi^0 \nu$, $\pi^\pm \pi^+ \pi^- \nu$ or $\pi^\pm \pi^0 \pi^0 \nu$. The $\pi^\pm \pi^+ \pi^- \nu$ and $\pi^\pm \pi^0 \pi^0 \nu$ modes are additional tag reconstructions from the previous paper. The total branching fractions of these τ_{tag} decay is about 90%. Because the neutrinos cannot be detected, I require missing momentum in the tag side. This τ_{tag} information enables the suppression of background events that have no neutrinos.

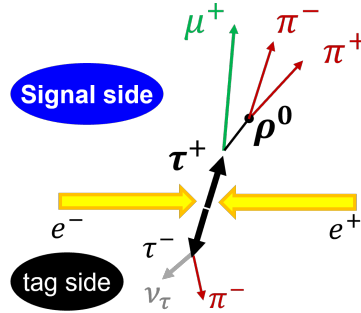


Figure 4.1: The $\tau \rightarrow \mu \rho^0$ and τ_{tag} decays.

The signal $\tau \rightarrow \ell V^0$ events have a unique kinematical feature; the ℓV^0 invariant mass ($M_{\ell V^0}$)

is close to the τ mass and the difference of the ℓV^0 energy from the beam energy in the c.m. frame (ΔE) is close to zero. The e^+e^- collider has advantage of the fixed τ energy, and this ΔE becomes powerful to suppress background events. The signal events within $1.65 \text{ GeV}/c^2 < M_{\ell V^0} < 1.90 \text{ GeV}/c^2$ and $|\Delta E| < 0.5 \text{ GeV}$ are reconstructed (figures 4.2 and 4.3 show the distributions of the signal MC events).

I follow a blind analysis approach in this search until finalizing the event selection and background estimation. In this approach, signal candidates in the data are hidden to avoid making bias through the event selection. The blind region is $1.75 \text{ GeV}/c^2 \leq M_{\ell V^0} < 1.81 \text{ GeV}/c^2$ and $|\Delta E| < 0.08 \text{ GeV}$ for the $\mu\rho^0$, $\mu\phi$ and $\mu K^{*0}(\bar{K}^{*0})$ modes, and $1.74 \text{ GeV}/c^2 \leq M_{\ell V^0} < 1.82 \text{ GeV}/c^2$ and $|\Delta E| < 0.1 \text{ GeV}$ for the other modes.

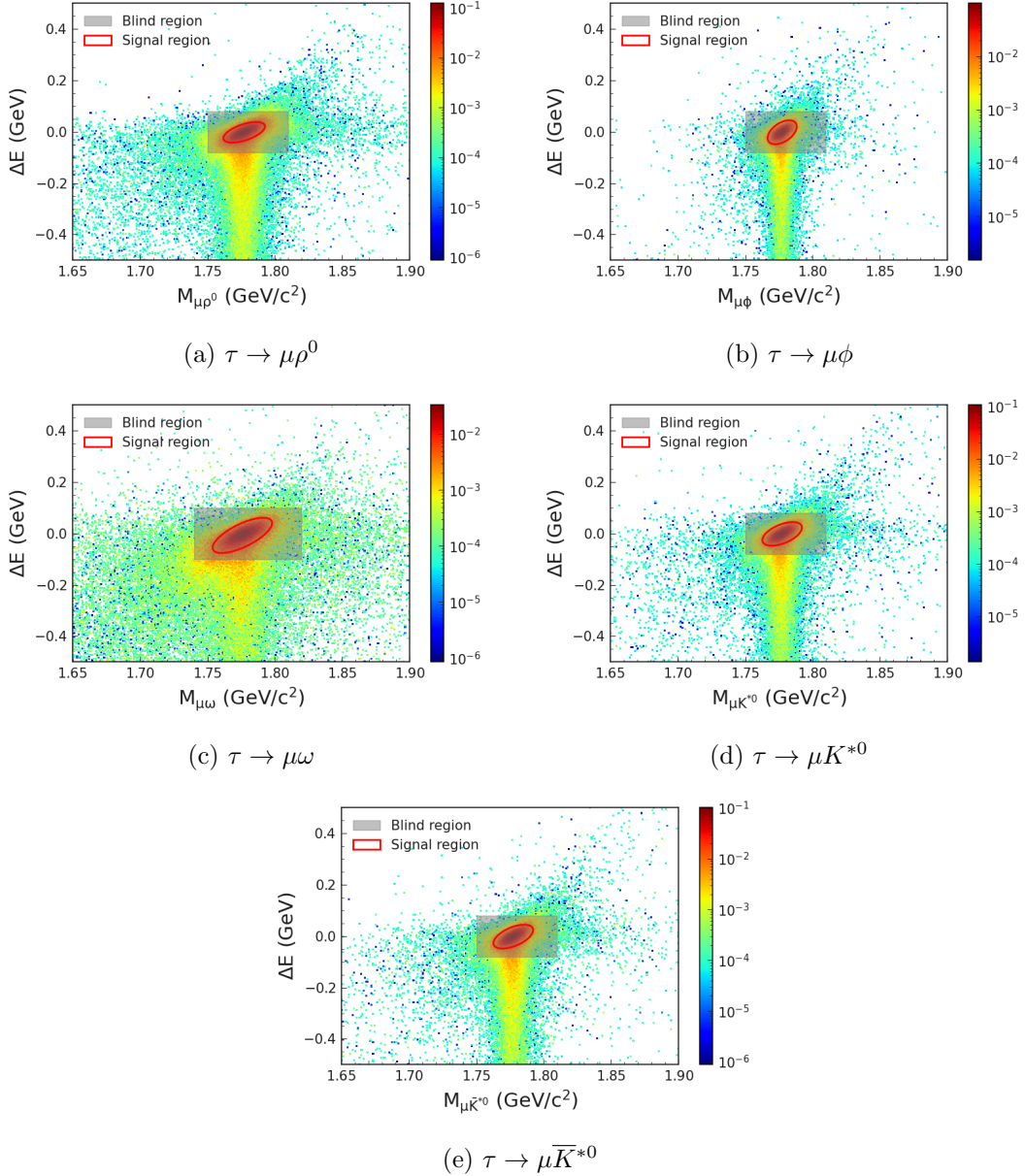


Figure 4.2: Reconstructed $M_{\ell V^0}$ and ΔE of the μV^0 signal MC events.

The gray square is the blind region for the data. The red ellipse line is the signal region. The tail events of ΔE below -0.1 GeV are e^+e^- collisions with some energy losses by bremsstrahlung before colliding.

Charged tracks, photons and π^0 s should satisfy the following selections. Each charged track

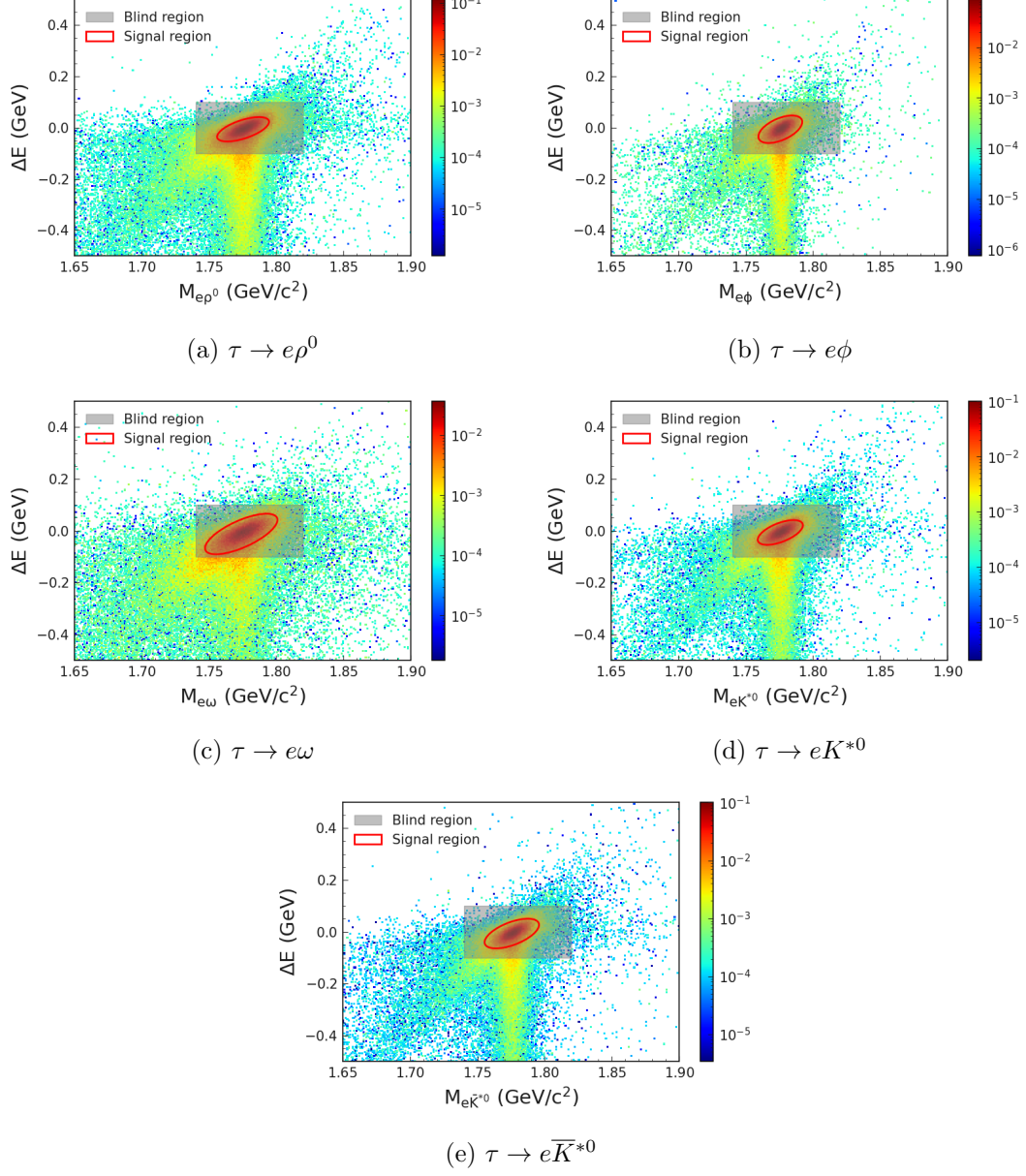


Figure 4.3: Reconstructed $M_{\ell V^0}$ and ΔE of the $e V^0$ signal MC events.

The gray square is the blind region for the data. The red ellipse line is the signal region. The tail events of ΔE below -0.1 GeV are e^+e^- collisions with some energy losses by bremsstrahlung before colliding. The events in the region of lower $M_{\ell V^0}$ and lower ΔE lose bremsstrahlung photons mainly from electrons or positrons.

or photon is within the Belle detector's fiducial volume defined by $-0.866 < \cos \theta < 0.956$, where θ is the polar angle with respect to the direction opposite to the e^+ beam in the laboratory frame. Charged tracks come from the e^+e^- beam interaction point; the distance of the closest point from the interaction point is less than 0.5 cm in the transverse direction and less than 3.0 cm in the longitudinal direction. Each π^0 is reconstructed from two photons inside the same hemisphere and the photon energy (E_γ) should be larger than 0.05 GeV. The π^0 mass window is $0.12 \text{ GeV}/c^2 < M_{\gamma\gamma} < 0.15 \text{ GeV}/c^2$. At the π^0 reconstruction, the daughter photon energies are corrected by assuming that the π^0 mass is $0.135 \text{ GeV}/c^2$, it improves the π^0 energy resolution.

After reconstructing the signal and tag τ 's, no extra charged tracks are allowed. I count the number of photons (n_γ) with E_γ larger than 0.1 GeV, and reject the events with n_γ more than seven. The n_γ in the signal side is required to be less than four for the $\ell\omega$ mode, which includes a $\pi^0 \rightarrow \gamma\gamma$, and less than two for the other modes.

I require features of a generic τ_{tag} decay. The invariant mass of the particles including all photons in the tag hemisphere should be less than the τ mass ($1.777 \text{ GeV}/c^2$). For τ_{tag} decays into $\pi\pi^0\nu$ ($\pi^\pm\pi^+\pi^-\nu$ or $\pi^\pm\pi^0\pi^0\nu$), the reconstructed mass of those pions is required to be $0.4 \text{ GeV}/c^2 < M_{\pi\pi^0} < 1.3 \text{ GeV}/c^2$ ($0.7 \text{ GeV}/c^2 < M_{3\pi} < 1.7 \text{ GeV}/c^2$), which corresponds to the mass of $\rho^\pm$ ($a_1^\pm$).

Particle identification is effective to reduce main background events of three hadron tracks. The lepton identification criteria are $\mathcal{P}(e) > 0.9$ for electrons, and $\mathcal{P}(\mu) > 0.95$ and the momentum is larger than $0.6 \text{ GeV}/c$ for muons. The electron (muon) identification efficiency is 90% (75%), whereas the probability of misidentifying a pion as an electron (muon) is 0.1% (2%). Electrons and positrons sometimes emit bremsstrahlung photons while traveling. If daughter particles from τ decays have bremsstrahlung photon, part of their energy is lost. I sum up the energy of every photon within 0.05 radians from the signal side electron track direction and $E_\gamma \geq 0.05 \text{ GeV}$ into the electron momentum. To suppress low multiplicity background events (e.g., Bhabha, $ee \rightarrow eeee$, and $ee\mu\mu$), an electron veto ($\mathcal{P}(e) < 0.9$) is applied to all hadron tracks.

The kaon identification criteria are $\mathcal{P}(K|\pi) > 0.6$ for both charged kaons from ϕ decay and $\mathcal{P}(K|\pi) > 0.8$ for the charged kaon from K^{*0} and $\bar{K}^{*0}$ decays. The kaon identification efficiency is 86% (77%), whereas the probability of misidentifying a charged pion as a kaon is 4% (2%) for the kaons from ϕ (K^{*0} , $\bar{K}^{*0}$). A kaon veto ($\mathcal{P}(K|\pi) < 0.6$) is applied to both charged pions from ρ^0 in the signal side, and 96% of pions are retained, whereas 14% of kaons are not vetoed. To suppress muons from kaon decays inside the CDC ($K^\pm \rightarrow \mu^\pm\nu$), the kaon veto is also applied to the signal side muon track for the hadronic tags ($\tau_{\text{tag}} \rightarrow \pi\nu$, $\pi\pi^0\nu$, $\pi\pi\pi\nu$ or $\pi\pi^0\pi^0\nu$). For the μV^0 modes with the hadronic tags, a proton veto ($\mathcal{P}(p|K) < 0.6$ and $\mathcal{P}(p|\pi) < 0.6$) is applied for the tag side tracks.

The signal event has one or two neutrinos from the τ_{tag} decay. The missing momentum due to the neutrino(s) is calculated by subtracting the vector sum of the momenta of all tracks and photons from the sum of the beam momenta. The transverse missing momentum is required to be larger than $0.5 \text{ GeV}/c$. The missing particles should be on the tag side and hence $\cos \theta_{\text{miss-tag}}^{\text{c.m.}} > 0$, where $\theta_{\text{miss-tag}}^{\text{c.m.}}$ is the angle between the missing momentum and the vector sum of tag side tracks and photons in the c.m. frame. The neutrino angle with respect to the τ_{tag} momentum direction is restricted in a τ_{tag} two body decay; thus $\cos \theta_{\text{miss-tag}}^{\text{c.m.}} < 0.85$ is also applied for the $\ell\rho^0$ modes with $\tau_{\text{tag}} \rightarrow \pi\nu$.

Events with missing particles other than neutrinos should be rejected as background events. These non-neutrino missing particles can arise in two ways: neutral particles pass through the gaps between the barrel and endcap ECLs, and any particles travel near the beam pipes where the particles are not detected by Belle. Thus, the direction of the missing momentum is required not to point to such regions.

After the event reconstruction above, remaining background events are from $q\bar{q}$ ($q = u, d, s, c$), $\tau^+\tau^-$ with at least one of τ decays to $\pi^\pm\pi^+\pi^-\nu$, and low multiplicity events. First, the low multiplicity events are rejected in section 4.2. The other $q\bar{q}$ and $\tau^+\tau^-$ background events are

mainly three charged pion with lepton mis-identification. I introduce a new event selection variable, which is neutrino mass squared (M_ν^2), to suppress the background events (section 4.3).

4.2 Low multiplicity event suppression

The low multiplicity events remain in the events with electrons: $\tau \rightarrow eV^0$ or $\tau_{\text{tag}} \rightarrow e\nu\bar{\nu}$. Because the missing particles of these events are the bremsstrahlung photons from electrons or positrons in the tag side, $\cos\theta_{\text{miss-tag}}^{\text{c.m.}}$ is close to one. Such bremsstrahlung photons can be missed when their ECL hits overlap with that of $e^\pm$ tracks. The missing energy in the c.m. frame ($E_{\text{miss}}^{\text{c.m.}}$) is also used to reduce the low multiplicity events, which do not have any missing particles actually. For all ℓV^0 modes, $E_{\text{miss}}^{\text{c.m.}}$ is required to be larger than 0.0 GeV (figure 4.4). For the $\mu\rho^0$ mode with $\tau_{\text{tag}} \rightarrow e\nu\bar{\nu}$, $\cos\theta_{\text{miss-tag}}^{\text{c.m.}} < 0.99$ (figure 4.5) and $E_{\text{miss}}^{\text{c.m.}} > 0.4$ GeV selections are applied. For the eV^0 modes with $\tau_{\text{tag}} \rightarrow e\nu\bar{\nu}$ or $\pi\nu$, $\cos\theta_{\text{miss-tag}}^{\text{c.m.}} < 0.97$ is applied. For the eV^0 modes with $\tau_{\text{tag}} \rightarrow e\nu\bar{\nu}$, $E_{\text{miss}}^{\text{c.m.}}$ should be larger than 0.4, 2.0, and 1.5 GeV for $e\phi$, $e\rho^0$ and the other eV^0 modes, respectively.

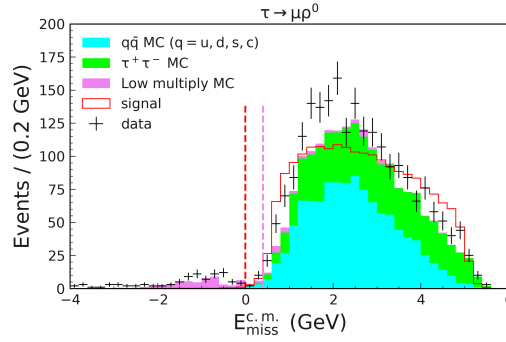


Figure 4.4: Distribution of $E_{\text{miss}}^{\text{c.m.}}$ in the $\tau \rightarrow \mu\rho^0$ mode. The events are in $1.65 < M_{\ell V^0} < 1.9$ GeV/c^2 and $|\Delta E| < 0.5$ GeV excluding the blind region. The violet dotted line is the lower limit for electron tag. The red dashed line is the lower limit for the other tags. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

A photon sometimes converts to an electron-positron pair in the beam pipe wall or in the SVD materials, and they are detected as two charged tracks in CDC. If a single-track background event is accompanied by such photon conversion, it can create a large amount of background events especially for the eV^0 modes. To suppress these events, the invariant mass of the electron and one of the tracks from the V^0 , assigned the electron-mass hypothesis ($M_{ee}(e^-\pi^+)$ or $M_{ee}(e^-K^+)$), should be larger than 0.2 GeV/c^2 (figure 4.6, left). In addition, for the eK^{*0} and $e\bar{K}^{*0}$ modes, the invariant mass of the two tracks from the V^0 , each assigned the electron-mass hypothesis ($M_{ee}(K\pi)$), is required to be larger than 0.1 GeV/c^2 (figure 4.6, right).

Most of the low multiplicity events are suppressed and the remaining background events are from $q\bar{q}$ and $\tau^+\tau^-$ events. I establish two event selection criteria to reduce these background events. The first one (section 4.3) is cut-base selection, which specifies the event selections for each variable. For the second one (section 4.4), I use a multivariate analysis library and get further improvement.

4.3 Cut-base selection

I make new variables of neutrino kinematics for the event selection: neutrino mass squared (M_ν^2) and neutrino momentum in the c.m. frame ($P_\nu^{\text{c.m.}}$). They are calculated from the momenta of

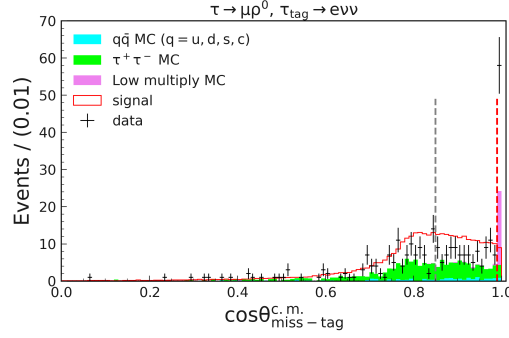


Figure 4.5: Distribution of $\cos\theta_{\text{miss-tag}}^{\text{c.m.}}$ in the $\tau \rightarrow \mu\rho^0, \tau_{\text{tag}} \rightarrow e\nu\bar{\nu}$ mode. The events are in $1.65 < M_{\ell V^0} < 1.9 \text{ GeV}/c^2$ and $|\Delta E| < 0.5 \text{ GeV}$ excluding the blind region. The $E_{\text{miss}}^{\text{c.m.}}$ selection is applied. The gray dotted line is the old upper limit that was applied for any tag side particles in the previous paper. The red dashed line is the new upper limit for the electron tag. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

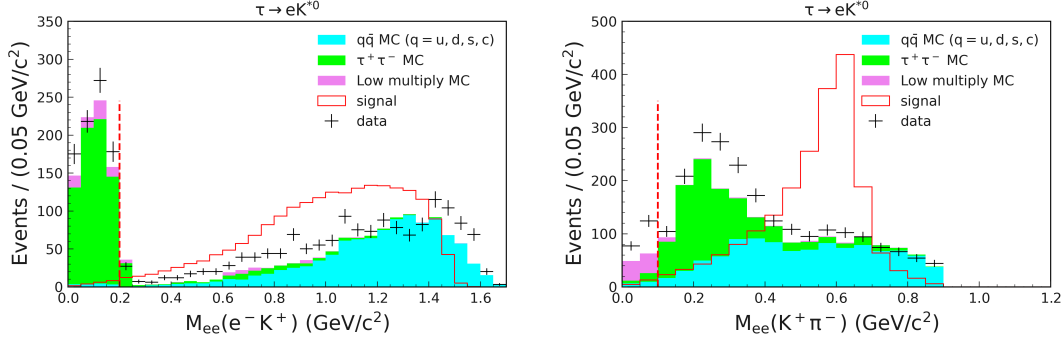


Figure 4.6: $M_{ee}(e^-K^+)$ (left) and $M_{ee}(K^+\pi^-)$ (right) of the eK^{*0} mode for the events of $1.65 < M_{\ell V^0} < 1.9 \text{ GeV}/c^2$ and $|\Delta E| < 0.5 \text{ GeV}$ excluding the blind region. The $E_{\text{miss}}^{\text{c.m.}}$ and $\cos\theta_{\text{miss-tag}}$ selections are applied. The red dashed lines are the lower limits. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

the reconstructed τ_{signal} and τ_{tag} , where the energy of τ_{signal} is fixed to the half of the beam energy in the c.m. frame. The $q\bar{q}$ background events can be effectively suppressed by a M_ν^2 selection in the hadronic tags, involving only one neutrino (figure 4.7).

The following event selections are only for the cut-base selection, they are not applied in the selection in section 4.4. The V^0 mass windows for the cut-base selection are shown in figures 4.8, 4.9 and table 4.1. The lower energy of the two photon from the signal side π^0 decay (E_γ^{lower}) is required to be larger than 0.1 (0.07) GeV for the $\mu(e)\omega$ mode. The M_ν^2 selections are shown in table 4.2. I do not require M_ν^2 selection for the $\ell\phi$, eK^{*0} and $e\bar{K}^{*0}$ modes because the number of background events is enough small.

For the $\ell\rho^0$, $\ell\phi$ and ℓK^{*0} modes, the number of events in the sideband region is less in the background MC than in the data. That may be due to the difficulty of hadronization in the MC generator. In addition, $ee \rightarrow DD$ events are not generated in the $c\bar{c}$ MC. Such events can be background events especially for the $\ell\phi$ and ℓK^{*0} modes.

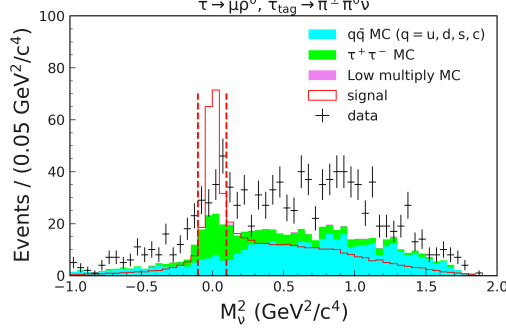


Figure 4.7: M_V^2 of the $\tau \rightarrow \mu\rho^0, \tau_{tag} \rightarrow \pi\pi^0\nu$ mode for the events of $1.65 < M_{\ell V^0} < 1.9 \text{ GeV}/c^2$ and $|\Delta E| < 0.5 \text{ GeV}$ excluding the blind region. Events between the two red dashed lines are inside the cut-base selection. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events. The tail events of the signal MC events originate from wrong π^0 or contamination from $\tau_{tag} \rightarrow \pi\pi^0\pi^0\nu$.

Table 4.1: The V^0 mass window for the cut-base selection. 3σ region for ϕ , 2σ region for the other V^0 s.

V^0	Invariant mass (GeV/c^2)
ρ^0	$0.587 < M_{\pi\pi} < 0.962$
ϕ	$1.009 < M_{KK} < 1.031$
ω	$0.757 < M_{\pi\pi\pi^0} < 0.808$
$K^{*0}(\bar{K}^{*0})$	$0.842 < M_{K\pi} < 0.956$

4.4 Selection with Boosted Decision Tree

For further improvement from the cut-base selection, I use a two-class Boosted Decision Tree (BDT) for signal and background classification. The BDT is a multivariate analysis tool, which estimate a signal probability from multiple input variables. I use LightGBM for the BDT library [78]. The BDT needs training to learn the difference between the signal and background events. Section 4.4.1 is about control samples for the training purpose.

Some event selections are loosened from the cut-base selection before the BDT training and selection. The V^0 mass windows for BDT are in table 4.3, and M_V^2 selections in table 4.2 are not applied. The selection of $E_{\gamma}^{\text{lower}}$ is also loosen to be larger than 0.05 GeV.

4.4.1 Control samples of background events

The BDT learns not only the event features but also the statistical fluctuation in the training data set. The performance to separate signal from background events usually becomes worse for unknown data sets than for the training data set. That is why the BDT training samples should be independent of the $\tau \rightarrow \ell V^0$ analysis data sets. Otherwise, the number of background events after the BDT selection can be under-estimated.

The training samples must have similar feature with that of $\tau \rightarrow \ell V^0$ analysis data sets. The larger the number of training samples, the better the performance of the BDT. I focus on the source of the background events; most of them are from the $q\bar{q}$ and $\tau^+\tau^-$ events that contain three charged hadron in the signal side. Such events are suppressed by the variable of a lepton probability in the $\tau \rightarrow \ell V^0$ analysis data sets. In other words, I have large number of three charged hadron events, which have smaller lepton probabilities than that of $\tau \rightarrow \ell V^0$ analysis

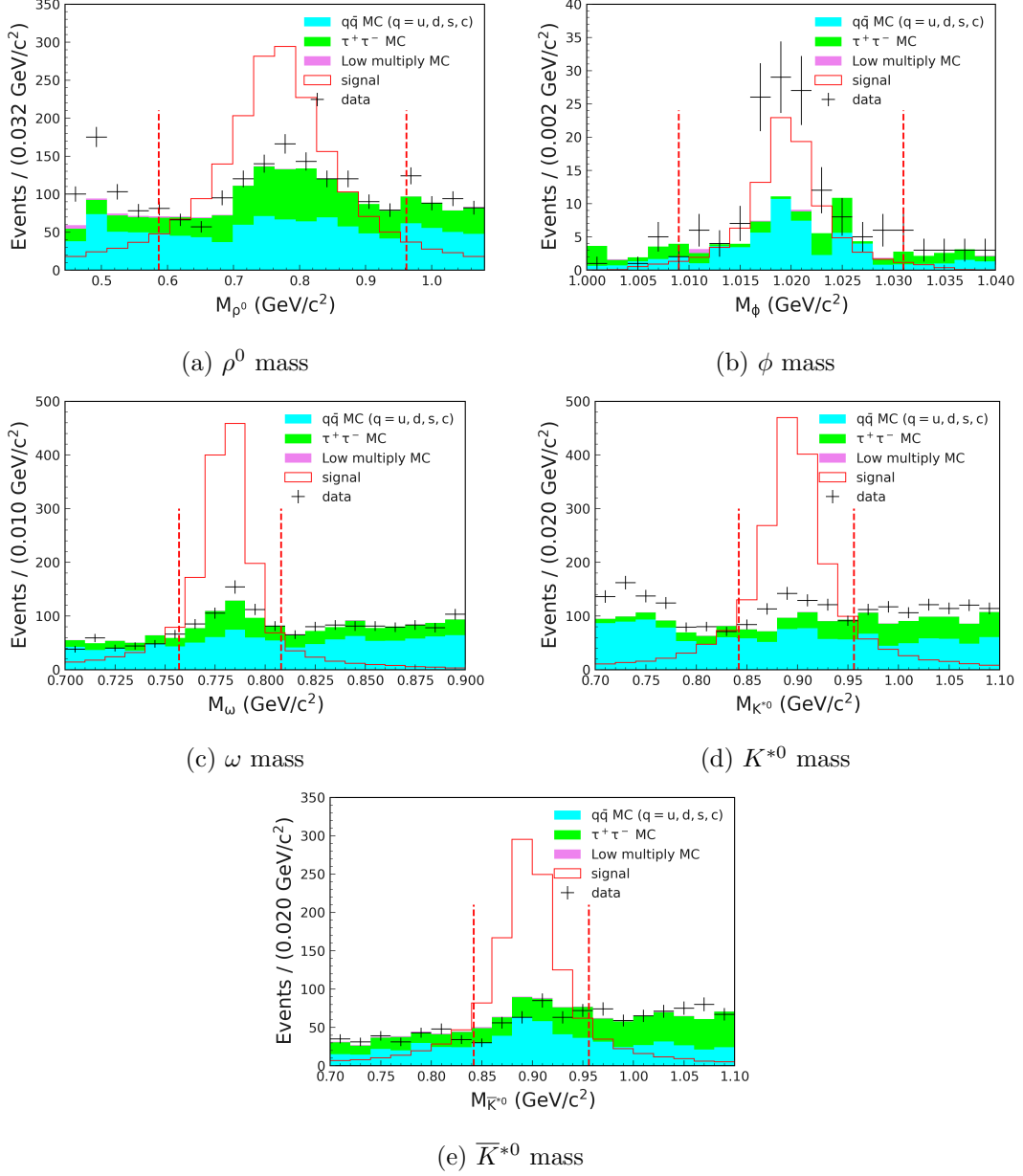


Figure 4.8: Invariant mass distribution of the V^0 for the μV^0 modes. The data and background MC events are in the sideband region where $1.65 < M_{\ell V^0} < 1.9 \text{ GeV}/c^2$ and $|\Delta E| < 0.5 \text{ GeV}$ excluding the blind region. Events between the two red dashed lines are inside the cut-base selection. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

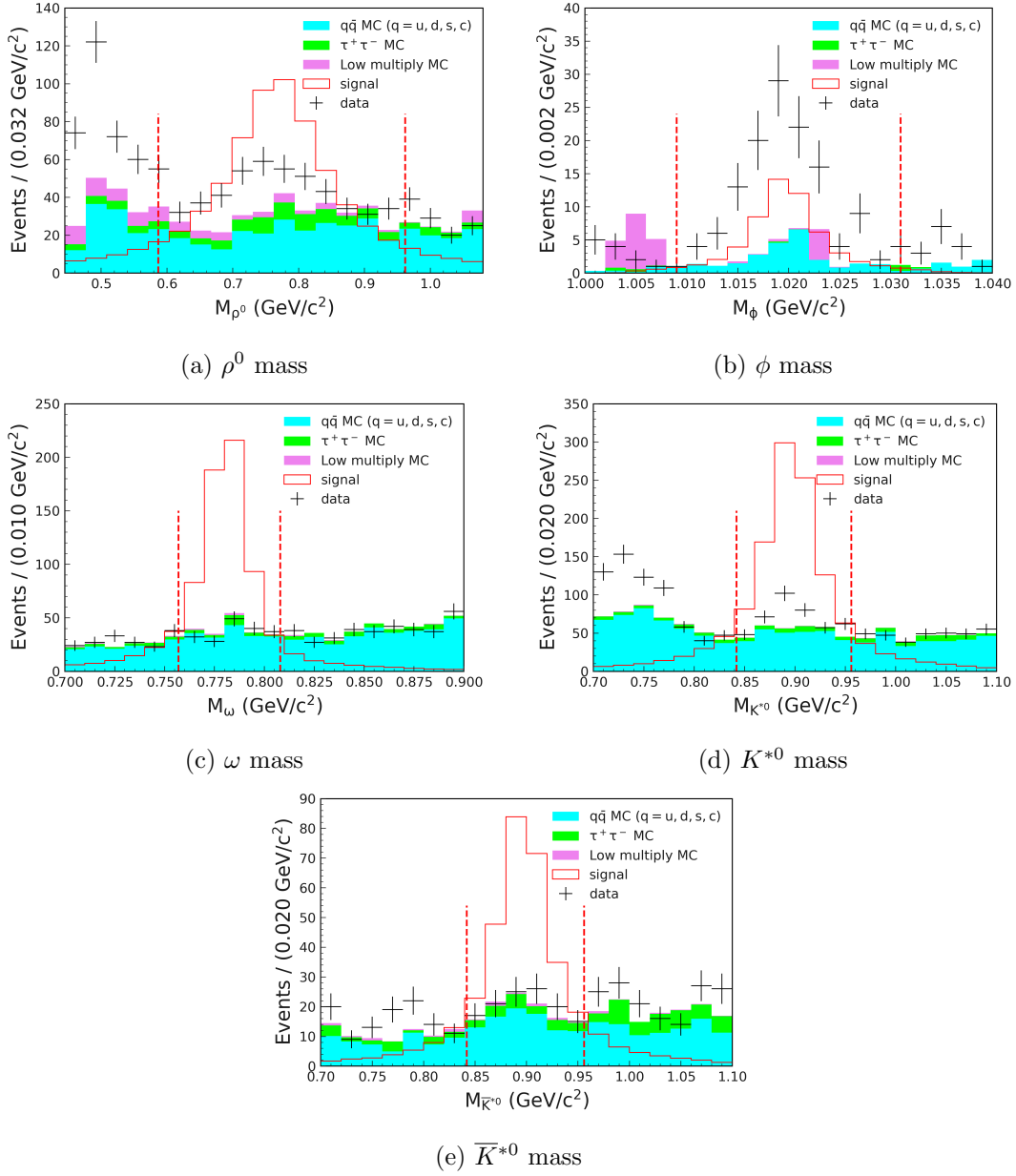


Figure 4.9: Invariant mass distribution of the V^0 for the eV^0 modes. The data and background MC events are in the sideband region where $1.65 < M_{\ell V^0} < 1.9 \text{ GeV}/c^2$ and $|\Delta E| < 0.5 \text{ GeV}$ excluding the blind region. Events between the two red dashed lines are inside the cut-base selection. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

Table 4.2: The M_ν^2 window.

LFV mode	tag decay mode	M_ν^2 (GeV ² /c ⁴)
$\ell\rho^0$ $\mu K^{*0}(\bar{K}^{*0})$	$\pi\nu$	$[-0.1, 0.25]$
	$\pi\pi^0\nu$	$[-0.1, 0.1]$
	$3\pi\nu$	$[-0.05, 0.05]$
	$\pi\pi^0\pi^0\nu$	$[-0.05, 0.05]$
$\mu\omega$	$\pi\nu$	$[-0.1, 0.5]$
	$\pi\pi^0\nu$	$[-0.2, 0.5]$
	$3\pi\nu$	$[-0.1, 0.1]$
	$\pi\pi^0\pi^0\nu$	$[-0.1, 0.2]$
$e\omega$	$\pi\nu$	$[-0.2,]$
	$\pi\pi^0\nu$	$[-0.2,]$
	$3\pi\nu$	$[-0.2, 0.6]$
	$\pi\pi^0\pi^0\nu$	$[-0.4,]$

Table 4.3: The V^0 mass region for the BDT preselection.

V^0	Invariant mass (GeV/c ²)
ρ^0	$0.445 < M_{\pi\pi} < 1.08$
ϕ	$1.0 < M_{KK} < 1.04$
ω	$0.7 < M_{\pi\pi\pi^0} < 0.9$
$K^{*0}(\bar{K}^{*0})$	$0.7 < M_{K\pi} < 1.1$

data sets. Figure 4.10 shows the muon probability distribution of the muon-assigned track in $\mu\rho^0$ mode. The events with muon (electron) probability less than 0.95 (0.9) are available as control samples of the $q\bar{q}$ and $\tau^+\tau^-$ background events. The amount of the control samples are about 100 times larger than that of the $\tau \rightarrow \ell V^0$ analysis data sets.

I must care that the lepton probability of a hadron track depends on the momentum of the track. This results in a small kinematical feature gap between the control samples and the $\tau \rightarrow \ell V^0$ analysis data sets. I consider it as a systematic uncertainty of the BDT selection in section 5.2.2.

4.4.2 BDT training and outputs

The BDT training is done with 40% of the signal MC events as signal samples and 80% of the control samples as background samples. I monitor the area under curve (AUC) of the Receiver Operating Characteristic curve as an indicator of the BDT performance. The validation samples, which have 10% of the signal MC events as signal samples and 20% of the control samples as background samples, are used for this monitoring purpose during the training. If the AUC is not improved for 50 rounds of the training, I stop training and choose the BDT with the best AUC score. The rest of the signal MC events (50%) are used to evaluate the signal efficiency after the BDT selection. The events are distributed to each data set so that the experiment numbers are equal. During the BDT training and validation, a event weight is applied to each of the signal MC events such that the sum of the weights is equal to the number of background events.

The followings are input variables for BDT:

- M_{V^0} , M_ν^2 , $P_\nu^{\text{c.m.}}$, T , P_ℓ^{signal} , $E_{\text{tag}}^{\text{hemisphere}}$, $\cos\theta_{\text{miss-tag}}^{\text{c.m.}}$
- (categorical variables) τ_{tag} decay mode, collision energy

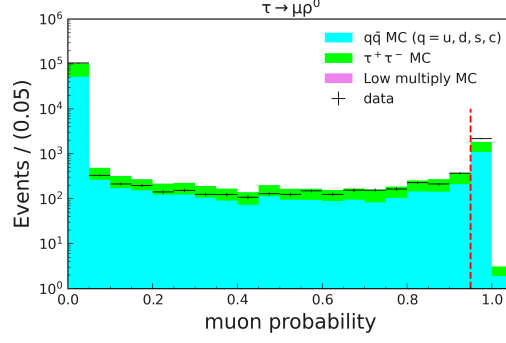


Figure 4.10: The muon probability of the track in the $\tau \rightarrow \mu\rho^0$ mode for the events of $1.65 < M_{\ell V^0} < 1.9\text{GeV}/c^2$ and $|\Delta E| < 0.5\text{ GeV}$. The red dashed line is the lower limit of tracks that is identified as muons. The black line with an error bar is the data events. The stacked filled-histograms are the background MC events. The data samples, whose muon probabilities are less than 0.95 (the red dashed line), are the control samples for the BDT training.

- (additional for the $\ell\omega$ modes) $P_{\pi^0}^{\text{signal}}$, $E_{\gamma}^{\text{lower}}$,

where M_{V^0} is the invariant mass of the vector meson, T is the magnitude of the thrust vector, P_{ℓ}^{signal} is the momentum of the lepton in the signal side, $E_{\text{tag}}^{\text{hemisphere}}$ is the sum of energy of tracks and photons in the tag hemisphere and $P_{\pi^0}^{\text{signal}}$ is the momentum of π^0 from ω . The hyper-parameters of BDT algorithm are as follows: a learning rate is 0.1, the max number of leaves in a tree is 128, the sum of sample weights in each leaf should be more than 5, and the sub-sample is 0.8. The sub-sample means 80% of the training samples are selected randomly for each tree training stage.

Figure 4.11 and Figure 4.12 show the AUC value of the validation samples for each BDT tree training stage (number of BDT trees). Figures 4.13–4.22 show the BDT output distribution of the $\tau \rightarrow \ell V^0$ analysis data sets and its feature importance in the BDT. The event selection with the BDT output (BDT selection) is determined only by the signal efficiency of the cut-base selection instead of optimizing with the signal-to-noise ratio, because the number of background events is too small. The expected number of background events inside the signal region (N_{BG}) is approximately one or less after the cut-base selection (section 5.2.2). The signal efficiency with the BDT selection is set as relatively 5% larger than that with the cut-base selection, because I expect improvement in separating the signal events from the background events.

The finalized BDT selection shows similar N_{BG} to that of the cut-base selection (section 5.2.2). The BDT selection is not applied to the $\ell\phi$ modes because N_{BG} in each of the two modes is small enough.

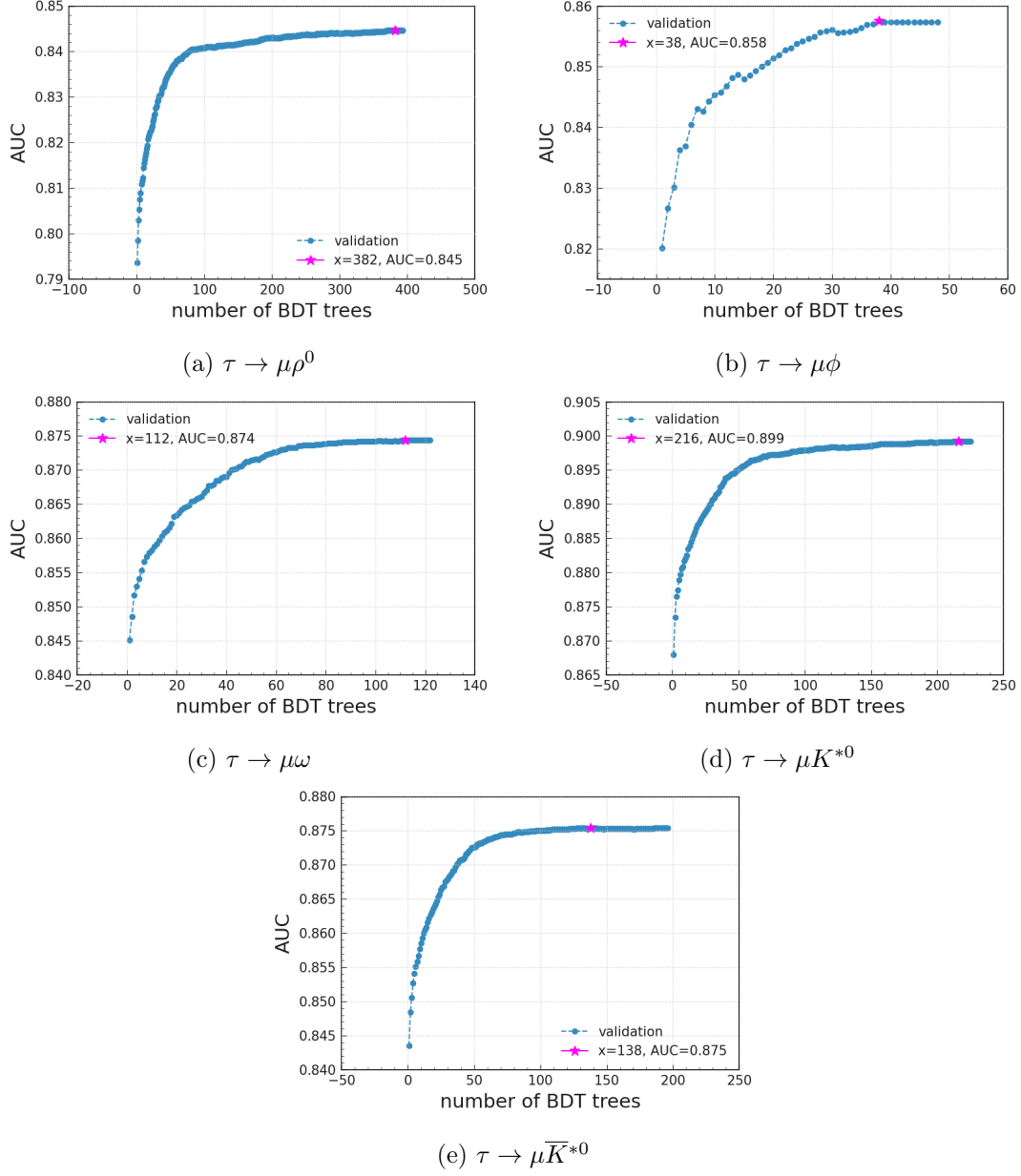
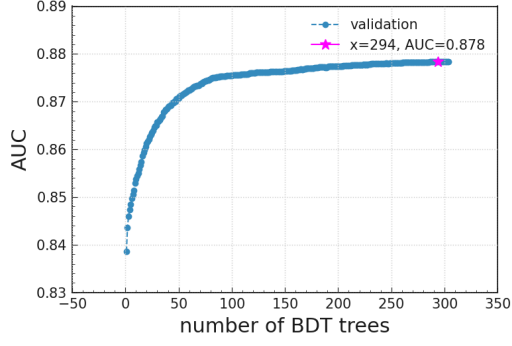
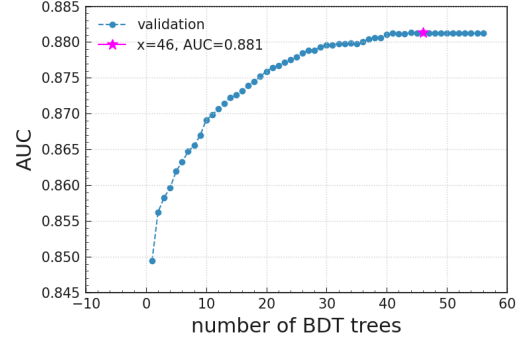


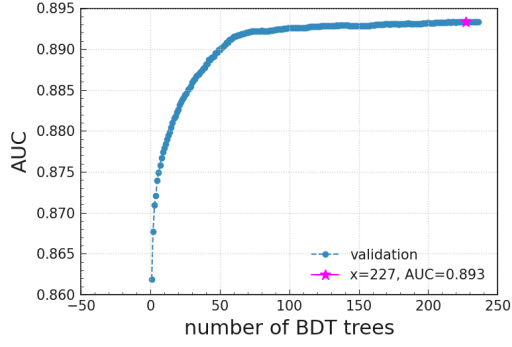
Figure 4.11: The BDT performance indicator (AUC of the validation samples) on each training stage (number of BDT tree) for the μV^0 modes. The violet star mark is the best AUC.



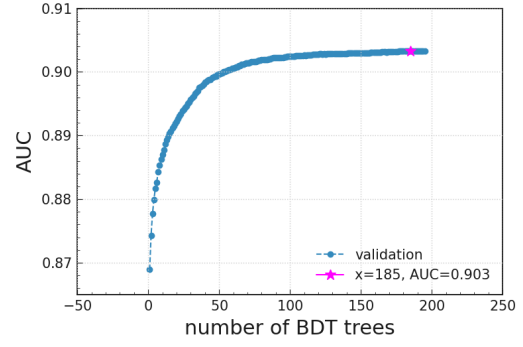
(a) $\tau \rightarrow e\rho^0$



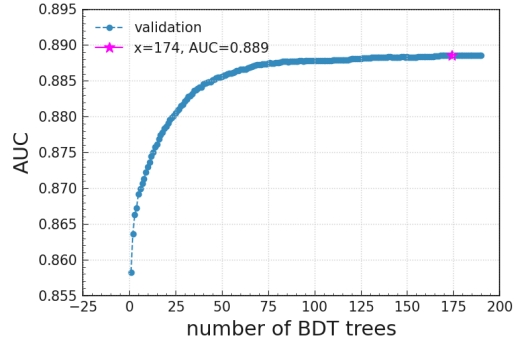
(b) $\tau \rightarrow e\phi$



(c) $\tau \rightarrow e\omega$



(d) $\tau \rightarrow eK^{*0}$



(e) $\tau \rightarrow e\bar{K}^{*0}$

Figure 4.12: The BDT performance indicator (AUC of the validation samples) on each training stage (number of BDT tree) for the eV^0 modes. The violet star mark is the best AUC.

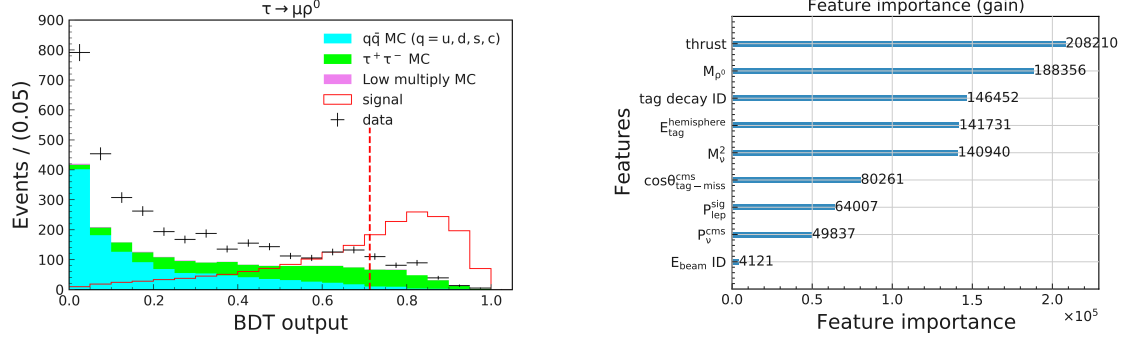


Figure 4.13: The BDT output distribution (left) and its feature importance (right, type is gain that is an indicator of background suppression power) of the $\mu \rho^0$ modes. The red dashed line in the left plot is the lower limit in the BDT selection. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

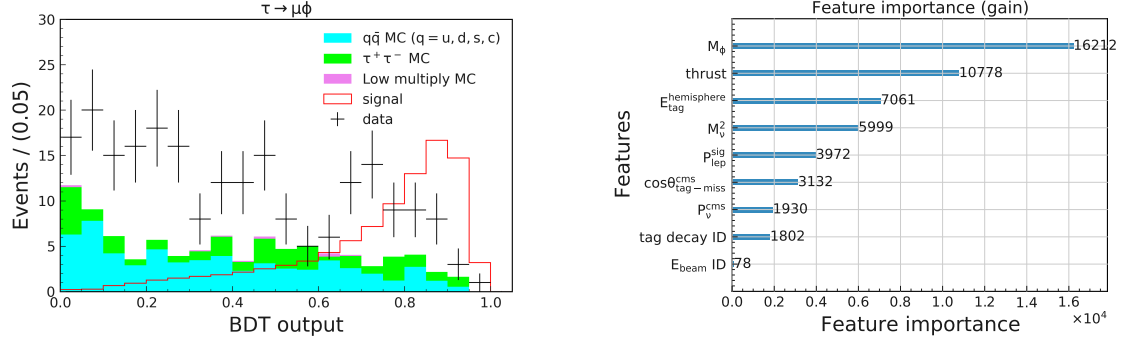


Figure 4.14: The BDT output distribution (left) and its feature importance (right, type is gain that is an indicator of background suppression power) of the $\mu \phi$ modes. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

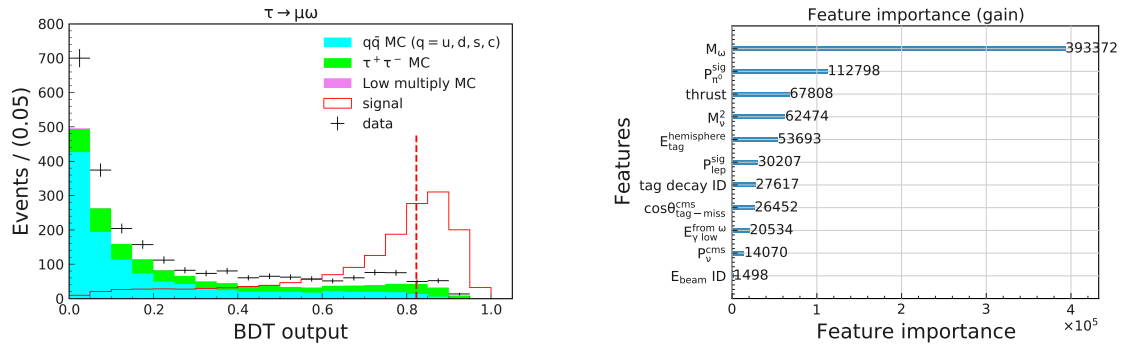


Figure 4.15: The BDT output distribution (left) and its feature importance (right, type is gain that is an indicator of background suppression power) of the $\mu \omega$ modes. The red dashed line in the left plot is the lower limit in the BDT selection. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

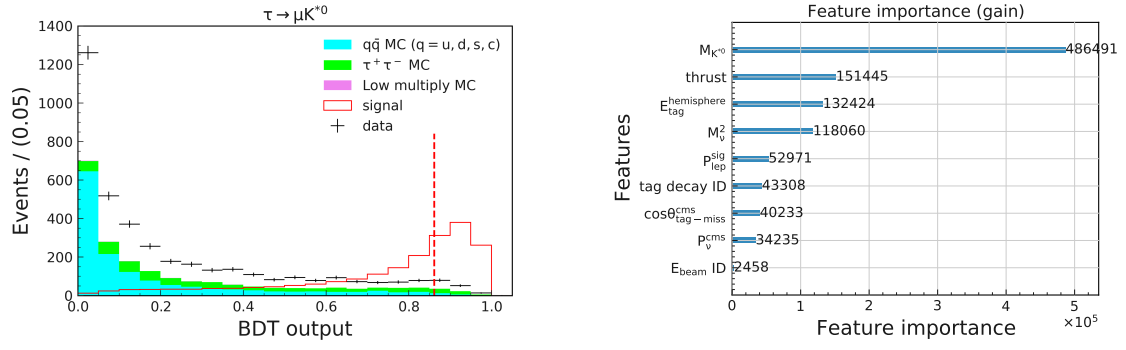


Figure 4.16: The BDT output distribution (left) and its feature importance (right, type is gain that is an indicator of background suppression power) of the μK^{*0} modes. The red dashed line in the left plot is the lower limit in the BDT selection. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

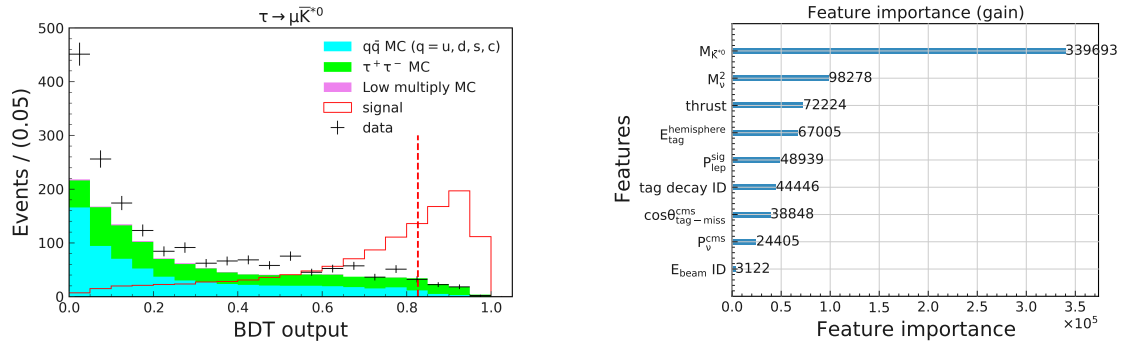


Figure 4.17: The BDT output distribution (left) and its feature importance (right, type is gain that is an indicator of background suppression power) of the $\mu \bar{K}^{*0}$ modes. The red dashed line in the left plot is the lower limit in the BDT selection. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

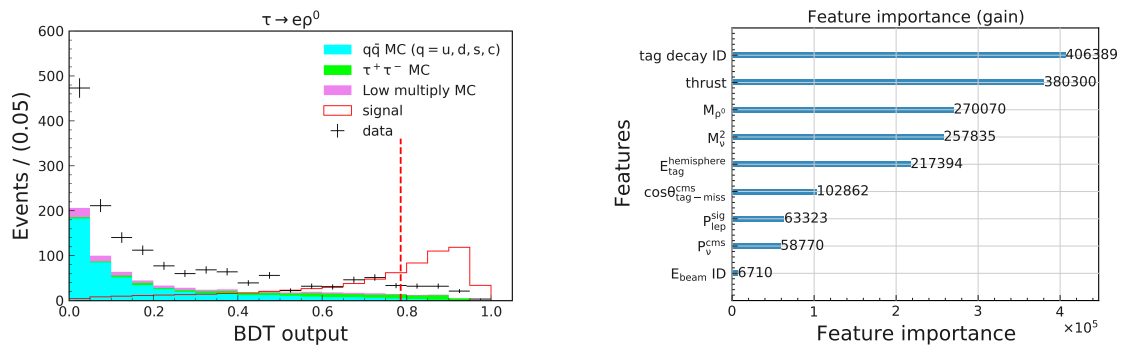


Figure 4.18: The BDT output distribution (left) and its feature importance (right, type is gain that is an indicator of background suppression power) of the $e \rho^0$ modes. The red dashed line in the left plot is the lower limit in the BDT selection. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

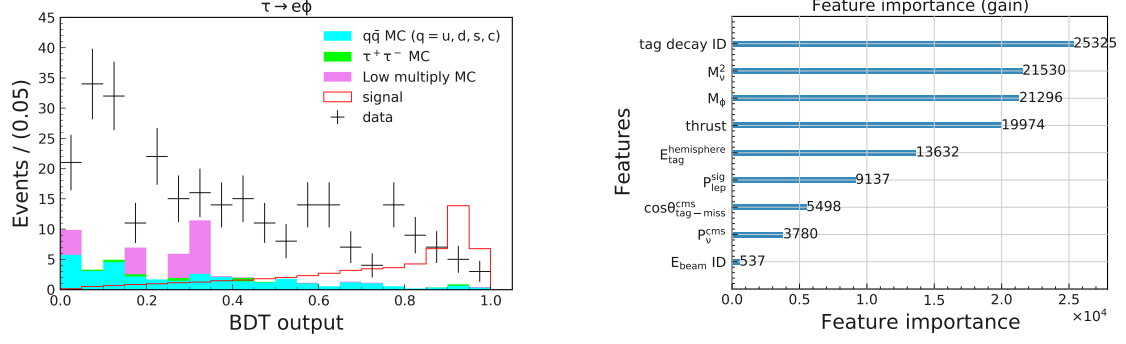


Figure 4.19: The BDT output distribution (left) and its feature importance (right, type is gain that is an indicator of background suppression power) of the $e\phi$ modes. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

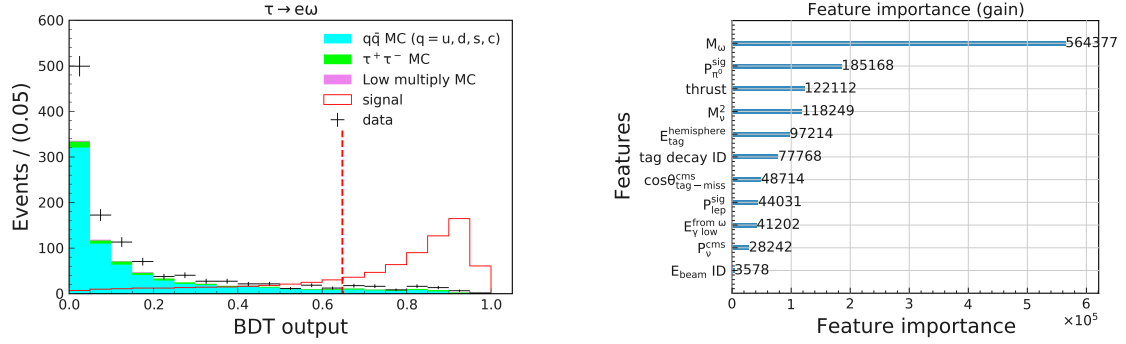


Figure 4.20: The BDT output distribution (left) and its feature importance (right, type is gain that is an indicator of background suppression power) of the $e\omega$ modes. The red dashed line in the left plot is the lower limit in the BDT selection. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

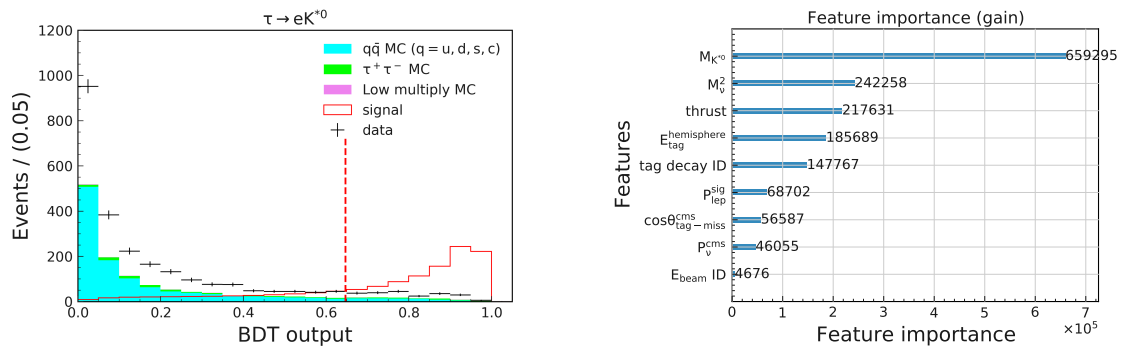


Figure 4.21: The BDT output distribution (left) and its feature importance (right, type is gain that is an indicator of background suppression power) of the eK^{*0} modes. The red dashed line in the left plot is the lower limit in the BDT selection. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

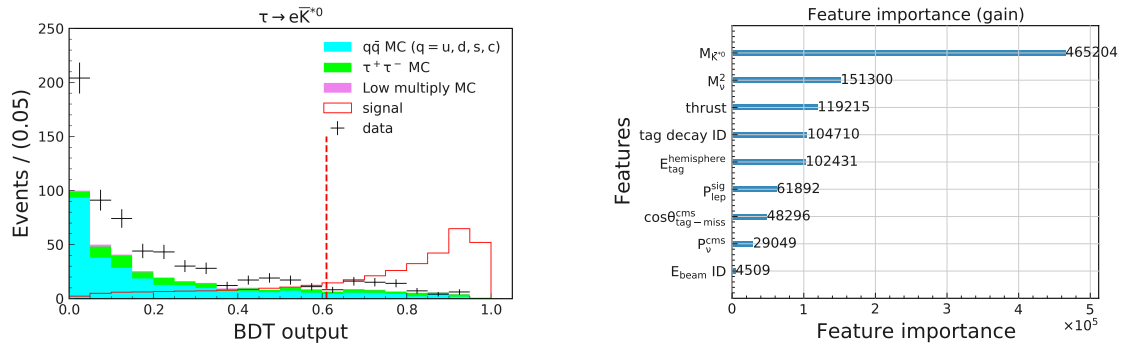


Figure 4.22: The BDT output distribution (left) and its feature importance (right, type is gain that is an indicator of background suppression power) of the $e\bar{K}^{*0}$ modes. The red dashed line in the left plot is the lower limit in the BDT selection. The black line with an error bar is the data events. The red solid histogram is the signal MC events (arbitrary scaled). The stacked filled-histograms are the background MC events.

Chapter 5

Signal efficiency and background estimation

5.1 Correction for the MC events

Although we simulate the Belle detector response and the trigger (section 3.2.7), the MC events have some differences from the collected data. The signal MC events need such data/MC corrections to estimate the $\tau \rightarrow \ell V^0$ signal efficiency in the actual Belle detector. I apply the following Belle official correction factors: the lepton identification efficiency, K/ π identification efficiency, tracking efficiency for slow tracks (momentum less than 0.2 GeV/c), and π^0 efficiency. They are obtained from analyses of well-known SM processes. Because the trigger simulation is not applied for some of MC events, these events are weighted by the trigger efficiency.

I find that the mass resolutions of $D^{(*)\pm}$ mesons are not consistent between the data and MC after similar event selection to the cut-base selection. The mass resolution of the data is worse than that of MC. The track momentum and photon energy resolutions contribute to the mass resolution. Although the track momentum resolution correction has been already studied, they depend on the track multiplicity in an event. The mass resolution of τ in my CLFV search should be also worse, and then the actual signal efficiency will decrease. I study the data/MC correction factor about the track momentum resolution and about the photon energy resolution.

Because the $D^{(*)+} \rightarrow K^-\pi^+\pi^+(\pi^0)$ events have similar event topology and mass with $\tau \rightarrow \ell V^0$ events, I reconstruct them with the tracks, photons and π^0 selections in section 4.1. The tag side selection is either one track or three tracks of $M_{K\pi\pi} < 2.2 \text{ GeV}/c^2$. The signal side tracks are required $\mathcal{P}(e) < 0.9$ and $\mathcal{P}(\mu) < 0.95$, the kaon track is also required K/ π ID > 0.6 . On the signal side, the number of tracks is restricted to be three and the number of photons with $E_\gamma > 0.1 \text{ GeV}$ is less than one (three) for $D^{(*)+} \rightarrow K^-\pi^+\pi^+ (K^-\pi^+\pi^+\pi^0)$. To suppress electron tracks from photon conversion, the invariant mass of the signal side K^- and π^+ assuming electron mass ($M_{ee}(K\pi)$) is more than $0.2 \text{ GeV}/c^2$. The invariant mass of D^+ (D^{*+}) is within 1.80–1.95 (1.95–2.10) GeV/c^2 . I reconstruct D^{*+} from $D^0\pi^+$, and reconstruct D^0 from $K^-\pi^+(\pi^0)$ that invariant mass is within 1.8–1.95 GeV/c^2 . The $\Delta E_{D^{(*)+}} (\equiv E_{D^{(*)+}} - E_{beam}/2)$ is required to $|\Delta E_{D^{(*)+}}| < 0.5 \text{ GeV}$.

I use the Belle $\Upsilon(4S)$ data set for this mass resolution study: 703 fb^{-1} of the data, 4216 fb^{-1} of the $q\bar{q}$ MC ($q = u, d, s, c$), 5619 fb^{-1} of the $\tau^+\tau^-$ MC. These MC samples are scaled to 703 fb^{-1} . Because the lack of some D meson events in $c\bar{c}$ MC, the number of reconstructed events are not consistent with the data. It will not affect to the mass resolution.

I fit the mass distributions of the $D^{(*)\pm}$ events with a Gaussian plus linear function. The σ of the Gaussian function (σ_{Gaus}) corresponds to the $D^{(*)\pm}$ mass resolution. The plots in figure 5.1 are the fit results to the reconstructed mass distributions of $D^{(*)+} \rightarrow K^-\pi^+\pi^+$ events. These $D^{(*)\pm}$ mass resolutions (σ_{Gaus}) are not affected by the photon energy resolution. The track momentum resolution in the MC events is corrected such that the mass resolution of $D^{(*)+}$ meson matches between the data and MC. I smear the track momentum in the MC events by scaling

the Belle official data/MC correction factor for the track momentum resolution. Figure 5.2 shows the $D^{(*)+}$ mass resolutions ($\sigma_{D^{(*)+}}$) in the MC with different track smearing factors. The smearing factor to meet the $D^{(*)+}$ mass resolutions between data and MC is interpolated by a linear function fit. Because the track smearing factor estimations have a gap between the D^+ and D^{*+} study, I roughly take the factor two and consider the difference as the systematic uncertainty of the track smearing factor.

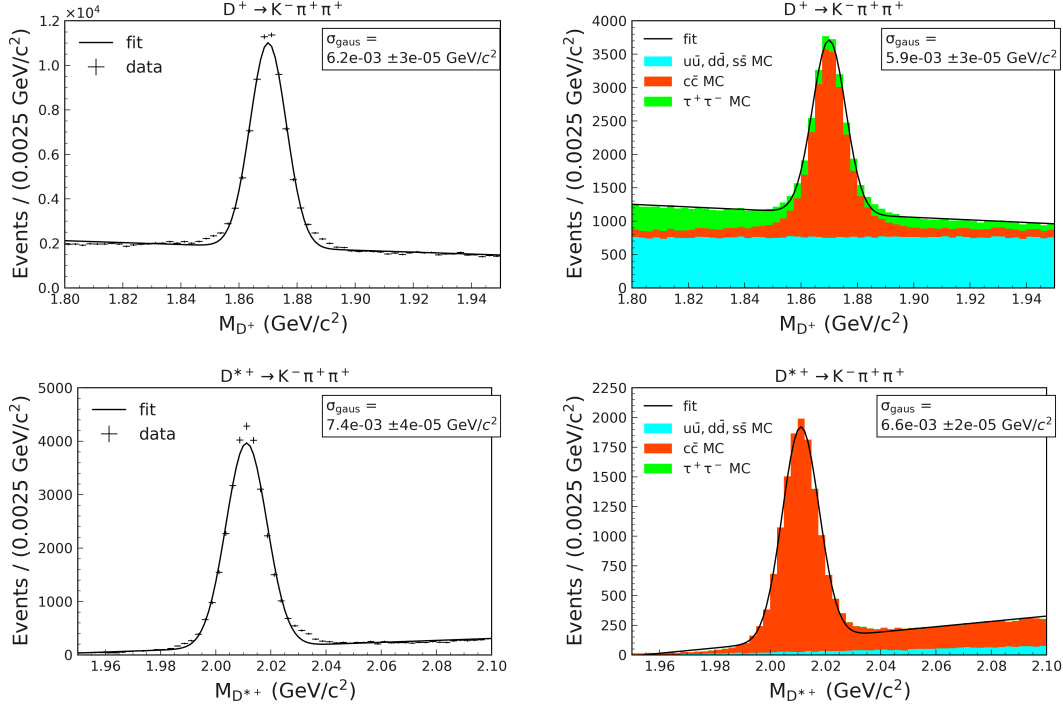


Figure 5.1: The mass distribution of $D^{(*)+}$ ($M_{D^{(*)+}}$) that is reconstructed from $K^-\pi^+\pi^+$, and the fit results for data (left) and MC (right). The fit function (the black line) is a gaussian plus linear function. The σ_{Gaus} corresponds to the $D^{(*)+}$ mass resolution.

The plots in figure 5.3 are the fit results to the reconstructed mass distributions of $D^{(*)+} \rightarrow K^-\pi^+\pi^+\pi^0$ events. Because the above track smearing factor is applied in the figure, the $D^{(*)+}$ mass resolution (σ_{Gaus}) difference between the data and MC is from the photon energy resolution in the π^0 reconstruction. For the data/MC correction about photon energy resolution, I smear the photon energy in the MC events by a constant rate. Figure 5.4 shows the $D^{(*)+}$ mass resolutions ($\sigma_{D^{(*)+}}$) in the MC with different photon energy smearing rate. The smearing rate to meet the $D^{(*)+}$ mass resolutions between data and MC is interpolated by a linear function fit, again. For the photon energy smearing, I choose the D^{*+} result because the error of D^+ study in figure 5.4 is very large. In summary, I apply a factor of 2 for the track smearing scale up and 2% for the photon energy smearing rate to the $\tau \rightarrow \ell V^0$ signal MC events. The signal MC efficiencies relatively decrease by 2.0–4.1% with this data/MC correction about the track momentum and photon energy resolutions.

5.2 Signal and background estimation

5.2.1 Signal region

As shown in figures 4.2 and 4.3, the $\tau \rightarrow \ell V^0$ events have a correlation between the $M_{\ell V^0}$ and ΔE . I define the signal region with an elliptical shape in the $M_{\ell V^0}-\Delta E$ plane. The signal MC samples are after the cut-base selection in section 4.3 and the data/MC correction in section 5.1. Most of the signal events cluster around $M_{\ell V^0} = 1.777 \text{ GeV}/c^2$ and $\Delta E = 0 \text{ GeV}$. The ellipse

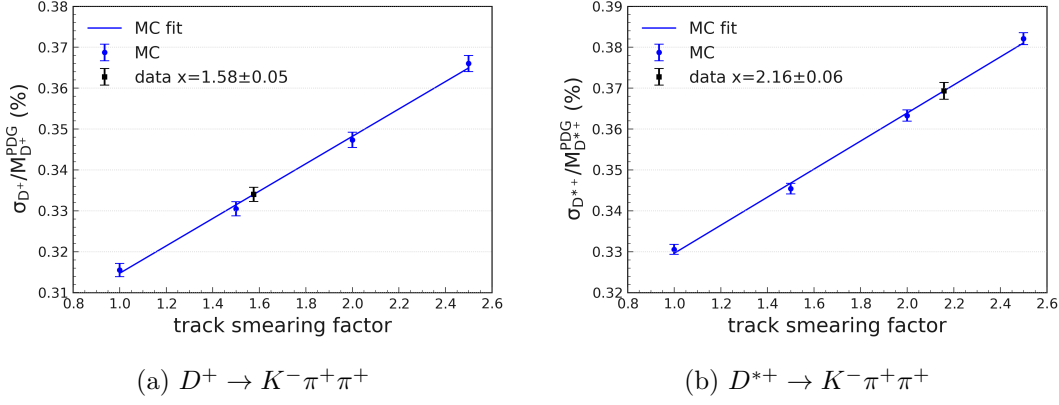


Figure 5.2: Study about the track smearing factor vs. $D^{(*)+}$ mass resolution ($\sigma_{D^{(*)+}}$) divided by the world average $D^{(*)+}$ mass in the particle data group ($M_{D^{(*)+}}^{\text{PDG}}$). The blue line represents the linear fit to the MC (blue points). The data (black point) is shown on the fit function to estimate the smearing factor. The track smearing factor one means the original Belle official data/MC correction factor.

oblateness and the rotation angle are calculated from the covariance matrix of the signal MC distribution. The center of the ellipse is the mean of the distribution. At this point, events with wrong reconstruction particles or with missing bremsstrahlung photons are removed by MC information because they might distort the ellipse. More $q\bar{q}$ and $\tau^+\tau^-$ background events distribute in the lower ΔE region and the lower $M_{\ell V^0}$ region (see section 5.2.2). The ellipse size is determined to maximize the figure of merit (FOM) [79],

$$FOM = \frac{\varepsilon}{\frac{\alpha}{2} + \sqrt{N_{\text{BG}}}}, \quad (5.1)$$

where ε is the signal efficiency inside the ellipse, α is the confidence coefficient ($\alpha = 1.64$ at 90% C.L.), and N_{BG} is the expected number of background events inside the signal region. Wrong MC matching or photon loss events in the signal MC are included in this scan step.

The signal efficiency of each Belle experiment number (Belle run periods) is shown in figures 5.5–5.8. The variations come from the Belle detector condition and so on. The signal efficiencies of the c.m. beam energies at $\Upsilon(1\text{--}3\text{S})$ become small due to the BDT selection. It may imply that the BDT training sample for these events are not enough.

5.2.2 Number of background events estimation

I estimate N_{BG} through interpolation from the sideband data. Here the sideband data is a set of data passing the event selection (cut-base or BDT) and inside the sideband region: $1.65 \text{ GeV}/c^2 < M_{\ell V^0} < 1.9 \text{ GeV}/c^2$ and $|\Delta E| < 0.1 \text{ GeV}$ outside of the blind region. The interpolation is based on a two-dimensional function that represents the background distribution in the $M_{\ell V^0}$ – ΔE plane.

To define the function, I use the background control samples in section 4.4.1 with the cut-base selection except the lepton identification. The two-dimensional function is obtained by fitting to the distribution of this control samples within $|\Delta E| < 0.1 \text{ GeV}$, and then it is scaled to the sideband data. The fitting region is limited within $|\Delta E| < 0.1 \text{ GeV}$ to simplify the fit function of $\tau^+\tau^-$ background. Figures 5.9–5.18 show the distributions of the control samples for the data and MC. The main background sources are the $q\bar{q}$ and $\tau^+\tau^-$ events. Some $D^\pm$ decay events cluster around the $D^\pm$ mass region for the $\ell\phi$ and ℓK^{*0} modes, which have $K^\pm$ track(s). The ΔE distribution of the $D^\pm$ events are different between the data and MC, because $ee \rightarrow D^+D^-$ events are lack in the MC. The $D^\pm$ from $D^{*\pm} \rightarrow (D\pi)^\pm$ events cluster at the $D^\pm$ mass and ΔE less than zero GeV.

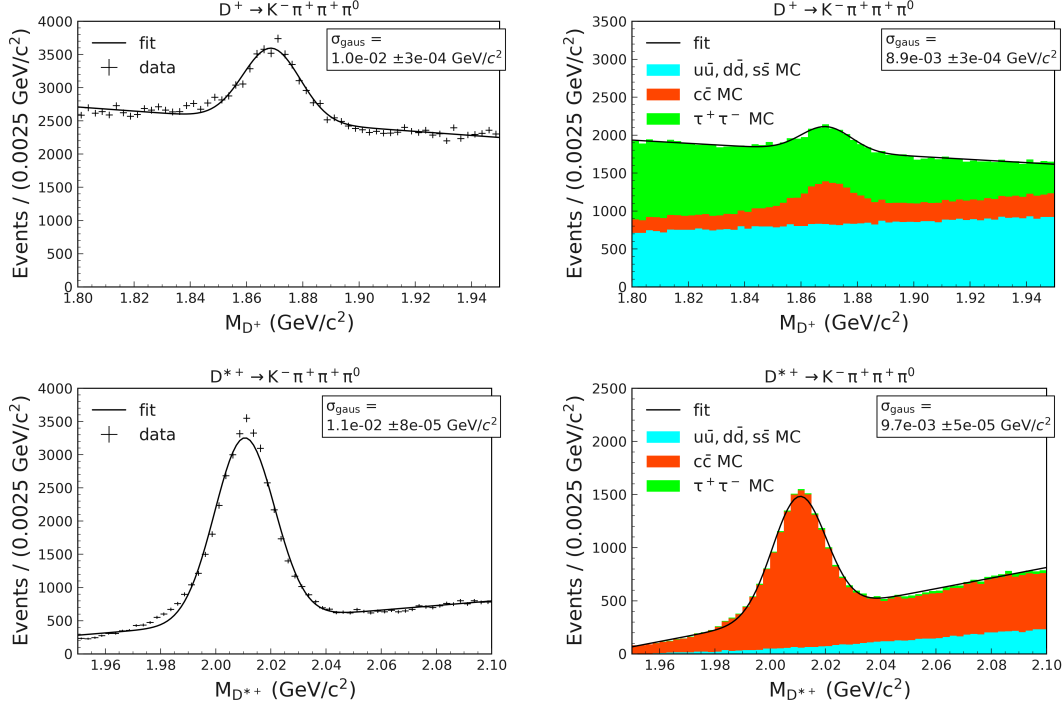


Figure 5.3: The mass distribution of $D^{(*)+}$ ($M_{D^{(*)+}}$) that is reconstructed from $K^-\pi^+\pi^+\pi^0$, and the fit results for data (left) and MC (right, the track smearing is applied). The fit function (the black line) is a gaussian plus linear function. The σ_{Gaus} corresponds to the $D^{(*)+}$ mass resolution.

The fit functions for each ℓV^0 mode are:

$$F(M_{\ell V^0}, \Delta E) = f(M_{\ell V^0}) \times \frac{1}{1 + \exp[a_y(\Delta E - y_0)]} + c_0^{\text{flat}}, \quad (5.2)$$

$$f(x) = \begin{cases} \int_{x-5\sigma}^{x+5\sigma} g(x') \times \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-x')^2}{2\sigma^2}\right] dx' & (V^0 = \rho^0, \omega) \\ c_1(x-x_0)^2 + c_0 & (V^0 = K^{*0}, \bar{K}^{*0}) \\ c_0 & (V^0 = \phi) \end{cases} \quad (5.3)$$

$$g(x) = \begin{cases} c_1[(x-x_0)^2 + k(x-x_0)] + c_0 & (x < x_0, V^0 = \rho^0) \\ c_1(x-x_0) + c_0 & (x < x_0, V^0 = \omega) \\ c_0 & (x \geq x_0) \end{cases}$$

where $f(x)$ represents the background distribution along with $M_{\ell V^0}$; c_1 , c_0 , x_0 and k are parameters that define the shape of the function; a_y represents sharpness of the sigmoid function along the ΔE axis; y_0 is the center of the sigmoid function; and c_0^{flat} is a term of flat background events in the $M_{\ell V^0}$ - ΔE plane. I define $f(x)$ for each V^0 in eq. (5.3), and the functions for the $\ell\rho^0$ ($\ell\omega$) modes are smeared by a gaussian with standard deviation, σ , of 6.6 (9.6) MeV/ c^2 estimated by the $D^{*\pm}$ mass fitting in the section 5.1. This σ corresponds to the mass resolution that affects the edge of the $M_{\ell V^0}$ distribution close to the τ mass for the $\tau^+\tau^-$ background events. Because the edge is broad for the other modes owing to wrong mass assignment of fake kaons, the smearing of the function is not necessary. The $\tau^+\tau^-$ background events for the $\ell\phi$ modes are included in c_0 because they are flat along the $M_{\ell V^0}$ axis in the $1.65 \text{ GeV}/c^2 < M_{\ell V^0} < 1.9 \text{ GeV}/c^2$ region.

I obtain the optimal fit parameters by a likelihood fit using MINUIT [80]. The following region is excluded from the fitting to avoid $D^+ \rightarrow K^-\pi^+\pi^+\pi^0$ and $D^+ \rightarrow \pi^+\phi$ background events, which cluster around the $D^\pm$ meson mass: $1.83(1.82) \text{ GeV}/c^2 \leq M_{\ell V^0} < 1.89 \text{ GeV}/c^2$ and $\Delta E <$

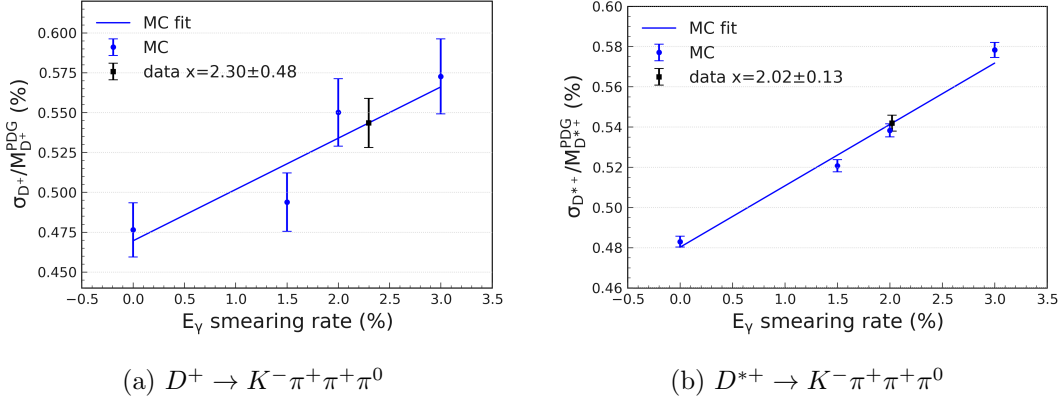


Figure 5.4: Study about the photon energy smearing rate (%) vs. $D^{(*)+}$ mass resolution ($\sigma_{D^{(*)+}}$) divided by the world average $D^{(*)+}$ mass in the particle data group ($M_{D^{(*)+}}^{\text{PDG}}$).

The blue line represents the linear fit to the MC (blue points). The data (black point) is shown on the fit function to estimate the smearing factor. The track smearing factor is applied to MC. The E_γ smearing rate zero means the original photon energy resolution of the Belle MC.

0.04 GeV for the μK^{*0} ($e K^{*0}$) and $\mu\phi$ ($e\phi$) modes. The fit results to the control samples are shown in figures 5.19 and 5.20.

The parameters of a_y , y_0 , k and x_0 are fixed at the fit results of the control samples within $|\Delta E| < 0.1$ GeV. The fit uncertainties of these fixed parameters are included in the systematic uncertainty of N_{BG} . The other fit parameters correspond to the scale factors of each background component: generic $\tau^+\tau^-$ (c_1), $q\bar{q}$ and low multiplicity background events (c_0 and c_0^{flat}). I fit the function floating these scale factors (c_1 , c_0 and c_0^{flat}) to the sideband data. Figures 5.21 and 5.22 are the fit results of the cut-base selection. Figures 5.23 and 5.24 are the fit results of the BDT selection. The same region around the D meson mass as for the fit to the control samples is excluded from the fitting for the $\ell\phi$ and ℓK^{*0} modes. The functions are integrated in the elliptical signal regions to deduce N_{BG} , which are in the range of 0.25–0.95 for the BDT selection (summarized in tables 6.1 and 6.2 in section 6.1).

The major contribution to N_{BG} comes from the $M_{\ell V^0}$ flat term in eq. (5.2) (c_0 and c_0^{flat}), which corresponds to the $q\bar{q}$ or low multiplicity background events. For the BDT selection, the contribution of the generic $\tau^+\tau^-$ background events, which depends on $M_{\ell V^0}$, is about one-third as large as the other background contributions. I cannot distinguish the $\ell\phi$ background components through the fit to the data, because the generic $\tau^+\tau^-$ background events are distributed evenly along the $M_{\ell V^0}$ axis.

The systematic uncertainty of N_{BG} is summarized in table 5.1. They are estimated by the following procedure:

Statistical uncertainty of N_{BG}

I generate 100 sets of pseudo-data for each mode in the $M_{\ell V^0}-\Delta E$ histogram. The content of each bin in these histogram is set randomly following a Poisson distribution, with the mean taken from the function fitted to the sideband data (figures 5.21–5.24). I fit the function to each set of the pseudo-data to deduce N_{BG} , and the standard deviation of these N_{BG} distribution is taken as the statistical uncertainty.

Systematic uncertainty of N_{BG} from the uncertainties of fit parameters

The fit functions has fixed parameters that determine the shape of $q\bar{q}$ and $\tau^+\tau^-$ background distribution. Those parameters are estimated using the control samples. I fluctuate one of the fixed parameters by Gauss random number with a sigma of its fit error. Then, repeat the procedure of fitting to the sideband data and estimate N_{BG} . I try this process for 100 times for each fixed parameter and get the standard deviation of N_{BG} .

Systematic uncertainty of N_{BG} from kinematical difference

Another systematic uncertainty of N_{BG} comes from difference of the $M_{\ell V^0} - \Delta E$ distributions between the control samples and the sideband data within $|\Delta E| < 0.1$ GeV. The difference mainly arises from the lepton identification fake rate that depends on the momentum and θ of the track. The sideband data have a pion misidentified as a lepton, which tends to have a lower momentum than the pions in the control samples. I evaluate a change of N_{BG} when the parameters— a_y , y_0 , k and x_0 —are re-determined with weighted control samples; the weight (w_{ij}) on each event is computed as the ratio of the numbers of events in the momentum- θ bin for the signal fake-lepton track between all the data passing the event selection and all the control samples in $|\Delta E| < 0.5$ GeV. ($w_{ij} = N_{ij}^{\text{obs}} / N_{ij}^{\text{BGenhanced}}$, i (j) is the pion momentum (theta) bin number, N^{obs} is number of the data after the event selection within $|\Delta E| < 0.5$ GeV, $N^{\text{BGenhanced}}$ is number of the control samples within $|\Delta E| < 0.5$ GeV. Both N^{obs} and $N^{\text{BGenhanced}}$ are normalized.) The amount of change of N_{BG} is taken as the systematic uncertainty of N_{BG} .

These N_{BG} uncertainties are summarized in tables 6.1 and 6.2 in section 6.1.

Mode	The uncertainty of N_{BG} from fit errors		N_{BG} difference with weighted control samples	
	cut-base	BDT	cut-base	BDT
$\mu^- \rho^0$	0.10	0.07	−0.08	+0.08
$e^- \rho^0$	0.01	0.01	+0.01	+0.02
$\mu^- \phi$	0.04	0.05	−0.02	−0.02
$e^- \phi$	0.00	0.00	0.00	0.00
$\mu^- \omega$	0.07	0.03	+0.01	−0.01
$e^- \omega$	0.05	0.01	−0.02	0.00
$\mu^- K^{*0}$	0.02	0.03	+0.19	+0.01
$e^- K^{*0}$	0.02	0.12	−0.03	−0.02
$\mu^- \bar{K}^{*0}$	0.03	0.02	−0.11	−0.05
$e^- \bar{K}^{*0}$	0.02	0.01	−0.01	−0.01

Table 5.1: The systematic uncertainties of N_{BG} . The second column is the standard deviation of N_{BG} with various fit function parameters. The third column is the differences of N_{BG} estimated with the weighted control samples and the nominal control samples.

5.2.3 Validation of number of background estimation

To validate the estimation method of number of background events in section 5.2.2, I compare the estimation and the actual number of events in a validation region. The validation region is inside the blind region but outside the elliptical signal region. The N_{BG} in the validation region is estimated by integrating the two-dimensional function obtained in section 5.2.2 in the validation region. The expected and observed number of events in the validation regions are listed in table 5.2.

I check the probability of N_{obs} excess from expected N_{BG} assuming Poisson distribution which mean is the expected N_{BG} . The probability of two modes are less than 5%, μK^{*0} and $\mu \bar{K}^{*0}$ modes with cut-base selection. These expected N_{BG} might need correction, if the number of observed events in the signal region is not zero. If they are zero, the 90% C.L. upper limits of $\tau^- \rightarrow \mu^- K^{*0} (\bar{K}^{*0})$ branching fraction are not affected by the expected N_{BG} .

5.2.4 Comparison with old background estimation method

In the previous paper, the number of background events are estimated assuming flat distribution along the ΔE axis. Here, I check how much the ΔE distribution assumption affects to the N_{BG} estimation.

I assume that the ΔE distribution of the background events is flat within -0.1 to 0.1 GeV, and interpolate the N_{BG} from the $M_{\ell V^0}$ sideband region in $|\Delta E| < 0.1$ GeV. For the $\ell\rho^0$ and $\mu\omega$ modes, I use the functions $f(x)$ in eq. 5.3. The function parameters except c_1 ($\tau^+\tau^-$ background component factor) and c_0 ($q\bar{q}$ and low multiplicity background component factor) are determined by fitting to the control samples in section 4.4.1 after the cut-base selection. In the same way as in section 5.2.2, the function is scaled to the sideband region after the cut-base or BDT selection. For the other modes, I expect the $M_{\ell V^0}$ distribution of the background events is flat within 1.65 to 1.90 GeV/ c^2 . In case of $\ell\phi$ and ℓK^{*0} modes, I blind the events from $D^\pm$ meson decay gathering around $D^\pm$ meson mass: $1.83(1.82)$ GeV/ $c^2 \leq M_{\ell V^0} < 1.89$ GeV/ c^2 region for the μK^{*0} ($e K^{*0}$) and $\mu\phi$ ($e\phi$) modes. Table 5.3 is the estimation of the number of background events in the signal region using this method. They are consistent with the two-dimensional function method in section 5.2.2.

5.3 Systematic uncertainty of the expected number of signal events

The systematic uncertainties of the expected number of signal events are listed in table 5.4. The dominant uncertainties are from the particle identifications.

The following systematic uncertainties are from the Belle official value or other Belle analyses. The uncertainties of lepton identification efficiency are 1.7(1.8%) per electron(muon). The uncertainties of K/ π identification efficiency are 0.8% per pion from ρ^0 , 0.9% per kaon from ϕ and 1.1% per kaon from $K^{*0}(\bar{K}^{*0})$. The uncertainty of charged track finding with $p \geq 200$ MeV/ c is estimated to be 0.35% per track. The bremsstrahlung photon energy correction is applied to about 5% of electron tracks on the signal side after the cut-base selection; the uncertainty of photon efficiency is 2.0% per photon, I set 0.1% on average for eV^0 modes. I use the π^0 efficiency data/MC correction, where the uncertainty is 2.2% per π^0 .

I also check the other systematic uncertainties with my $\tau \rightarrow \ell V^0$ signal MC events. The uncertainty of the electron veto on hadron tracks is set to the same value as the signal inefficiency due to the electron veto, i.e., 0.4–1.2% for each LFV mode. The uncertainty from the signal MC statistics is 0.3–0.5%. The branching fractions of each τ_{tag} generic decay mode have uncertainties of 0.04–0.10% in PDG 2022; they result in the systematic uncertainty about 0.1% for both of the cut-base and BDT selection.

About 34% of the generated signal MC samples are losing energy more than 0.1 GeV by bremsstrahlung photon(s) from the $e^\pm$ before collision (initial state radiation, ISR). Because most of them are out from the signal region due to energy loss, the uncertainty of the ISR

Table 5.2: N_{BG} estimation in the validation region after the cut-base (BDT) selection.

Mode	cut-base selection		BDT selection	
	expected N_{BG}	N_{obs}	expected N_{BG}	N_{obs}
$\tau^- \rightarrow \mu^- \rho^0$	$5.9 \pm 1.1(\text{stat}) \pm 0.4(\text{syst})$	8	$5.5 \pm 1.0(\text{stat}) \pm 0.9(\text{syst})$	7
$\tau^- \rightarrow e^- \rho^0$	$4.5 \pm 1.5(\text{stat}) \pm 0.0(\text{syst})$	2	$5.0 \pm 1.6(\text{stat}) \pm 0.1(\text{syst})$	5
$\tau^- \rightarrow \mu^- \phi$	$2.3 \pm 0.7(\text{stat}) \pm 0.1(\text{syst})$	2	$2.6 \pm 0.8(\text{stat}) \pm 0.1(\text{syst})$	3
$\tau^- \rightarrow e^- \phi$	$1.2 \pm 0.7(\text{stat}) \pm 0.0(\text{syst})$	3	$2.3 \pm 1.2(\text{stat}) \pm 0.0(\text{syst})$	5
$\tau^- \rightarrow \mu^- \omega$	$3.2 \pm 1.3(\text{stat}) \pm 0.2(\text{syst})$	0	$1.4 \pm 0.8(\text{stat}) \pm 0.1(\text{syst})$	0
$\tau^- \rightarrow e^- \omega$	$1.5 \pm 0.8(\text{stat}) \pm 0.1(\text{syst})$	4	$1.7 \pm 0.9(\text{stat}) \pm 0.0(\text{syst})$	2
$\tau^- \rightarrow \mu^- K^{*0}$	$2.5 \pm 0.8(\text{stat}) \pm 0.5(\text{syst})$	9	$3.3 \pm 1.0(\text{stat}) \pm 0.1(\text{syst})$	7
$\tau^- \rightarrow e^- K^{*0}$	$5.3 \pm 1.4(\text{stat}) \pm 0.1(\text{syst})$	9	$3.3 \pm 1.3(\text{stat}) \pm 0.8(\text{syst})$	5
$\tau^- \rightarrow \mu^- \bar{K}^{*0}$	$1.7 \pm 0.5(\text{stat}) \pm 0.3(\text{syst})$	5	$2.1 \pm 0.6(\text{stat}) \pm 0.1(\text{syst})$	3
$\tau^- \rightarrow e^- \bar{K}^{*0}$	$1.0 \pm 0.5(\text{stat}) \pm 0.1(\text{syst})$	2	$1.0 \pm 0.5(\text{stat}) \pm 0.0(\text{syst})$	2

Table 5.3: The expected number of background events with the flat distribution assumption along the ΔE axis for the cut-base and BDT selection. They are consistent with my results in section 5.2.2.

Mode	N_{BG} (cut-base)	N_{BG} (BDT)
$\tau^- \rightarrow \mu^- \rho^0$	0.90 ± 0.23	1.46 ± 0.33
$\tau^- \rightarrow e^- \rho^0$	0.57 ± 0.26	0.87 ± 0.25
$\tau^- \rightarrow \mu^- \phi$	0.37 ± 0.14	0.43 ± 0.15
$\tau^- \rightarrow e^- \phi$	0.23 ± 0.17	0.35 ± 0.21
$\tau^- \rightarrow \mu^- \omega$	0.58 ± 0.34	0.25 ± 0.16
$\tau^- \rightarrow e^- \omega$	0.66 ± 0.39	0.74 ± 0.41
$\tau^- \rightarrow \mu^- K^{*0}$	0.90 ± 0.24	1.11 ± 0.29
$\tau^- \rightarrow e^- K^{*0}$	0.81 ± 0.28	0.51 ± 0.23
$\tau^- \rightarrow \mu^- \bar{K}^{*0}$	0.57 ± 0.18	0.57 ± 0.18
$\tau^- \rightarrow e^- \bar{K}^{*0}$	0.35 ± 0.18	0.26 ± 0.16

event rate in the generator affects our $\tau \rightarrow \ell V^0$ signal efficiency. KK Monte Carlo generator (KKMC [71]) has uncertainty about the $ee \rightarrow ff\gamma$ cross section [81, 82]. I take 0.5% uncertainty as a $ee \rightarrow \tau\tau\gamma$ cross section conservatively, then the propagated uncertainty of the signal efficiency is 0.1%.

The track and photon energy resolution uncertainties are estimated from the $D^{(*)+}$ mass study in section 5.1. I generate signal MC sets with different smearing factors of track momentum and photon energy resolution: track/photon smearing factors are 1.6/2.0%, 2.2/2.0%, 2.0/1.9% and 2.0/2.2%. Because M_τ and ΔE resolutions differ among those samples, the signal efficiency is also different with uncertainties of 0.6–1.5% for the cut-base selection and 0.5–1.3% for the BDT selection. This track/photon smearing factors also affect the other kinematical variables, and thus these uncertainties include the cut-base and BDT selection uncertainty.

All the uncertainties in table 5.4 are summed in quadrature to yield the total systematic uncertainties shown in tables 6.1 and 6.2. The total systematic uncertainties are slightly decreased from the previous paper [18], thanks to the updated Belle tracking study.

Table 5.4: The systematic uncertainties of the expected number of signal events. The average number of tracks (particles) in the reconstructed $\tau^+\tau^-$ events for each signal mode is represented as $N_{\text{track}(\text{particle})}$. When the uncertainty is different mode by mode, I show the range of the uncertainty.

Source	σ_{syst} (%)
Belle integrated luminosity	1.4
$ee \rightarrow \tau\tau(\gamma)$ cross section	0.3
$\text{Br}(\phi \rightarrow KK)$ and $\text{Br}(\omega \rightarrow \pi\pi\pi^0)$	1.2 and 0.7
Trigger efficiency	0.2–0.9
Tracking efficiency	$0.35 \times N_{\text{track}}$
Electron identification efficiency	$1.7 \times N_{\text{electron}}$
Muon identification efficiency	$1.8 \times N_{\text{muon}}$
K and π identification efficiency	1.6 (ρ^0), 1.8 (ϕ) and 1.1 (K^{*0} and $\bar{K}^{*0}$)
photon efficiency for bremsstrahlung correction (eV^0)	0.1
π^0 efficiency	$2.2 \times N_{\pi^0}$
Electron veto for hadrons	0.4–1.2
MC statistics	0.3–0.5
τ_{tag} decay mode fraction	0.1
KKMC ISR uncertainty	0.1
Track and photon energy resolution	0.5–1.3

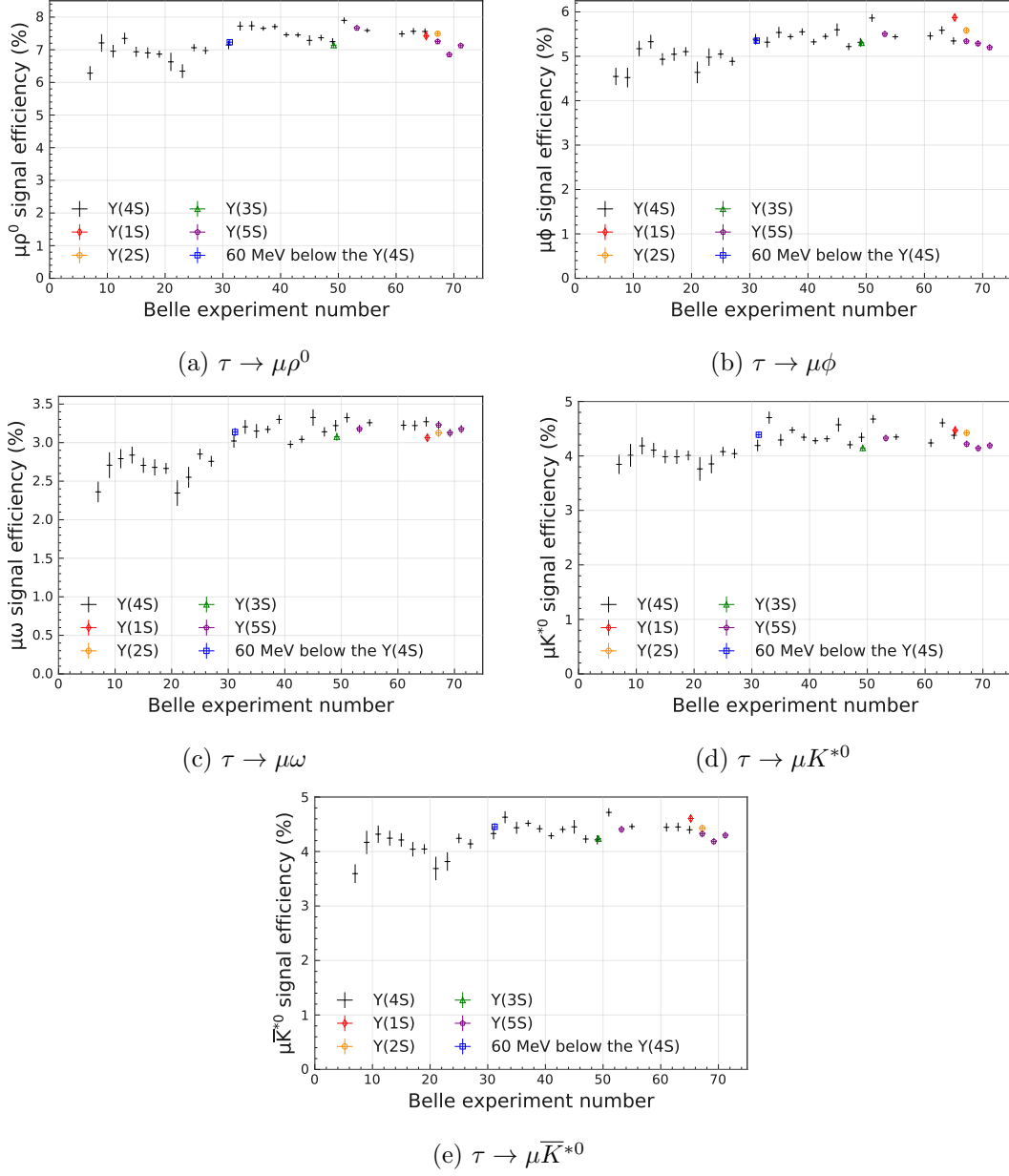


Figure 5.5: The μV^0 signal efficiencies for each Belle run period (the cut-base selection). The legends show the c.m. energy of the collision.

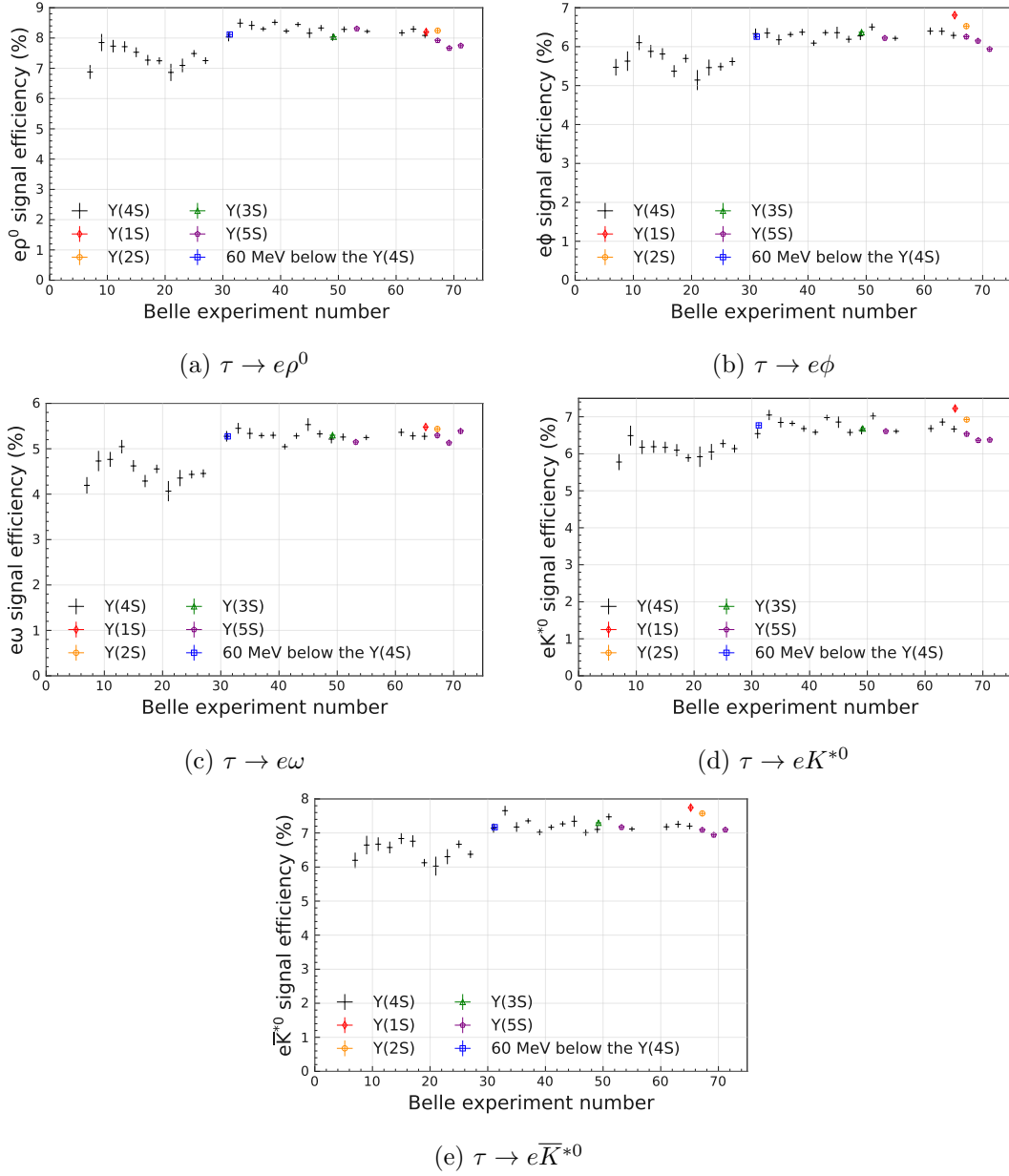


Figure 5.6: The eV^0 signal efficiencies for each Belle run period (the cut-base selection). The legends show the c.m. energy of the collision.

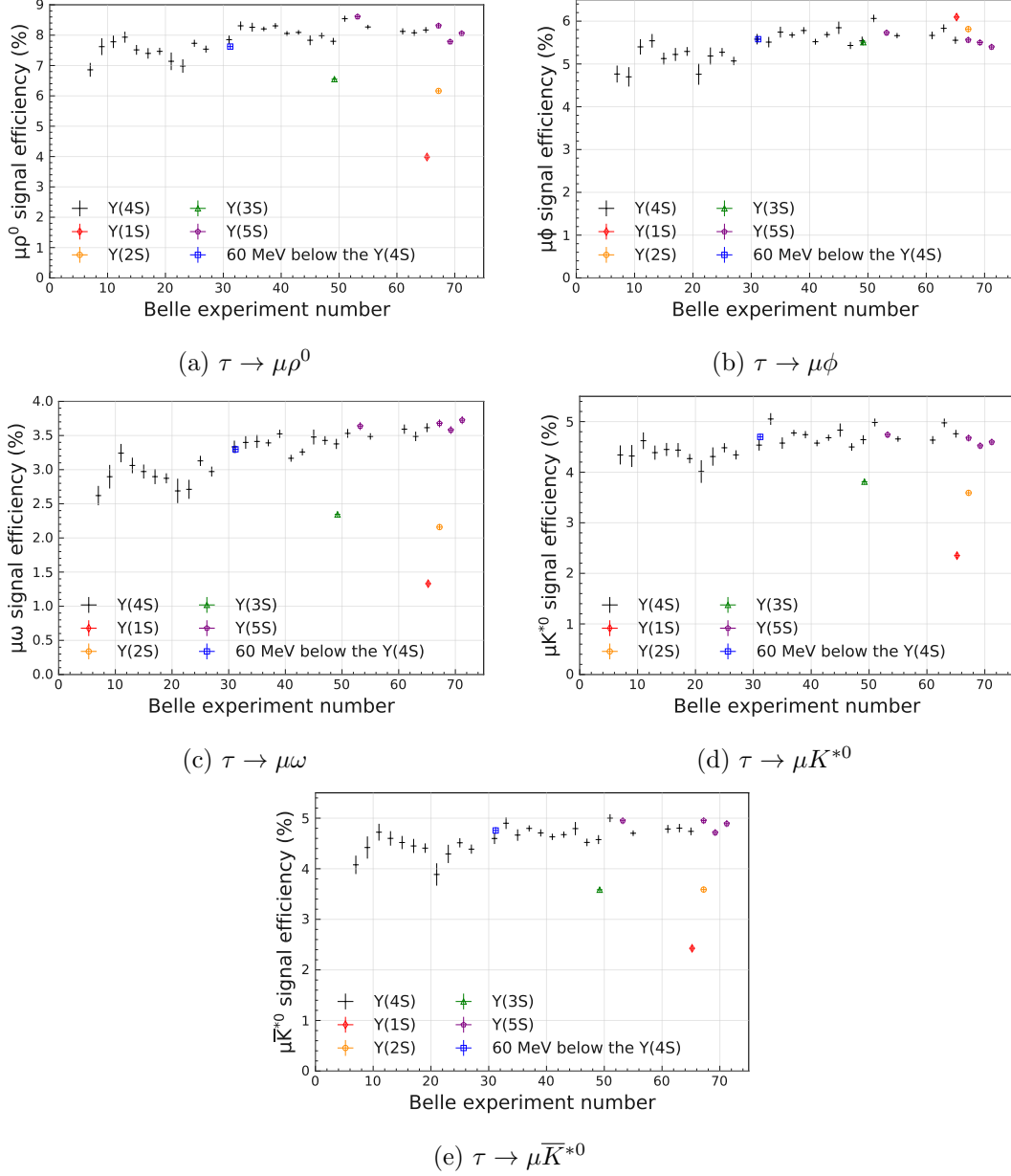


Figure 5.7: The μV^0 signal efficiencies for each Belle run period (the BDT selection). The legends show the c.m. energy of the collision. Some events are lost in Υ (1–3S) due to the BDT selection. No BDT cut for $\ell \phi$ modes because N_{BG} is enough small.

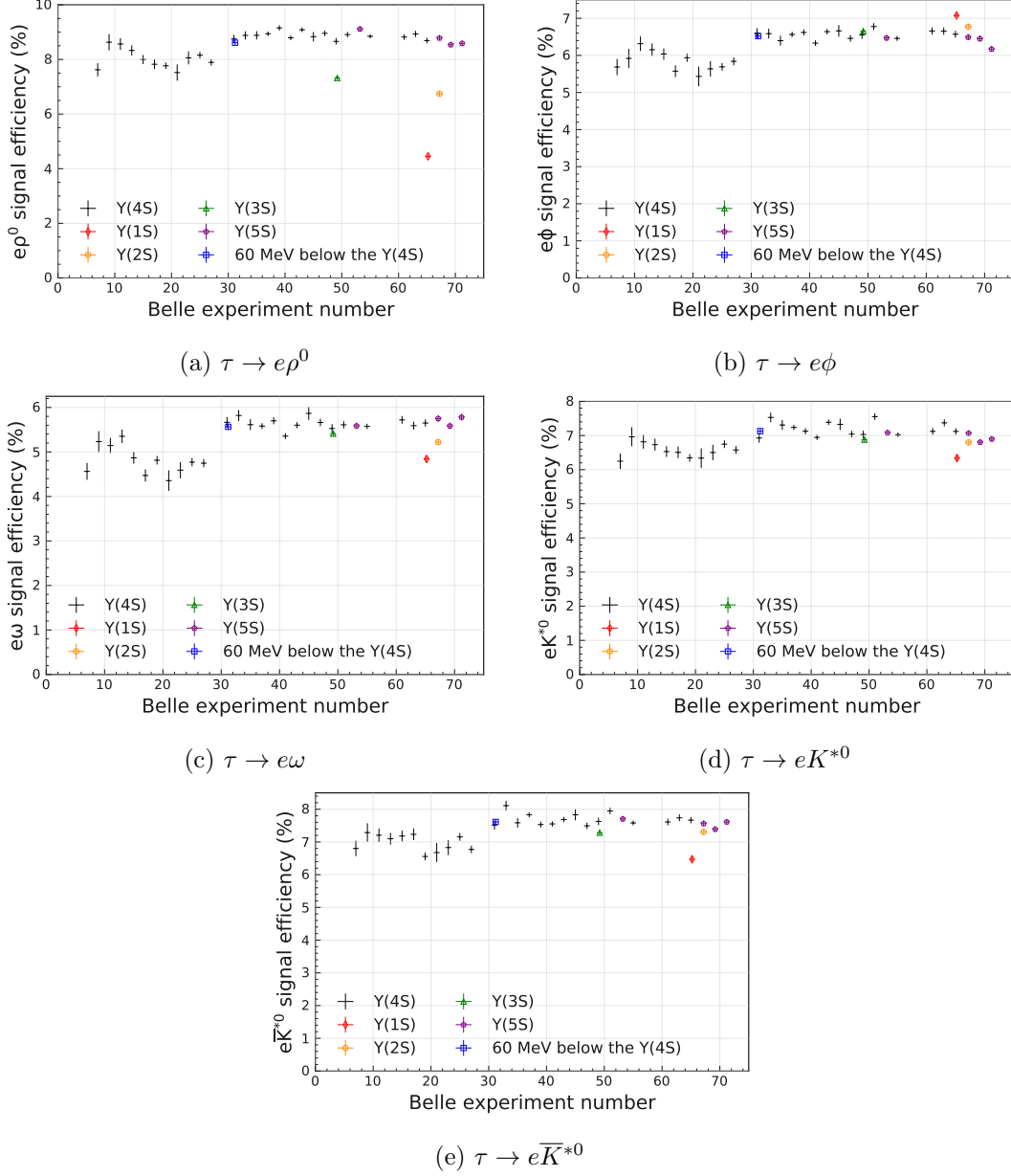


Figure 5.8: The eV^0 signal efficiencies for each Belle run period (the BDT selection). The legends show the c.m. energy of the collision. Some events are lost in Υ (1–3S) due to the BDT selection. No BDT cut for $\ell\phi$ modes because N_{BG} is enough small.

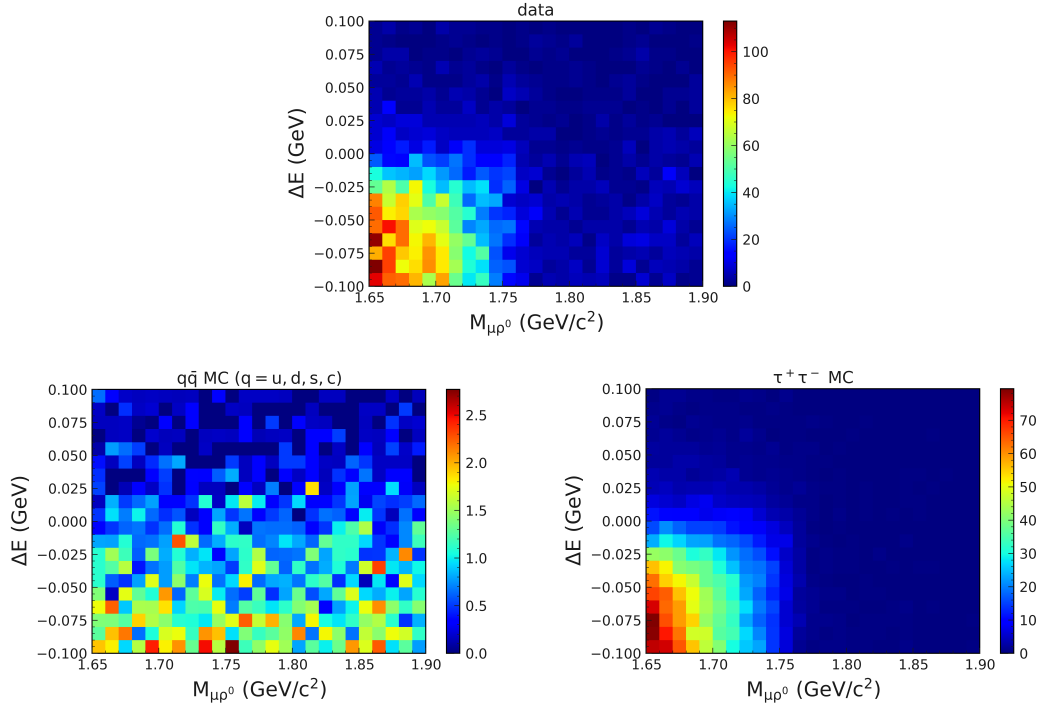


Figure 5.9: distribution of the $\tau \rightarrow \mu \rho^0$ control samples after the cut-base selection. Data samples (upper figure), $q\bar{q}$ MC ($q = u, d, s, c$, lower left), and $\tau^+ \tau^-$ MC (lower right).

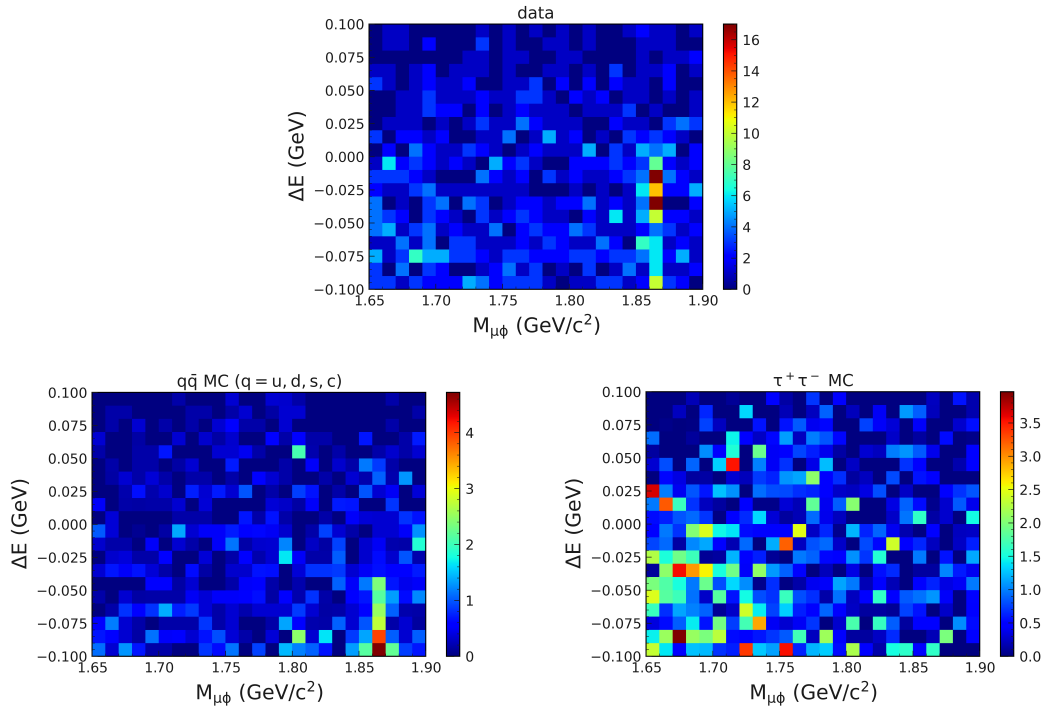


Figure 5.10: distribution of the $\tau \rightarrow \mu \phi$ control samples after the cut-base selection. Data samples (upper figure), $q\bar{q}$ MC ($q = u, d, s, c$, lower left), and $\tau^+ \tau^-$ MC (lower right).

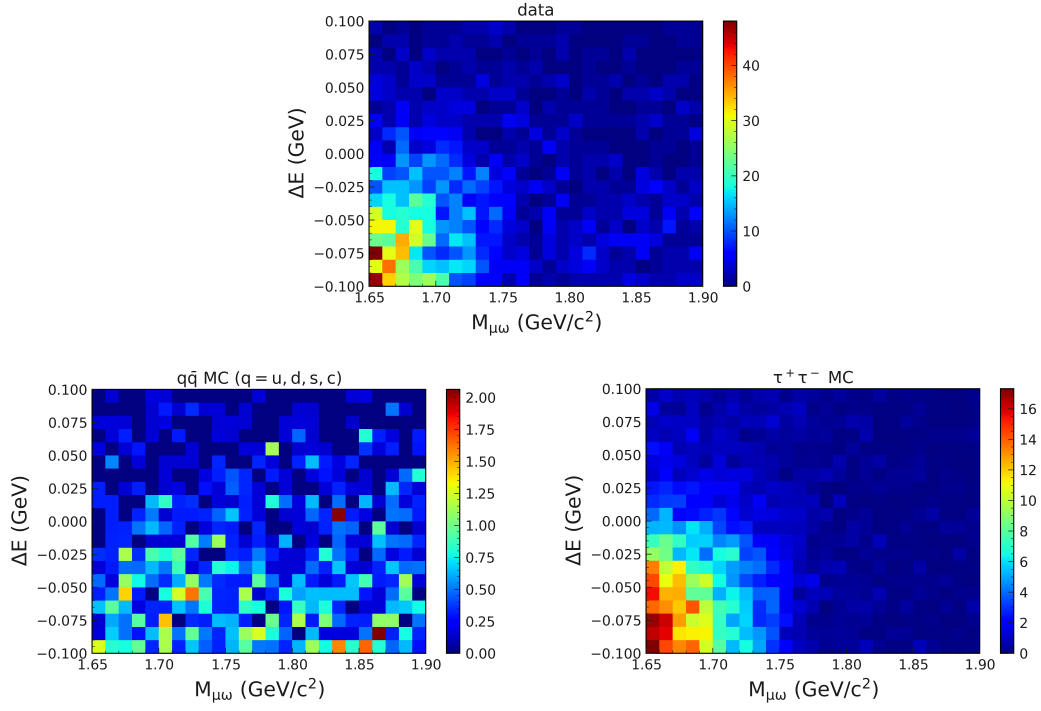


Figure 5.11: distribution of the $\tau \rightarrow \mu\omega$ control samples after the cut-base selection. Data samples (upper figure), $q\bar{q}$ MC ($q = u, d, s, c$, lower left), and $\tau^+\tau^-$ MC (lower right).

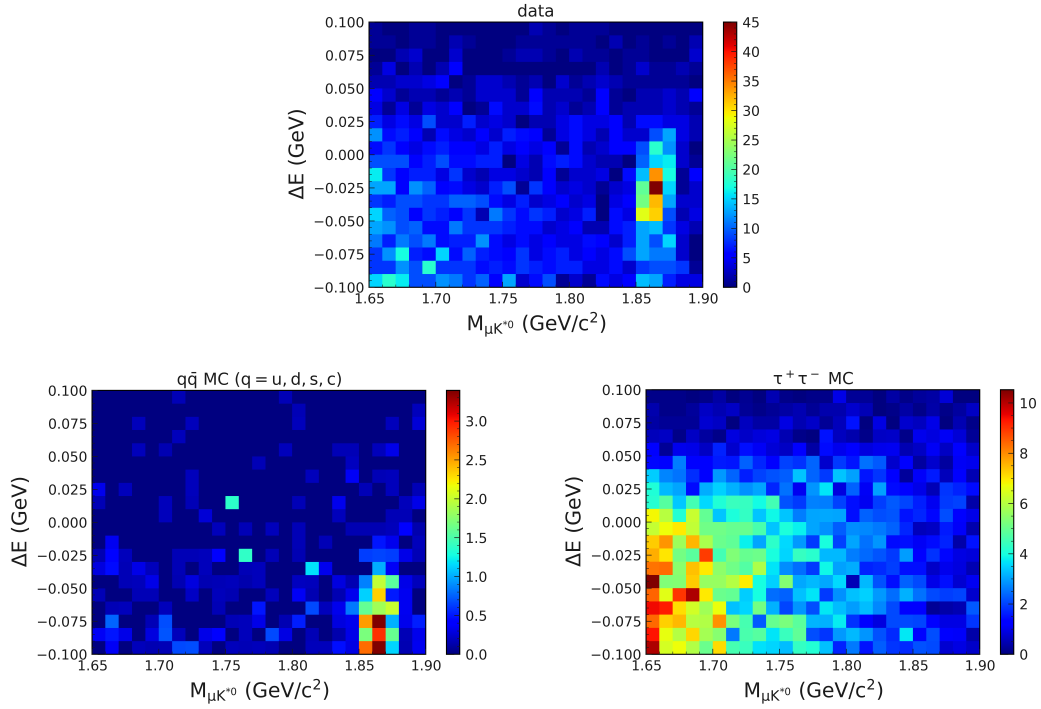


Figure 5.12: distribution of the $\tau \rightarrow \mu K^{*0}$ control samples after the cut-base selection. Data samples (upper figure), $q\bar{q}$ MC ($q = u, d, s, c$, lower left), and $\tau^+\tau^-$ MC (lower right).

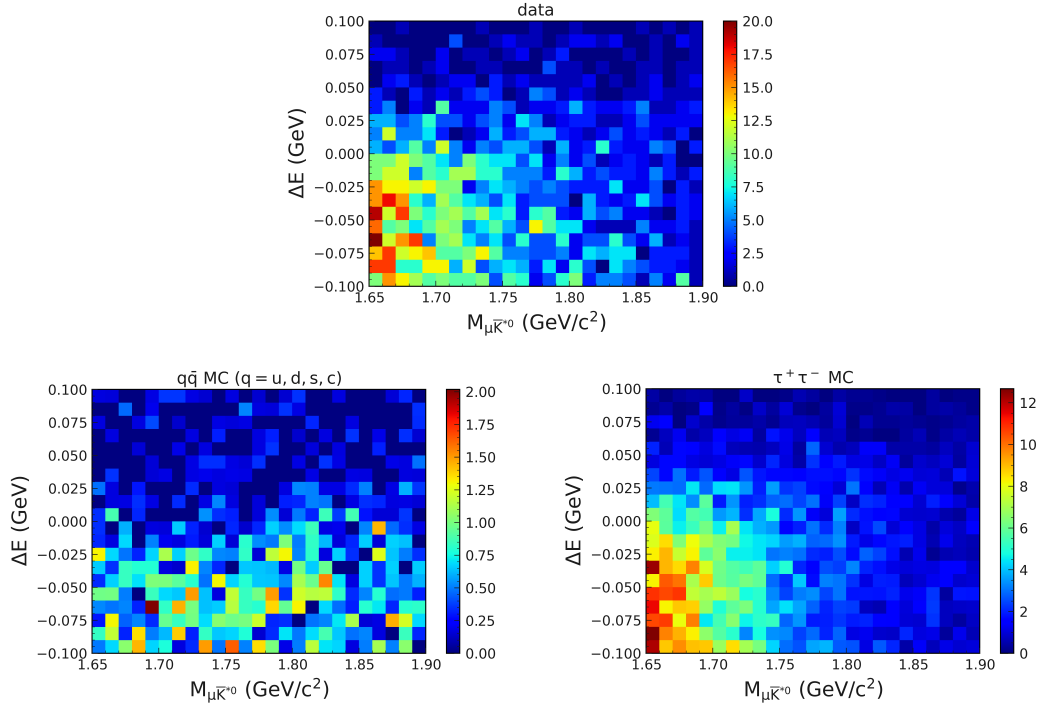


Figure 5.13: distribution of the $\tau \rightarrow \mu \bar{K}^{*0}$ control samples after the cut-base selection. Data samples (upper figure), $q\bar{q}$ MC ($q = u, d, s, c$, lower left), and $\tau^+\tau^-$ MC (lower right).

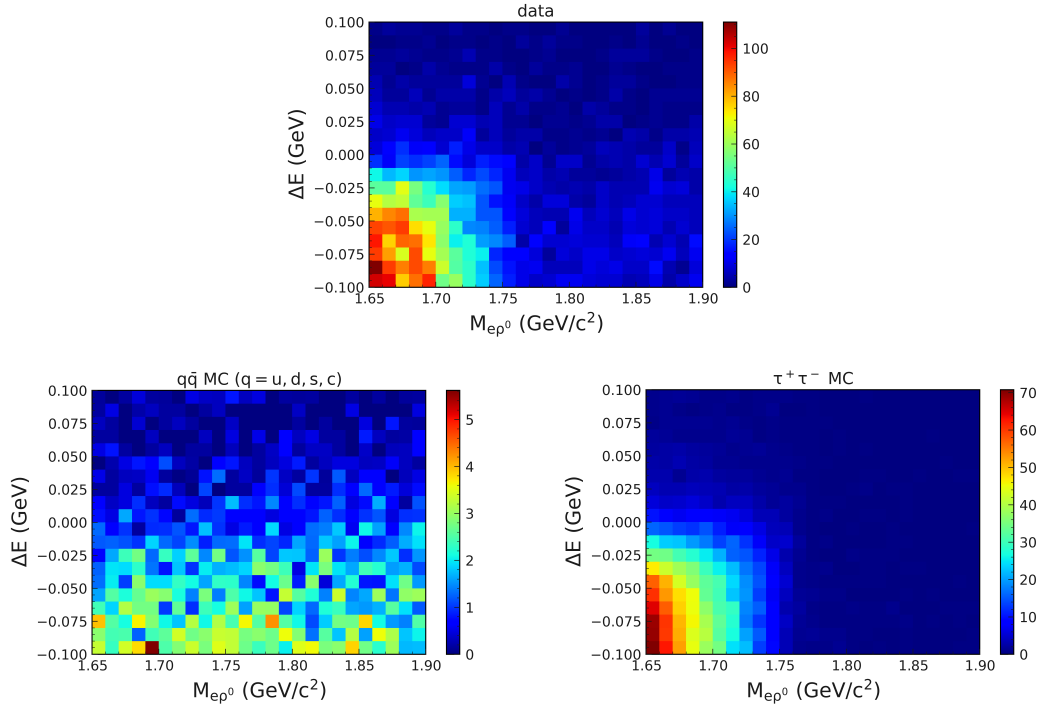


Figure 5.14: distribution of the $\tau \rightarrow e \rho^0$ control samples after the cut-base selection. Data samples (upper figure), $q\bar{q}$ MC ($q = u, d, s, c$, lower left), and $\tau^+\tau^-$ MC (lower right).

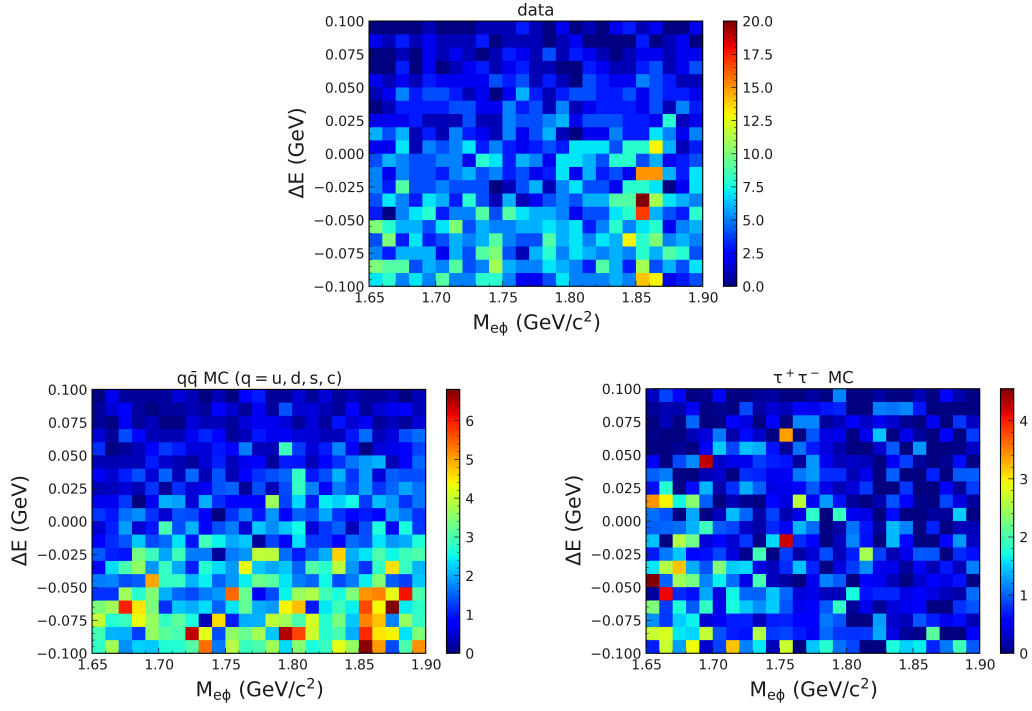


Figure 5.15: distribution of the $\tau \rightarrow e\phi$ control samples after the cut-base selection. Data samples (upper figure), $q\bar{q}$ MC ($q = u, d, s, c$, lower left), and $\tau^+\tau^-$ MC (lower right).

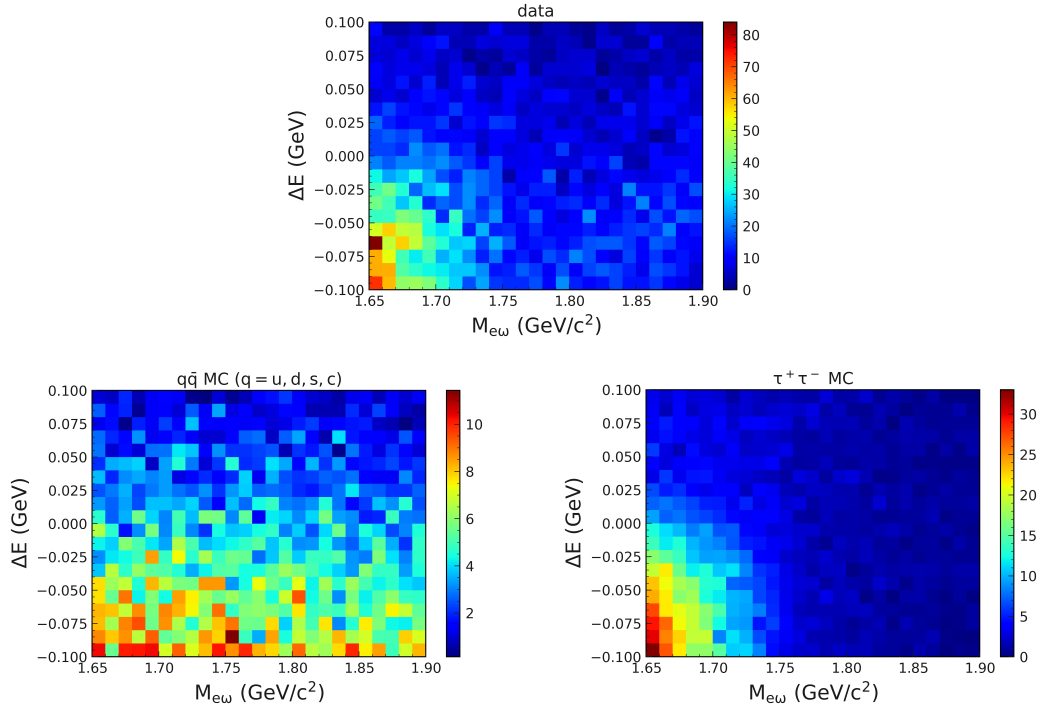


Figure 5.16: distribution of the $\tau \rightarrow e\omega$ control samples after the cut-base selection. Data samples (upper figure), $q\bar{q}$ MC ($q = u, d, s, c$, lower left), and $\tau^+\tau^-$ MC (lower right).

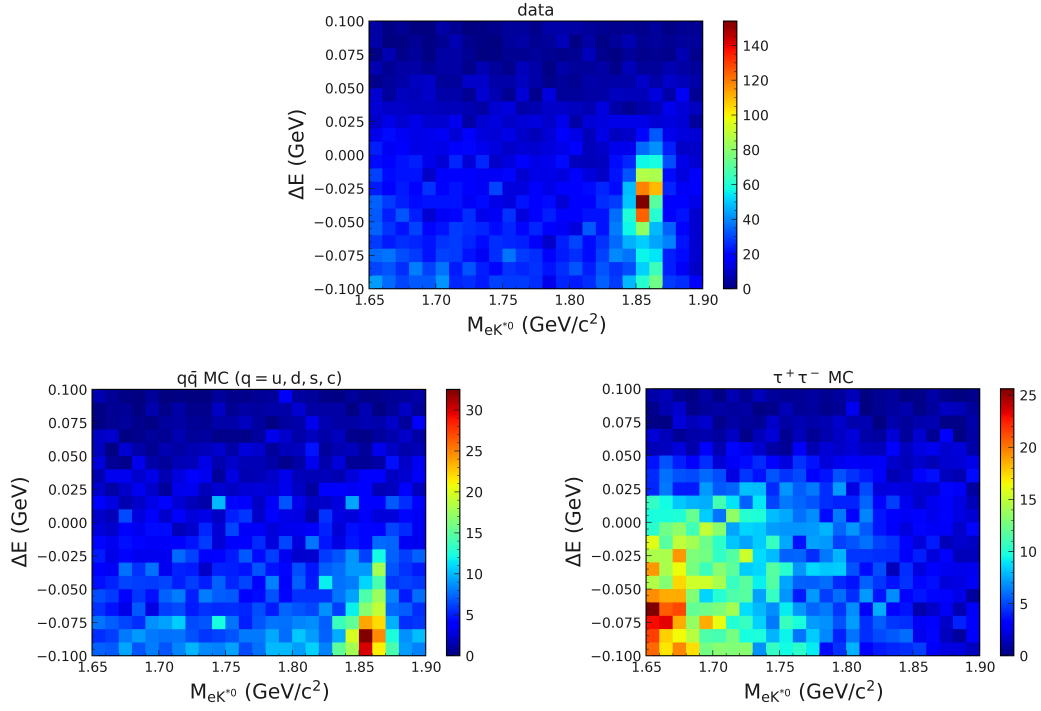


Figure 5.17: distribution of the $\tau \rightarrow e K^{*0}$ control samples after the cut-base selection. Data samples (upper figure), $q\bar{q}$ MC ($q = u, d, s, c$, lower left), and $\tau^+ \tau^-$ MC (lower right).

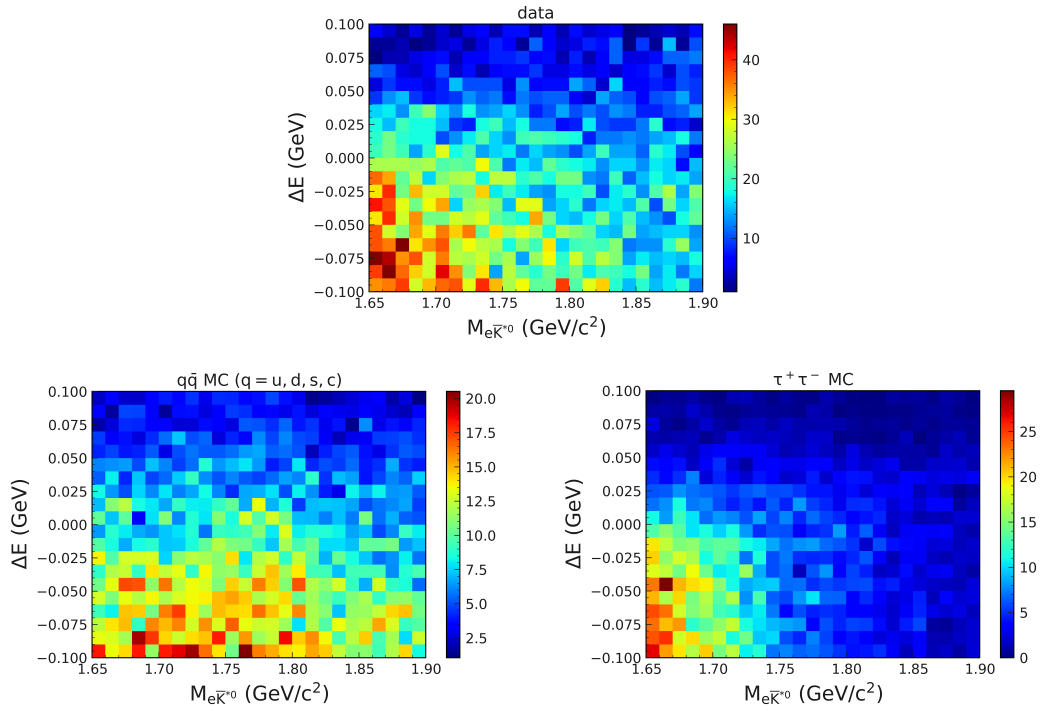
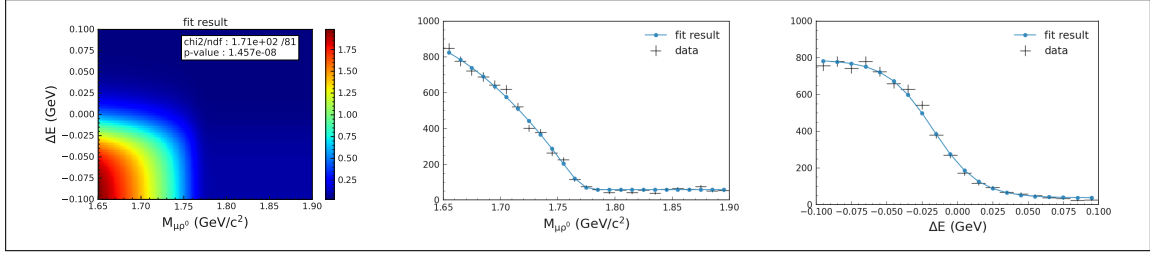
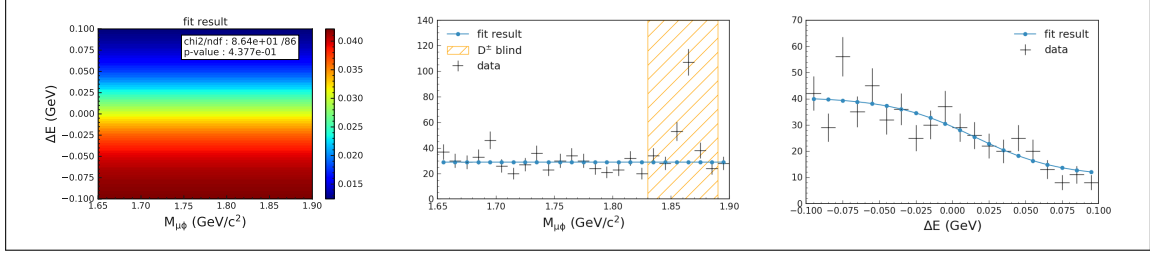


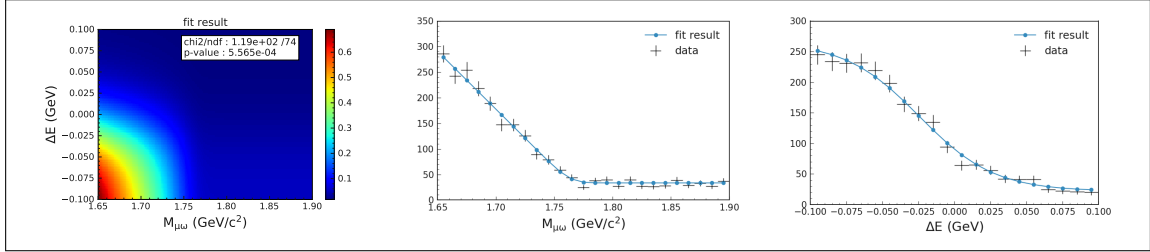
Figure 5.18: distribution of the $\tau \rightarrow e \bar{K}^{*0}$ control samples after the cut-base selection. Data samples (upper figure), $q\bar{q}$ MC ($q = u, d, s, c$, lower left), and $\tau^+ \tau^-$ MC (lower right).



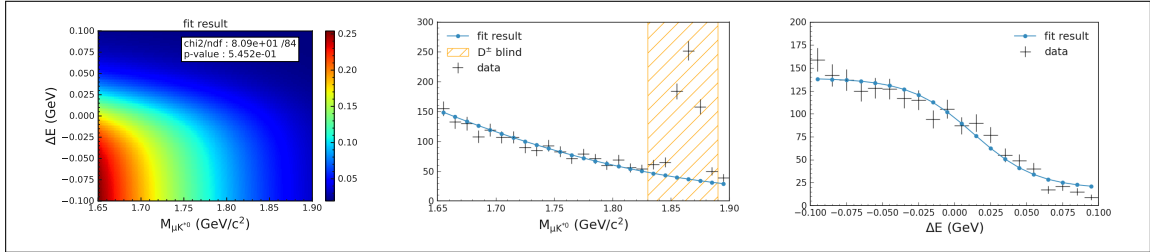
(a) $\tau \rightarrow \mu \rho^0$



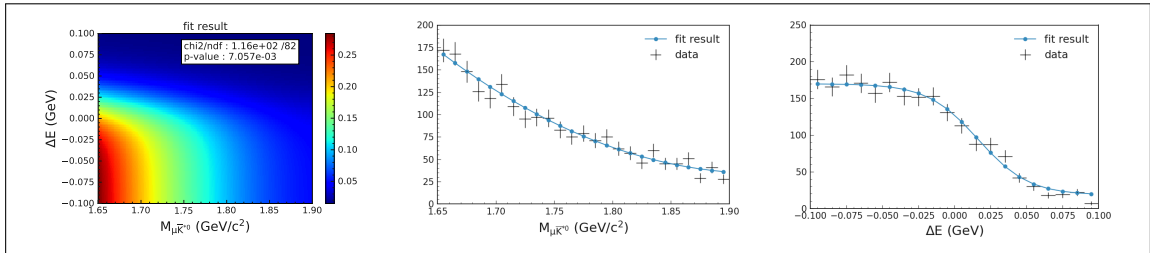
(b) $\tau \rightarrow \mu \phi$



(c) $\tau \rightarrow \mu \omega$

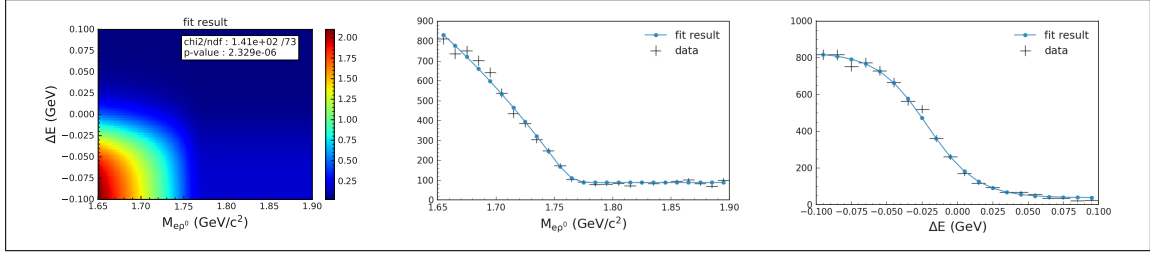


(d) $\tau \rightarrow \mu K^{*0}$

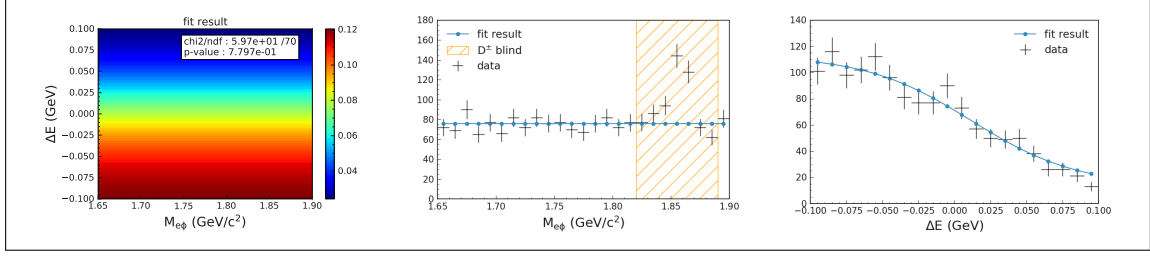


(e) $\tau \rightarrow \mu \bar{K}^{*0}$

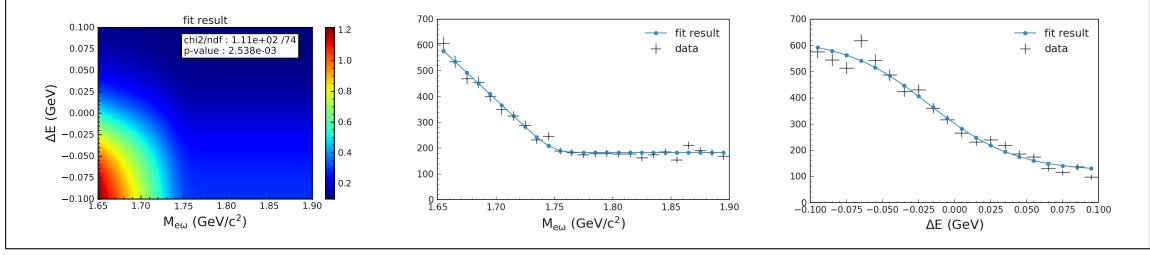
Figure 5.19: The fitting results to the μV^0 control samples (left). The projection plots to the $M_{\ell V^0}$ (middle) and ΔE (right). The orange hatched regions in the $\mu \phi$ and μK^{*0} modes are blinded to avoid the $D^\pm$ meson event peaks. They are removed from the fitting and the plot of ΔE projection.



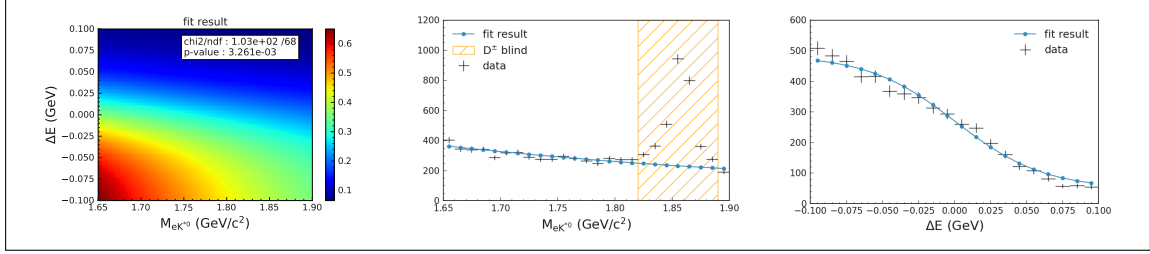
(a) $\tau \rightarrow e\rho^0$



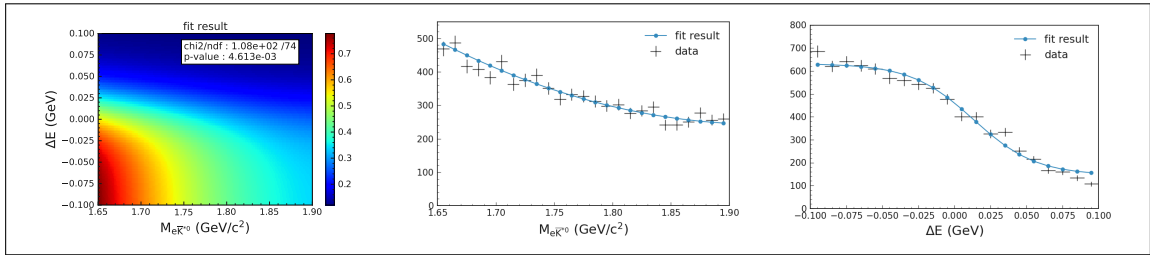
(b) $\tau \rightarrow e\phi$



(c) $\tau \rightarrow e\omega$

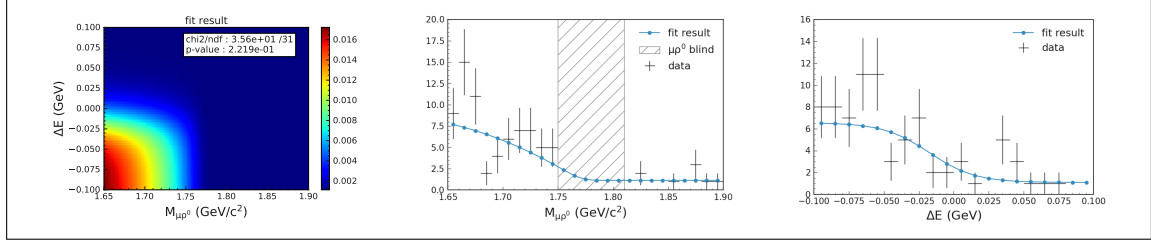


(d) $\tau \rightarrow eK^{*0}$

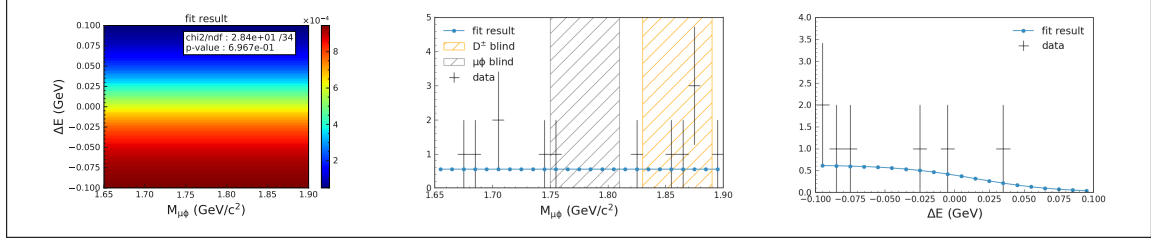


(e) $\tau \rightarrow e\bar{K}^{*0}$

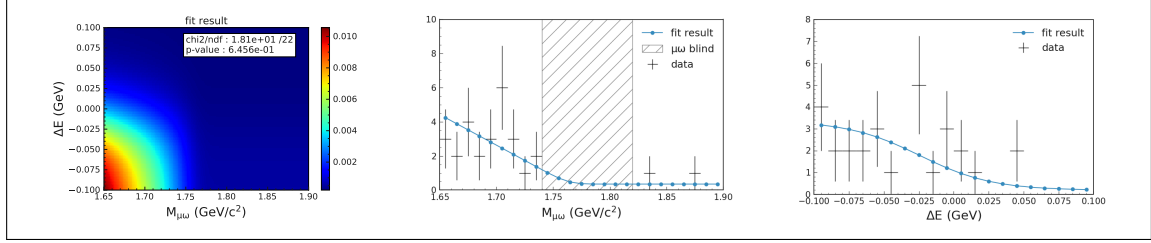
Figure 5.20: The fitting results to the eV^0 control samples (left). The projection plots to the M_{eV^0} (middle) and ΔE (right). The orange hatched regions in the $e\phi$ and eK^{*0} modes are blinded to avoid the $D^\pm$ meson event peaks. They are removed from the fitting and the plot of ΔE projection.



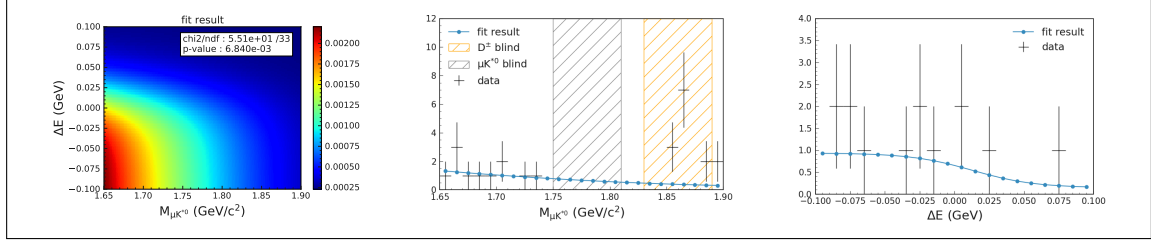
(a) $\tau \rightarrow \mu\rho^0$



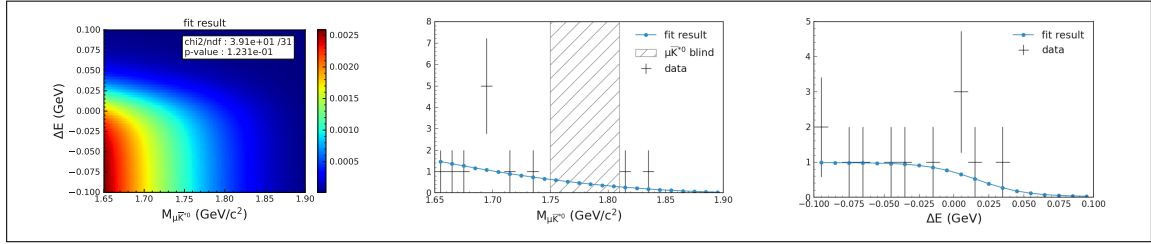
(b) $\tau \rightarrow \mu\phi$



(c) $\tau \rightarrow \mu\omega$

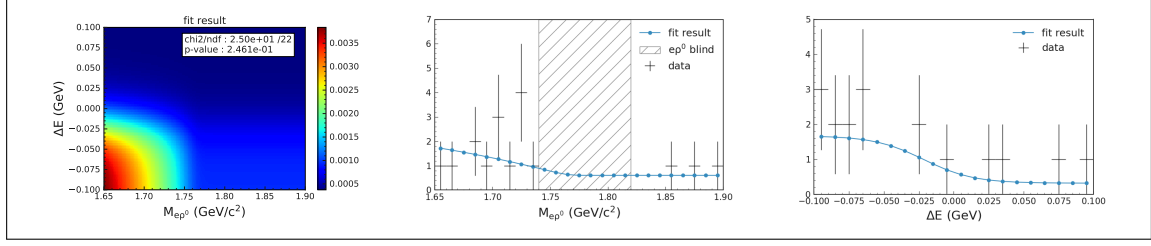


(d) $\tau \rightarrow \mu K^{*0}$

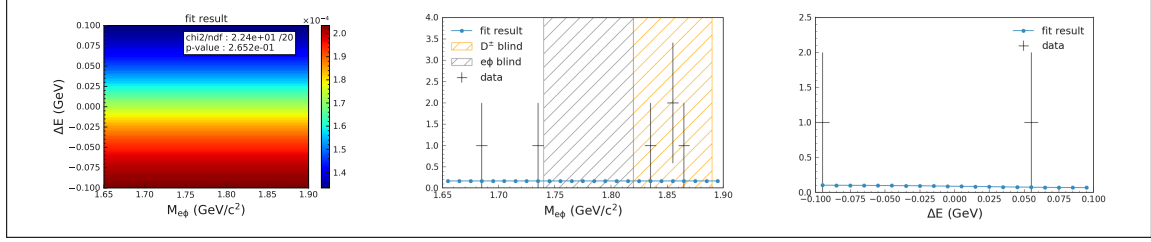


(e) $\tau \rightarrow \mu\bar{K}^{*0}$

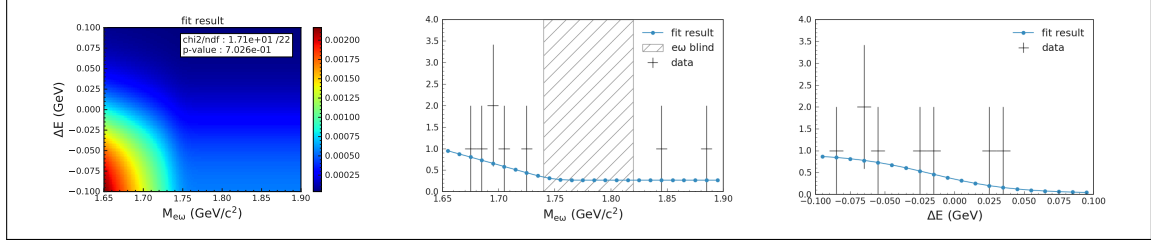
Figure 5.21: The fitting results to the μV^0 sideband data samples after the cut-base selection (left). The projection plots to the $M_{\ell V^0}$ (middle) and ΔE (right). The gray hatched regions are the blind regions for the signal events. The orange hatched regions in the $\mu\phi$ and μK^{*0} modes are blinded to avoid the $D^\pm$ meson event peaks. Both hatched regions are removed from the fitting and the plot of ΔE projection.



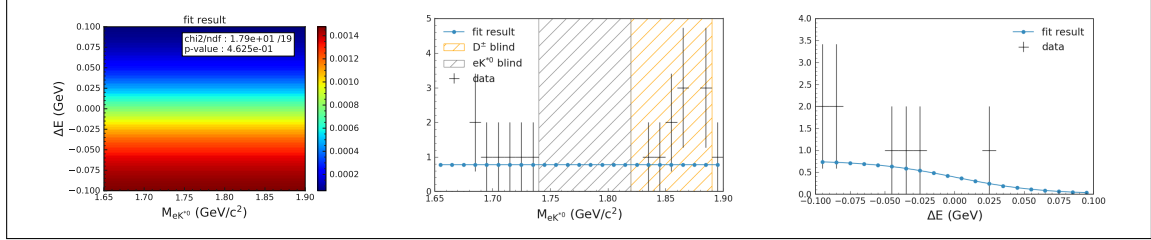
(a) $\tau \rightarrow e\rho^0$



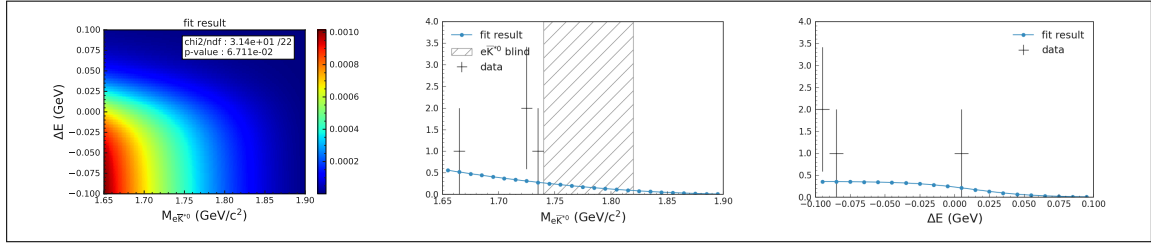
(b) $\tau \rightarrow e\phi$



(c) $\tau \rightarrow e\omega$

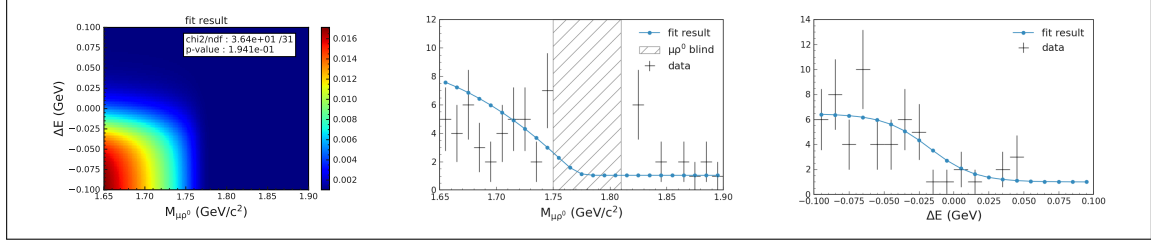


(d) $\tau \rightarrow eK^{*0}$

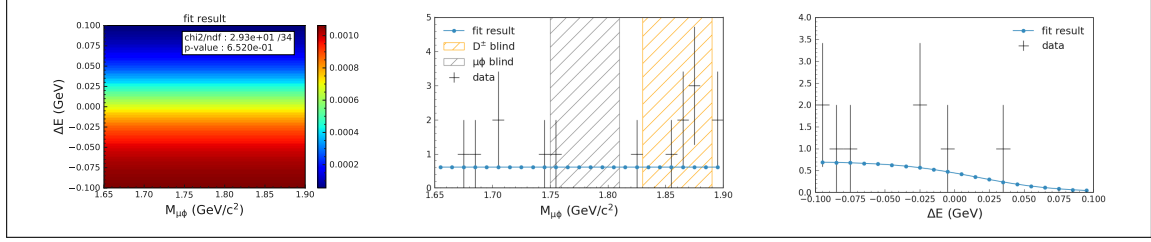


(e) $\tau \rightarrow e\bar{K}^{*0}$

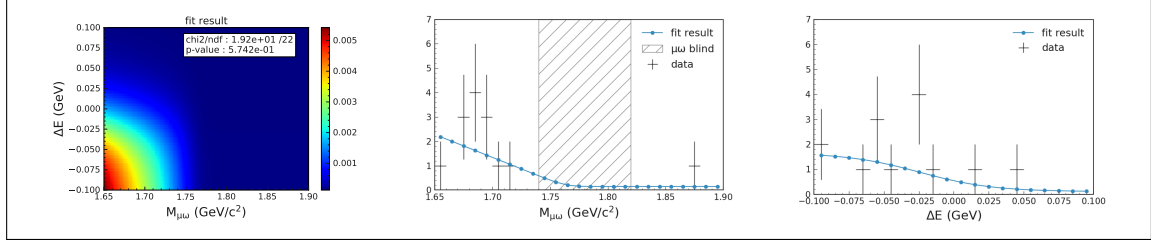
Figure 5.22: The fitting results to the eV^0 sideband data samples after the cut-base selection (left). The projection plots to the $M_{\ell V^0}$ (middle) and ΔE (right). The gray hatched regions are the blind regions for the signal events. The orange hatched regions in the $e\phi$ and eK^{*0} modes are blinded to avoid the $D^\pm$ meson event peaks. Both hatched regions are removed from the fitting and the plot of ΔE projection.



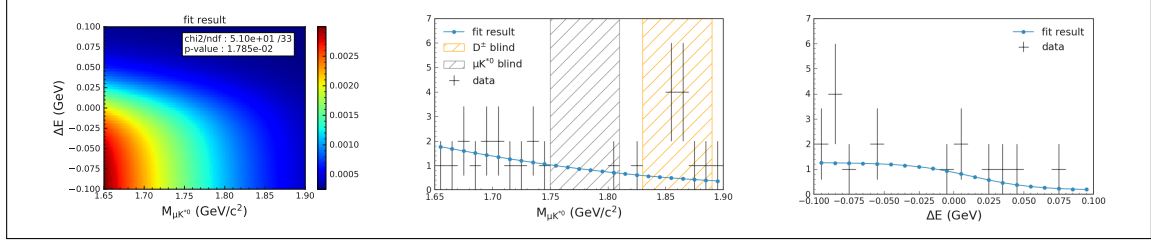
(a) $\tau \rightarrow \mu\rho^0$



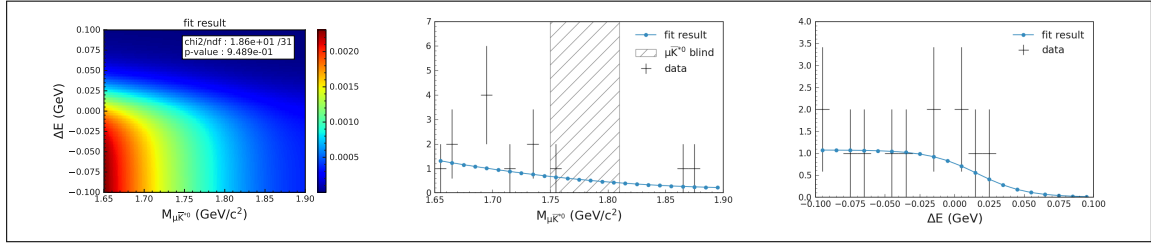
(b) $\tau \rightarrow \mu\phi$



(c) $\tau \rightarrow \mu\omega$

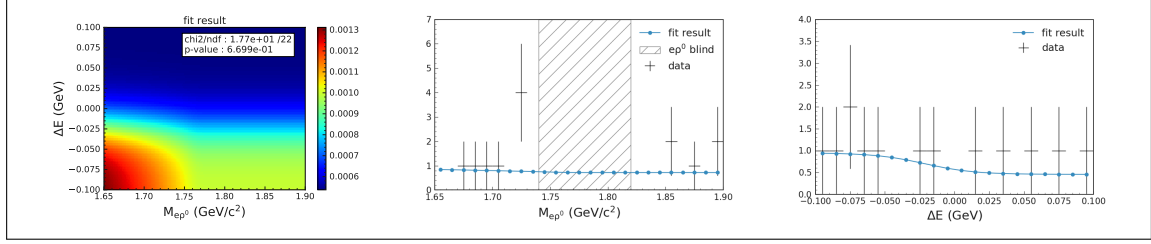


(d) $\tau \rightarrow \mu K^{*0}$

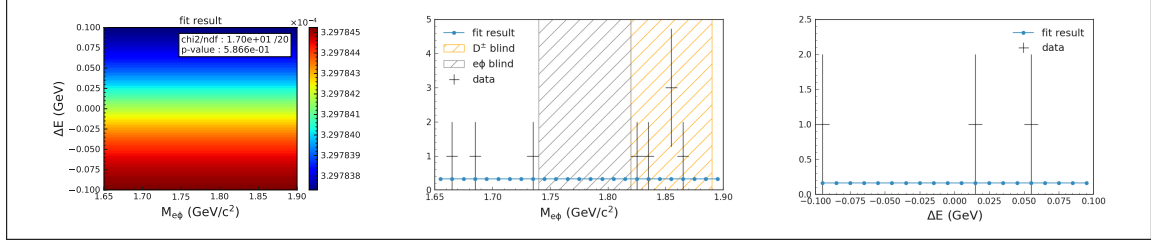


(e) $\tau \rightarrow \mu\bar{K}^{*0}$

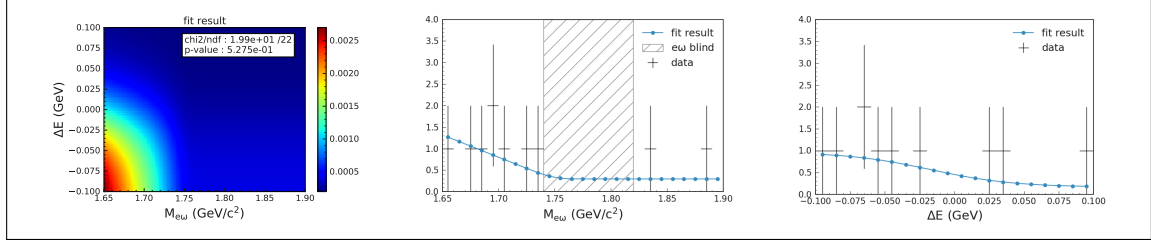
Figure 5.23: The fitting results to the μV^0 sideband data samples after the BDT selection (left). The projection plots to the $M_{\mu V^0}$ (middle) and ΔE (right). The gray hatched regions are the blind regions for the signal events. The orange hatched regions in the $\mu\phi$ and μK^{*0} modes are blinded to avoid the $D^\pm$ meson event peaks. Both hatched regions are removed from the fitting and the plot of ΔE projection.



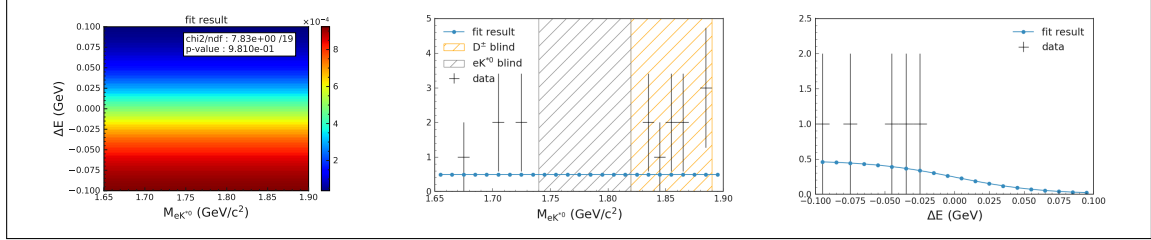
(a) $\tau \rightarrow e\rho^0$



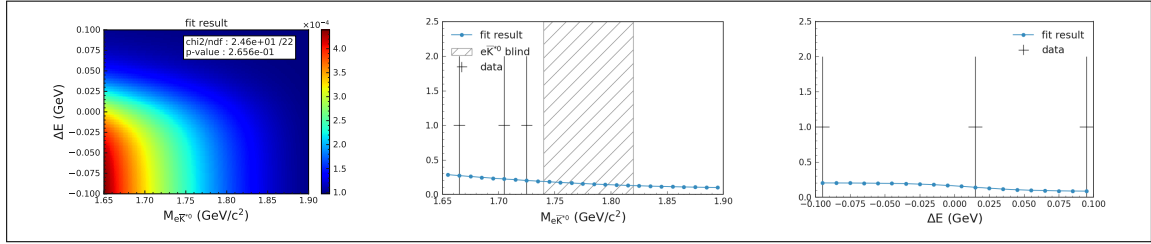
(b) $\tau \rightarrow e\phi$



(c) $\tau \rightarrow e\omega$



(d) $\tau \rightarrow eK^{*0}$



(e) $\tau \rightarrow e\bar{K}^{*0}$

Figure 5.24: The fitting results to the eV^0 sideband data samples after the BDT selection (left). The projection plots to the M_{eV^0} (middle) and ΔE (right). The gray hatched regions are the blind regions for the signal events. The orange hatched regions in the $e\phi$ and eK^{*0} modes are blinded to avoid the $D^\pm$ meson event peaks. Both hatched regions are removed from the fitting and the plot of ΔE projection.

Chapter 6

Results

I get better signal efficiency by the BDT selection (section 4.4) than by the cut-base selection (section 4.3), while the expected number of background events (N_{BG}) are similar values. Compared with the previous Belle paper [18], the signal efficiency is improved by 40% on the average of ten ℓV^0 modes for the BDT selection, whereas the average N_{BG} increases 9.5%. Here, the N_{BG} in the previous paper is scaled to 980 fb^{-1} .

I apply the $\tau \rightarrow \ell V^0$ reconstruction and event selection for the data samples inside the signal region. Figures 6.1–6.4 are the observed data samples after the cut-base selection and BDT selection. The observed number of events in the signal region (N_{obs}) has no excess over N_{BG} in section 5.2.2.

6.1 LFV decay branching fraction upper limits

I set 90% C.L. upper limits on the branching fractions based on a Bayesian method using Bayesian Analysis Toolkit (BAT) [83]. BAT is based on the Markov Chain Monte Carlo method to search the probability density function (PDF) of each parameters. The PDF of the branching fraction ($\mathcal{B}(\tau \rightarrow \ell V^0)$) is calculated assuming that N_{obs} follows a poisson distribution function whose mean value is the expected number of events (N_{exp}),

$$N_{\text{exp}} = L \times 2\sigma_{\tau\tau} \mathcal{B}(\tau \rightarrow \ell V^0) \times \varepsilon + N_{\text{BG}}, \quad (6.1)$$

where L is the integrated luminosity ($980.4 \pm 13.7 \text{ fb}^{-1}$), $\sigma_{\tau\tau}$ is the cross section of τ -pair production that is calculated with KKMC [84] (the weighted average of $\sigma_{\tau\tau}$ at all the beam energies is $0.916 \pm 0.003 \text{ nb}$) and ε is the signal efficiency including the branching fraction of the V^0 . I assume that these values ($L, \sigma_{\tau\tau}, \varepsilon$ and N_{BG}) follow a gaussian distribution of sigma equal to the uncertainty of each value. The lower limits of N_{BG} are set as zero. I get 90% C.L. upper limits of the $\mathcal{B}(\tau \rightarrow \ell V^0)$ PDF. The expected upper limits ($\mathcal{B}_{\text{exp}}$) are the values when $N_{\text{obs}} = N_{\text{BG}}$.

The upper limits on $\mathcal{B}(\tau \rightarrow \ell V^0)$ with the cut-base selection are listed in table 6.1. The average of the observed (expected) limits is better than that of the previous paper using 854 fb^{-1} [18] by 23(31)%. The limits with the BDT selection are listed in table 6.2. The average of the observed (expected) limits is better than that of the previous paper by 30(34)%. I choose the BDT selection for our final result. The upper limits comparison between the previous paper and the BDT selection is shown in figure 6.5. All expected upper limits are improved, and eight observed upper limits are improved.

Table 6.1: Results of the cut-base selection. The signal efficiency (ε), the expected number of background events (N_{BG}), total systematic uncertainty on the expected number of signal events (σ_{syst}), the expected 90% C.L. upper limits ($\mathcal{B}_{\text{exp}}$ (10^{-8})), the number of observed events in the signal region (N_{obs}), the observed 90% C.L. upper limits ($\mathcal{B}_{\text{obs}}$ (10^{-8})) and the expected upper limits of previous study with 854 fb^{-1} ($\mathcal{B}_{\text{exp}}^{\text{prev}}$).

Mode	ε (%)	N_{BG}	σ_{syst} (%)	$\mathcal{B}_{\text{exp}}$	N_{obs}	$\mathcal{B}_{\text{obs}}$	$\mathcal{B}_{\text{exp}}^{\text{prev}}$
$\tau^- \rightarrow \mu^- \rho^0$	7.40	$1.03 \pm 0.22(\text{stat}) \pm 0.13(\text{syst})$	4.9	2.5	0	1.8	3.1
$\tau^- \rightarrow e^- \rho^0$	8.07	$0.68 \pm 0.25(\text{stat}) \pm 0.01(\text{syst})$	4.5	2.1	0	1.6	2.3
$\tau^- \rightarrow \mu^- \phi$	5.37	$0.42 \pm 0.12(\text{stat}) \pm 0.05(\text{syst})$	4.8	2.9	0	2.4	4.9
$\tau^- \rightarrow e^- \phi$	6.18	$0.20 \pm 0.14(\text{stat}) \pm 0.00(\text{syst})$	4.6	2.3	0	2.1	4.3
$\tau^- \rightarrow \mu^- \omega$	3.11	$0.75 \pm 0.36(\text{stat}) \pm 0.07(\text{syst})$	5.1	5.7	0	4.2	7.9
$\tau^- \rightarrow e^- \omega$	5.15	$0.71 \pm 0.39(\text{stat}) \pm 0.05(\text{syst})$	4.5	3.4	0	2.5	5.9
$\tau^- \rightarrow \mu^- K^{*0}$	4.30	$0.65 \pm 0.20(\text{stat}) \pm 0.20(\text{syst})$	4.5	4.0	0	3.0	5.4
$\tau^- \rightarrow e^- K^{*0}$	6.61	$0.86 \pm 0.27(\text{stat}) \pm 0.04(\text{syst})$	4.1	2.7	2	3.9	3.9
$\tau^- \rightarrow \mu^- \bar{K}^{*0}$	4.36	$0.49 \pm 0.15(\text{stat}) \pm 0.12(\text{syst})$	4.6	3.7	0	3.0	4.9
$\tau^- \rightarrow e^- \bar{K}^{*0}$	7.09	$0.31 \pm 0.15(\text{stat}) \pm 0.03(\text{syst})$	4.1	2.1	1	2.9	3.6

Table 6.2: Results of the BDT selection. The signal efficiency (ε), the expected number of background events (N_{BG}), total systematic uncertainty on the expected number of signal events (σ_{syst}), the expected 90% C.L. upper limits ($\mathcal{B}_{\text{exp}}$ (10^{-8})), the number of observed events in the signal region (N_{obs}), the observed 90% C.L. upper limits ($\mathcal{B}_{\text{obs}}$ (10^{-8})), the number of observed events in the previous study ($N_{\text{obs}}^{\text{prev}}$) and the observed upper limits of previous study with 854 fb^{-1} ($\mathcal{B}_{\text{obs}}^{\text{prev}}$).

Mode	ε (%)	N_{BG}	σ_{syst} (%)	$\mathcal{B}_{\text{exp}}$	N_{obs}	$\mathcal{B}_{\text{obs}}$	$N_{\text{obs}}^{\text{prev}}$	$\mathcal{B}_{\text{obs}}^{\text{prev}}$
$\tau^- \rightarrow \mu^- \rho^0$	7.78	$0.95 \pm 0.20(\text{stat}) \pm 0.11(\text{syst})$	4.6	2.3	0	1.7	0	1.2
$\tau^- \rightarrow e^- \rho^0$	8.49	$0.80 \pm 0.27(\text{stat}) \pm 0.02(\text{syst})$	4.4	2.1	1	2.2	0	1.8
$\tau^- \rightarrow \mu^- \phi$	5.59	$0.47 \pm 0.15(\text{stat}) \pm 0.05(\text{syst})$	4.8	2.9	0	2.3	1	8.4
$\tau^- \rightarrow e^- \phi$	6.45	$0.38 \pm 0.21(\text{stat}) \pm 0.00(\text{syst})$	4.5	2.4	0	2.0	0	3.1
$\tau^- \rightarrow \mu^- \omega$	3.27	$0.32 \pm 0.23(\text{stat}) \pm 0.03(\text{syst})$	4.8	4.7	0	3.9	0	4.7
$\tau^- \rightarrow e^- \omega$	5.41	$0.74 \pm 0.43(\text{stat}) \pm 0.01(\text{syst})$	4.5	3.2	0	2.4	0	4.8
$\tau^- \rightarrow \mu^- K^{*0}$	4.52	$0.84 \pm 0.25(\text{stat}) \pm 0.03(\text{syst})$	4.3	4.0	0	2.9	1	7.2
$\tau^- \rightarrow e^- K^{*0}$	6.94	$0.54 \pm 0.21(\text{stat}) \pm 0.12(\text{syst})$	4.1	2.4	0	1.9	0	3.2
$\tau^- \rightarrow \mu^- \bar{K}^{*0}$	4.58	$0.58 \pm 0.17(\text{stat}) \pm 0.06(\text{syst})$	4.3	3.6	1	4.2	1	7.0
$\tau^- \rightarrow e^- \bar{K}^{*0}$	7.45	$0.25 \pm 0.11(\text{stat}) \pm 0.01(\text{syst})$	4.1	2.0	0	1.7	0	3.4

6.2 Discussion

6.2.1 Improvements from the previous paper

Compared with the previous paper [18], the signal efficiency is improved by 40% on the average of ten ℓV^0 modes for the BDT selection, whereas the average N_{BG} increases 9.5%. The observed $\mathcal{B}(\tau \rightarrow \ell V^0)$ upper limits are improved 30% on average. The improvements come from the followings: (1) The integrated luminosity increases 15%; (2) Additional reconstructions of $\tau_{\text{tag}}^\pm \rightarrow \pi^\pm \pi^+ \pi^- \nu$, $\pi^\pm \pi^0 \pi^0 \nu$ increase the signal efficiency by 9.6%; (3) reconsideration of the event selection criteria with the new variables (M_ν^2 , $P_\nu^{c.m.}$); (4) A multivariate analysis using BDT is introduced to the event selection; and (5) The total number of the observed events reduced from three to two. Although the numbers of background events are also by the first and second items, the third and fourth items reduce them. The last item is just a statistical effect.

On the other hand, the observed upper limits on $\mathcal{B}(\tau \rightarrow \ell \rho^0)$ are worse than those of the previous results, though the expected upper limits before unblinding are better than those of the

previous Belle paper. In the case of the $e\rho^0$ mode, it is just due to the one event observation in this paper. In the case of the $\mu\rho^0$ mode, the upper limit calculation method is affected. I use the Bayesian limits instead of the Frequentist limits [85] that was used in the previous Belle paper. The Frequentist limits are negatively proportional to N_{BG} when N_{obs} is fixed. In the previous paper of the $\mu\rho^0$ mode, the upper limit was more stringent than the expected one because N_{obs} was zero, whereas N_{BG} was 1.48.

The improvements in observed upper limits are 4% smaller than that of expected. This is also due to the difference between the Bayesian limits and the Frequentist limits. Table 6.3 is the results of the previous paper and re-calculated upper limits using BAT with the items in the table. The expected upper limits are similar between the Bayesian and Frequentist method, whereas some of the observed upper limits have gaps. Compared with the Bayesian upper limits of the previous paper, the upper limits in this paper (BDT selection, observed $\mathcal{B}$) are relatively improved by 40% on average. This is larger than that of expected due to the slightly larger N_{BG} on average and the reduction of the number of total N_{obs} within the statistical error.

Table 6.3: Comparison of Bayesian method and frequentist method to calculate upper limits using the previous paper results. The signal efficiency (ε), the expected number of background events (N_{BG}), total systematic uncertainty (σ_{syst}), the expected 90% C.L. upper limits ($\mathcal{B}_{\text{exp}}$ (10^{-8})), the number of observed events in the signal region (N_{obs}) and the observed 90% C.L. upper limits ($\mathcal{B}_{\text{obs}}$ (10^{-8})) are from previous paper. The last two columns are re-calculated expected and observed 90% C.L. upper limits using Bayesian analysis toolkit ($\mathcal{B}_{\text{exp}}^{\text{Bayesian}}$ and $\mathcal{B}_{\text{obs}}^{\text{Bayesian}}$ (10^{-8})).

Mode	ε (%)	N_{BG}	σ_{syst} (%)	$\mathcal{B}_{\text{exp}}$	N_{obs}	$\mathcal{B}_{\text{obs}}$	$\mathcal{B}_{\text{exp}}^{\text{Bayesian}}$	$\mathcal{B}_{\text{obs}}^{\text{Bayesian}}$
$\tau^- \rightarrow \mu^- \rho^0$	7.09	1.48 ± 0.35	5.3	3.1	0	1.2	3.3	2.1
$\tau^- \rightarrow e^- \rho^0$	7.58	0.29 ± 0.15	5.4	2.3	0	1.8	2.3	2.0
$\tau^- \rightarrow \mu^- \phi$	3.21	0.06 ± 0.06	5.8	4.9	1	8.4	4.8	7.7
$\tau^- \rightarrow e^- \phi$	4.18	0.47 ± 0.19	5.9	4.3	0	3.1	4.4	3.5
$\tau^- \rightarrow \mu^- \omega$	2.38	0.72 ± 0.18	6.1	7.9	0	4.7	8.4	6.2
$\tau^- \rightarrow e^- \omega$	2.92	0.30 ± 0.14	6.2	5.9	0	4.8	5.9	5.1
$\tau^- \rightarrow \mu^- K^{*0}$	3.39	0.53 ± 0.20	5.5	5.4	1	7.2	5.5	6.6
$\tau^- \rightarrow e^- K^{*0}$	4.37	0.29 ± 0.14	5.6	3.9	0	3.2	4.0	3.4
$\tau^- \rightarrow \mu^- \bar{K}^{*0}$	3.60	0.45 ± 0.17	5.5	4.9	1	7.0	5.1	6.3
$\tau^- \rightarrow e^- \bar{K}^{*0}$	4.41	0.08 ± 0.08	5.6	3.6	0	3.4	3.6	3.4

6.2.2 Prospect for the future Belle II experiment

The limit for the $\tau \rightarrow \mu\phi$ (2.3×10^{-8} , improved by 73%) sets more constraint on the single leptoquark model [36]. The other limits for the $\tau \rightarrow \ell V^0$ branching fractions (except $V^0 = \rho^0$) also set more constraint on different NP models [44–47]. Although the allowed region still remains, the further $\tau \rightarrow \ell V^0$ search can be performed by Belle II experiment. The integrated luminosity of 50 ab^{-1} is scheduled to be collected, 50 times as large as Belle experiment. That can provide a sensitivity from 7 times to a maximum of 50 times better than this paper, depending on the N_{BG} suppression. I calculate the expected upper limits in Belle II experiment conservatively. The integrated luminosity and the number of background events are 50 times larger than this paper. To reduce the background events, I use a tight BDT selection and evaluate the signal efficiency and the number of background events with the Belle data set.

Figure 6.6 shows the $\mathcal{B}(\tau \rightarrow \ell V^0)$ upper limits in this paper, that of Belle II prospects, and some NP predictions [36, 44, 46]. The observed upper limits in this paper are under the predictions of the single vector leptoquark model ($\mu\phi$) and the unparticle physics model ($\mu\rho^0$, $\mu\phi$ and $\mu\omega$). The expected upper limits in Belle II experiment reach the predictions of the Littlest Higgs model with T-parity.

Further background reduction in Belle II experiment is expected. New charged K/π identification detectors can suppress the pion fake rate for the $\ell\phi$, ℓK^{*0} and $\ell\bar{K}^{*0}$ modes. The track momentum resolution is also expected to be improved, it results in the improvements of τ mass and ΔE resolution. In addition, the BDT performance will improve as the background control samples for BDT training increase. Especially for the $\ell\phi$ modes, the number of background events in the Belle data set is not enough for the BDT training and for the evaluation of BDT performance. I expect much background reduction for the $\ell\phi$ modes and their sensitivity can be reached the branching fractions of $\mathcal{O}(10^{-10})$.

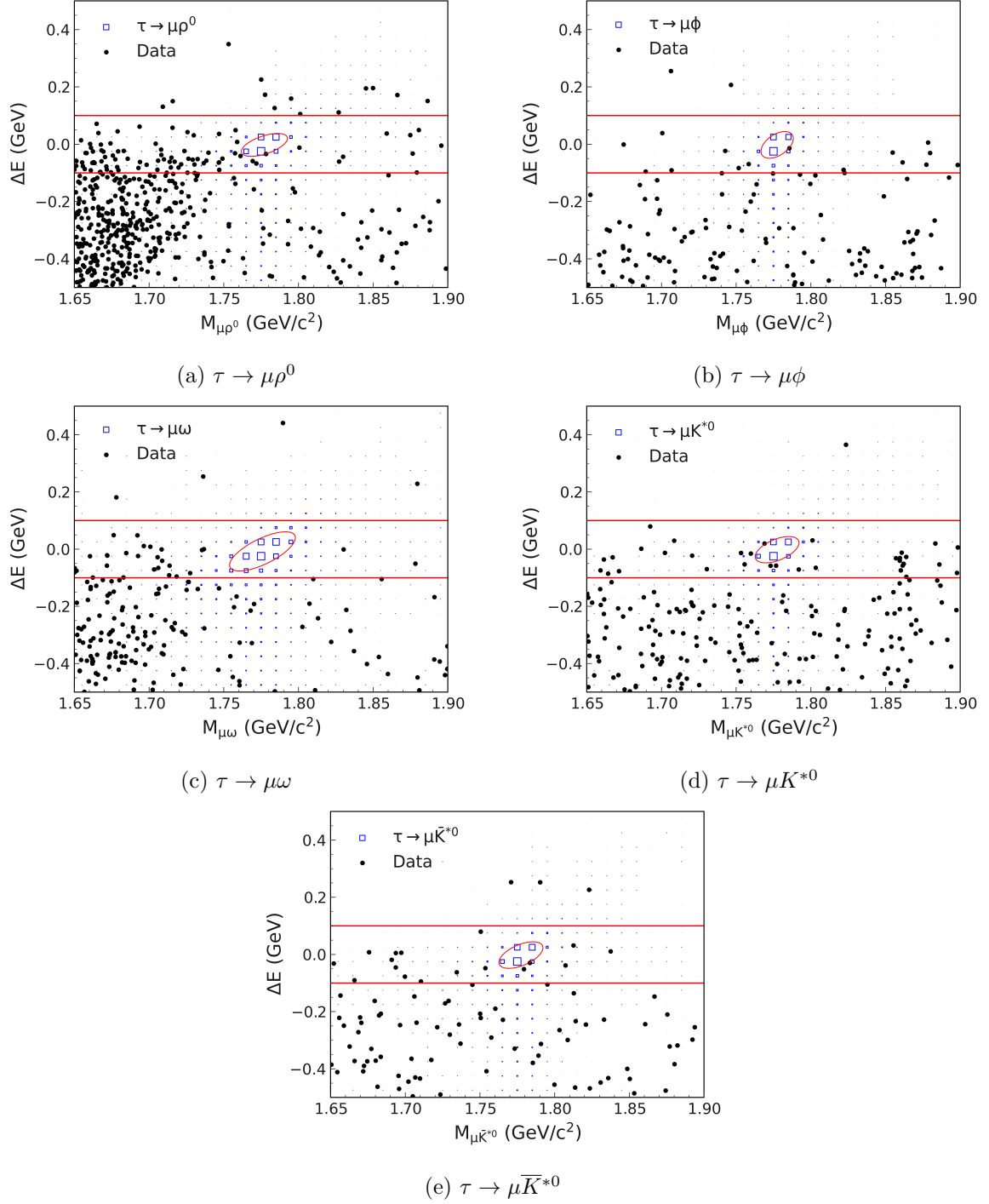


Figure 6.1: Observed event distributions in $M_{\ell V^0} - \Delta E$ plane after the $\tau \rightarrow \mu V^0$ cut-base selection. Black points are the data, blue squares show the signal MC distribution with an arbitrary normalization. The red ellipse is the signal region. The estimation of the number of background events is done using the data between the red horizontal lines except the blind region.

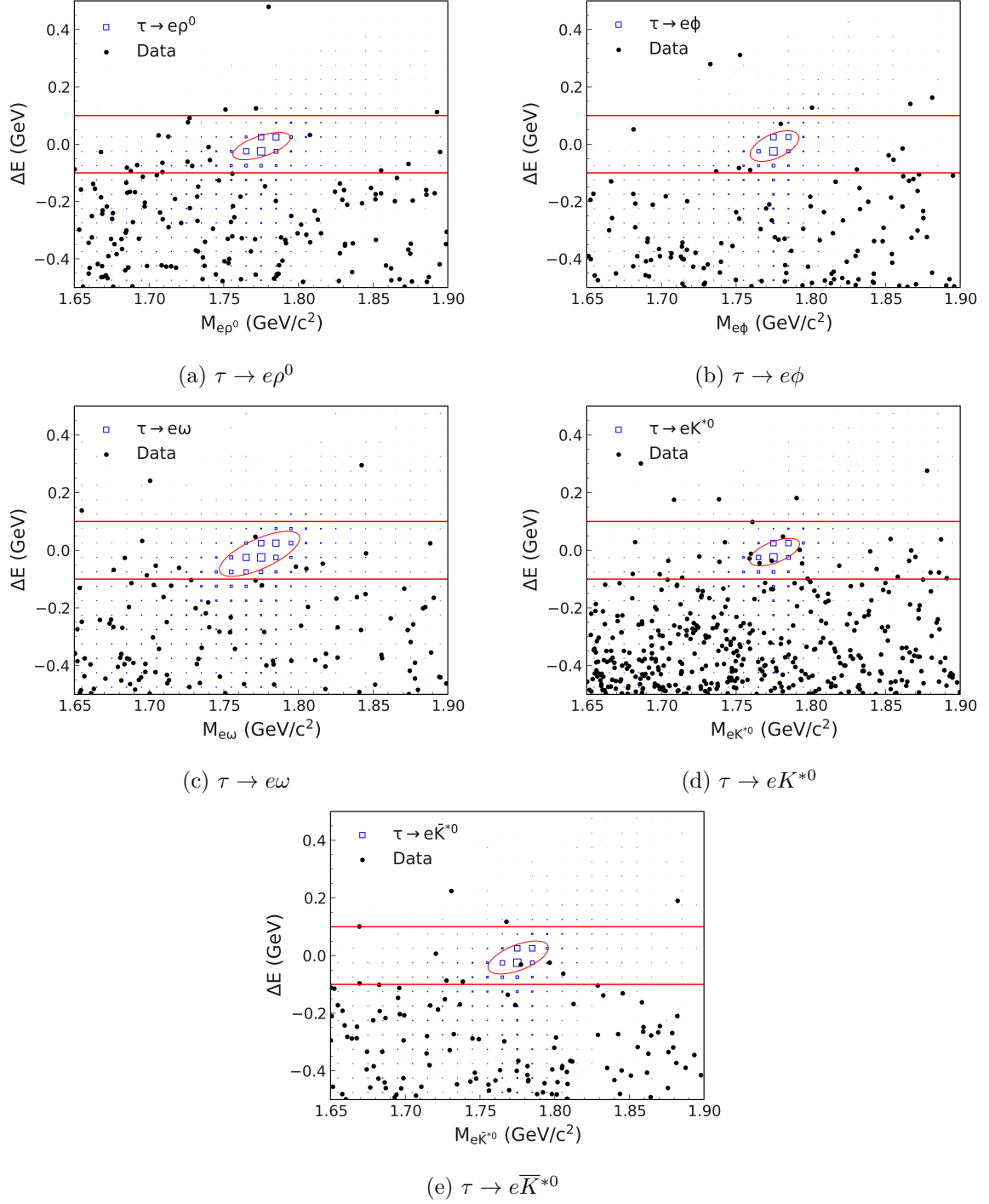


Figure 6.2: Observed event distributions in $M_{\ell V^0}$ – ΔE plane after the $\tau \rightarrow e V^0$ cut-based selection. Black points are the data, blue squares show the signal MC distribution with an arbitrary normalization. The red ellipse is the signal region. The estimation of the number of background events is done using the data between the red horizontal lines except the blind region.

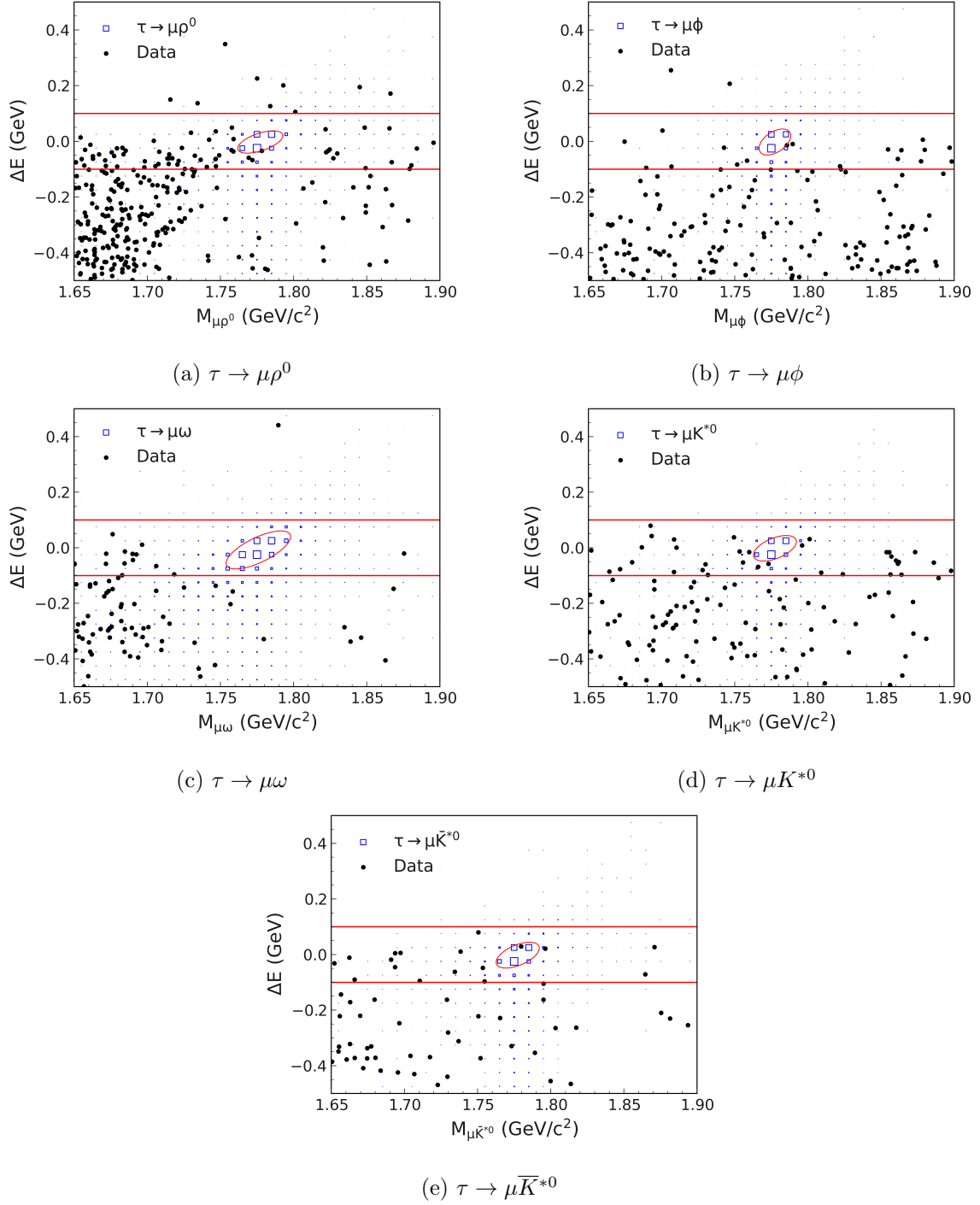
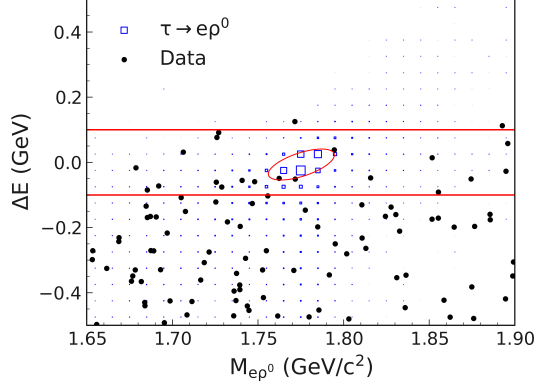
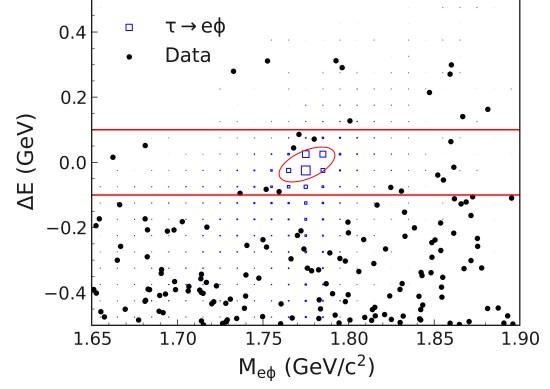


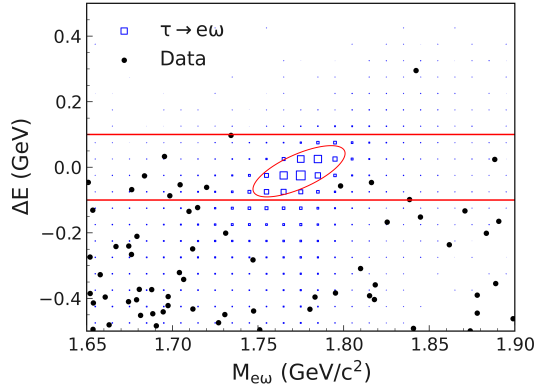
Figure 6.3: Observed event distributions in $M_{\ell V^0} - \Delta E$ plane after the $\tau \rightarrow \mu V^0$ BDT selection. Black points are the data, blue squares show the signal MC distribution with an arbitrary normalization. The red ellipse is the signal region. The estimation of the number of background events is done using the data between the red horizontal lines except the blind region.



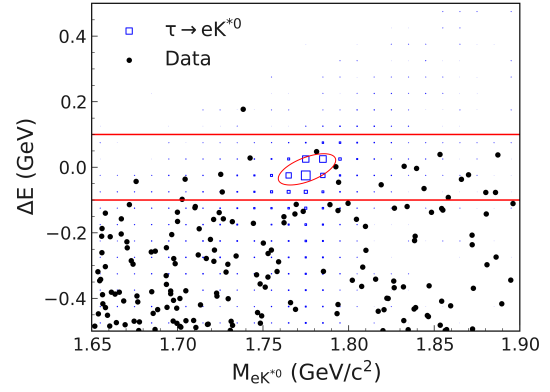
(a) $\tau \rightarrow e\rho^0$



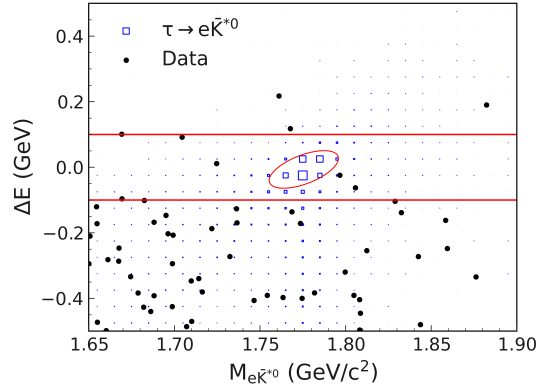
(b) $\tau \rightarrow e\phi$



(c) $\tau \rightarrow e\omega$



(d) $\tau \rightarrow eK^{*0}$



(e) $\tau \rightarrow e\bar{K}^{*0}$

Figure 6.4: Observed event distributions in $M_{\ell V^0}$ - ΔE plane after the $\tau \rightarrow eV^0$ BDT selection. Black points are the data, blue squares show the signal MC distribution with an arbitrary normalization. The red ellipse is the signal region. The estimation of the number of background events is done using the data between the red horizontal lines except the blind region.

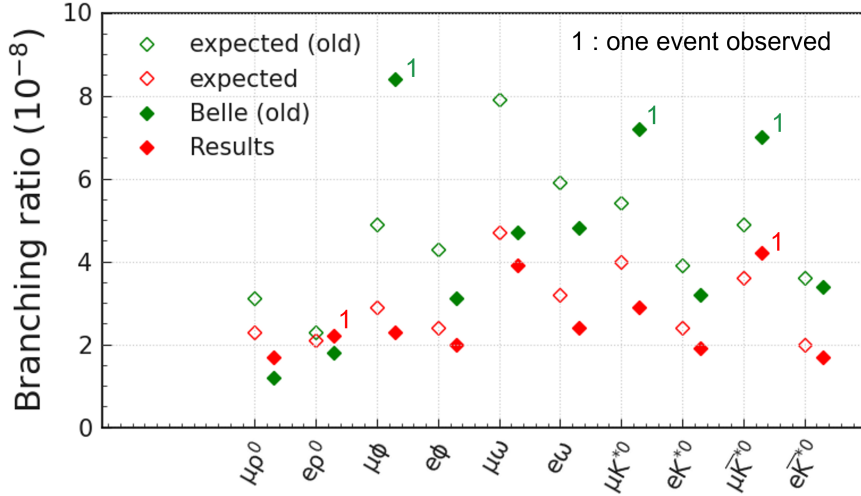


Figure 6.5: The comparison of the $\mathcal{B}(\tau \rightarrow \ell V^0)$ upper limits between the previous paper and the BDT selection. The open diamonds are the expected upper limits, and the filled diamonds are the observed upper limits. The mark "1" on the right shoulder of the filled diamond markers means one event observed in the mode. The other modes without the mark "1" have no event in the signal region.

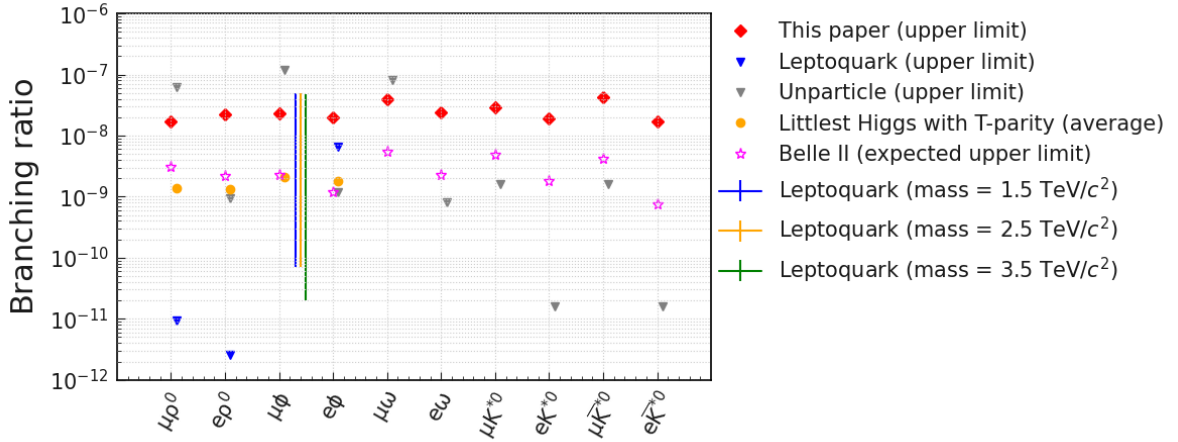


Figure 6.6: The comparison of the $\mathcal{B}(\tau \rightarrow \ell V^0)$ upper limits in this paper and some NP predictions. The red diamonds are the upper limits in this paper. The blue triangles are the upper limits of the single vector leptoquark model [36]. The gray triangles are the upper limits of the unparticle physics model [46]. The orange circles are the mean values of the littlest Higgs model with T-parity [44]. The open violet stars are the conservative Belle II prospects. The blue, orange, and green lines are the 90% C.L. region of the $\mu\phi$ mode in the single vector leptoquark model [36] in the case of that leptoquark mass is 1.5 TeV/ c^2 , 2.5 TeV/ c^2 , and 3.5 TeV/ c^2 , respectively.

Chapter 7

Conclusion

I search for a lepton-flavor violating τ decay into a lepton and a vector meson using the full 980 fb^{-1} of Belle data. I work on the following improvements on the previous results [18]:

- The integrated luminosity is increased by 15%
- Additional reconstructions of $\tau_{\text{tag}}^{\pm} \rightarrow \pi^{\pm}\pi^{+}\pi^{-}\nu$, $\pi^{\pm}\pi^0\pi^0\nu$ increase the signal efficiency by 9.6%
- $q\bar{q}$ background events are suppressed by neutrino mass square (M_{ν}^2) selection
- A multivariate analysis using BDT is introduced to the event selection
- The N_{BG} estimation method is updated utilizing the two-dimensional background probability density function in the $M_{\ell V^0}-\Delta E$ plane.

No statistically significant signal candidates are observed, and the 90% C.L. upper limits on the branching fraction are in the range of $(1.7\text{--}4.2) \times 10^{-8}$ for $\tau \rightarrow \mu V^0$ and $(1.7\text{--}2.4) \times 10^{-8}$ for $\tau \rightarrow e V^0$. We choose the BDT selection for our final result and the upper limits are improved by 30% on average from the previous results. This thesis sets the most stringent upper limits for the branching fractions of $\tau \rightarrow \ell\phi$, $\ell\omega$, ℓK^{*0} , and $\ell\bar{K}^{*0}$. In the future, the sensitivity of these searches can be improved by factor 7–50 with 50 ab^{-1} data set of the Belle II experiment. Then, the single leptoquark model and other NP models may be verified.

Acknowledgments

I would like to express the deepest appreciation to Prof. Toru Iijima and Prof. Kenji Inami for giving me an opportunity to work on the high energy physics experiment, and for guiding me for eight years. Without their encouragement, this paper would not have materialized. I am also grateful to Prof. Kodai Matsuoka for his advice and discussions. My colleagues, the N-lab members and Belle (II) collaborators, also greatly encourage my research activity. I thank all those people sincerely.

This paper is based on data collected using the KEKB accelerator and Belle detector, which was operated until June 2010. Nagoya University and KEK give me financial support through employing me as a research assistant.

Bibliography

- [1] SUPER-KAMIOKANDE collaboration, *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562 [[hep-ex/9807003](#)].
- [2] SNO collaboration, *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **89** (2002) 011301 [[nucl-ex/0204008](#)].
- [3] P.F. de Salas, D.V. Forero, S. Gariazzo, P. Martínez-Miravé, O. Mena, C.A. Ternes et al., *2020 global reassessment of the neutrino oscillation picture*, *JHEP* **02** (2021) 071 [[2006.11237](#)].
- [4] MUON G-2 collaboration, *Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL*, *Phys. Rev. D* **73** (2006) 072003 [[hep-ex/0602035](#)].
- [5] MUON G-2 collaboration, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, *Phys. Rev. Lett.* **126** (2021) 141801 [[2104.03281](#)].
- [6] BABAR collaboration, *Evidence for an excess of $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ decays*, *Phys. Rev. Lett.* **109** (2012) 101802 [[1205.5442](#)].
- [7] BABAR collaboration, *Measurement of an Excess of $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, *Phys. Rev. D* **88** (2013) 072012 [[1303.0571](#)].
- [8] BELLE collaboration, *Measurement of the branching ratio of $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, *Phys. Rev. D* **92** (2015) 072014 [[1507.03233](#)].
- [9] BELLE collaboration, *Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \rightarrow D^*\tau^-\bar{\nu}_\tau$* , *Phys. Rev. Lett.* **118** (2017) 211801 [[1612.00529](#)].
- [10] BELLE collaboration, *Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method*, *Phys. Rev. Lett.* **124** (2020) 161803 [[1910.05864](#)].
- [11] LHCb collaboration, *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)$* , *Phys. Rev. Lett.* **115** (2015) 111803 [[1506.08614](#)].
- [12] LHCb collaboration, *Measurement of the ratio of the $B^0 \rightarrow D^{*-}\tau^+\nu_\tau$ and $B^0 \rightarrow D^{*-}\mu^+\nu_\mu$ branching fractions using three-prong τ -lepton decays*, *Phys. Rev. Lett.* **120** (2018) 171802 [[1708.08856](#)].
- [13] LHCb collaboration, *Test of Lepton Flavor Universality by the measurement of the $B^0 \rightarrow D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays*, *Phys. Rev. D* **97** (2018) 072013 [[1711.02505](#)].
- [14] BELLE collaboration, *Lepton-Flavor-Dependent Angular Analysis of $B \rightarrow K^*\ell^+\ell^-$* , *Phys. Rev. Lett.* **118** (2017) 111801 [[1612.05014](#)].

- [15] LHCb collaboration, *Measurement of CP-Averaged Observables in the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ Decay*, *Phys. Rev. Lett.* **125** (2020) 011802 [[2003.04831](#)].
- [16] LHCb collaboration, *Angular Analysis of the $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ Decay*, *Phys. Rev. Lett.* **126** (2021) 161802 [[2012.13241](#)].
- [17] LHCb collaboration, *Branching Fraction Measurements of the Rare $B_s^0 \rightarrow \phi \mu^+ \mu^-$ and $B_s^0 \rightarrow f_2'(1525) \mu^+ \mu^-$ Decays*, *Phys. Rev. Lett.* **127** (2021) 151801 [[2105.14007](#)].
- [18] BELLE collaboration, *Search for Lepton-Flavor-Violating tau Decays into a Lepton and a Vector Meson*, *Phys. Lett. B* **699** (2011) 251 [[1101.0755](#)].
- [19] B.W. Lee and R.E. Shrock, *Natural Suppression of Symmetry Violation in Gauge Theories: Muon - Lepton and Electron Lepton Number Nonconservation*, *Phys. Rev. D* **16** (1977) 1444.
- [20] J. Ellis, *TikZ-Feynman: Feynman diagrams with TikZ*, *Comput. Phys. Commun.* **210** (2017) 103 [[1601.05437](#)].
- [21] M. Dohse, *TikZ-FeynHand: Basic User Guide*, [1802.00689](#).
- [22] MEG collaboration, *Search for the lepton flavour violating decay $\mu^+ \rightarrow e^+ \gamma$ with the full dataset of the MEG experiment*, *Eur. Phys. J. C* **76** (2016) 434 [[1605.05081](#)].
- [23] SINDRUM collaboration, *Search for the Decay $\mu^+ \rightarrow e^+ e^+ e^-$* , *Nucl. Phys. B* **299** (1988) 1.
- [24] SINDRUM II collaboration, *A Search for muon to electron conversion in muonic gold*, *Eur. Phys. J. C* **47** (2006) 337.
- [25] ATLAS collaboration, *Search for the lepton flavor violating decay $Z \rightarrow e \mu$ in pp collisions at $\sqrt{s}$ TeV with the ATLAS detector*, *Phys. Rev. D* **90** (2014) 072010 [[1408.5774](#)].
- [26] ATLAS collaboration, *Search for lepton-flavor-violation in Z-boson decays with τ -leptons with the ATLAS detector*, *Phys. Rev. Lett.* **127** (2022) 271801 [[2105.12491](#)].
- [27] ATLAS collaboration, *Search for the Higgs boson decays $H \rightarrow ee$ and $H \rightarrow e \mu$ in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Lett. B* **801** (2020) 135148 [[1909.10235](#)].
- [28] CMS collaboration, *Search for lepton-flavor violating decays of the Higgs boson in the $\mu \tau$ and $e \tau$ final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **104** (2021) 032013 [[2105.03007](#)].
- [29] HFLAV collaboration, *Averages of b-hadron, c-hadron, and τ -lepton properties as of 2018*, *Eur. Phys. J. C* **81** (2021) 226 [[1909.12524](#)].
- [30] L. Calibbi, A. Faccia, A. Masiero and S.K. Vempati, *Lepton flavour violation from SUSY-GUTs: Where do we stand for MEG, PRISM/PRIME and a super flavour factory*, *Phys. Rev. D* **74** (2006) 116002 [[hep-ph/0605139](#)].
- [31] J.R. Ellis, J. Hisano, M. Raidal and Y. Shimizu, *A New parametrization of the seesaw mechanism and applications in supersymmetric models*, *Phys. Rev. D* **66** (2002) 115013 [[hep-ph/0206110](#)].
- [32] A. Brignole and A. Rossi, *Lepton flavor violating decays of supersymmetric Higgs bosons*, *Phys. Lett. B* **566** (2003) 217 [[hep-ph/0304081](#)].

- [33] G. Cvetič, C. Dib, C.S. Kim and J.D. Kim, *On lepton flavor violation in tau decays*, *Phys. Rev. D* **66** (2002) 034008 [[hep-ph/0202212](#)].
- [34] A. Celis, V. Cirigliano and E. Passemar, *Model-discriminating power of lepton flavor violating τ decays*, *Phys. Rev. D* **89** (2014) 095014 [[1403.5781](#)].
- [35] BELLE collaboration, *Search for lepton-flavor-violating tau-lepton decays to $\ell\gamma$ at Belle*, *JHEP* **10** (2021) 19 [[2103.12994](#)].
- [36] C. Hati, J. Kriewald, J. Orloff and A.M. Teixeira, *The fate of $\mathbf{V}_1$ vector leptoquarks: the impact of future flavour data*, *Eur. Phys. J. C* **81** (2021) 1066 [[2012.05883](#)].
- [37] L. Di Luzio, J. Fuentes-Martin, A. Greljo, M. Nardecchia and S. Renner, *Maximal Flavour Violation: a Cabibbo mechanism for leptoquarks*, *JHEP* **11** (2018) 081 [[1808.00942](#)].
- [38] J. Kumar, D. London and R. Watanabe, *Combined Explanations of the $b \rightarrow s\mu^+\mu^-$ and $b \rightarrow c\tau^-\bar{\nu}$ Anomalies: a General Model Analysis*, *Phys. Rev. D* **99** (2019) 015007 [[1806.07403](#)].
- [39] A. Crivellin, C. Greub, D. Müller and F. Saturnino, *Importance of Loop Effects in Explaining the Accumulated Evidence for New Physics in B Decays with a Vector Leptoquark*, *Phys. Rev. Lett.* **122** (2019) 011805 [[1807.02068](#)].
- [40] A. Crivellin, D. Müller and F. Saturnino, *Flavor Phenomenology of the Leptoquark Singlet-Triplet Model*, *JHEP* **06** (2020) 020 [[1912.04224](#)].
- [41] P.S. Bhupal Dev, R. Mohanta, S. Patra and S. Sahoo, *Unified explanation of flavor anomalies, radiative neutrino masses, and ANITA anomalous events in a vector leptoquark model*, *Phys. Rev. D* **102** (2020) 095012 [[2004.09464](#)].
- [42] LHCb collaboration, *Test of lepton universality in beauty-quark decays*, *Nature Phys.* **18** (2022) 277 [[2103.11769](#)].
- [43] LHCb collaboration, *Test of lepton universality in $b \rightarrow s\ell^+\ell^-$ decays*, [2212.09152](#).
- [44] I. Pacheco and P. Roig, *Lepton flavour violation in hadron decays of the tau lepton within the littlest Higgs model with T-parity*, *JHEP* **09** (2022) 144 [[2207.04085](#)].
- [45] A. Arhrib, R. Benbrik and C.-H. Chen, *Lepton flavor violating tau decays in type-III seesaw mechanism*, *Phys. Rev. D* **81** (2010) 113003 [[0903.1553](#)].
- [46] Z.-H. Li, Y. Li and H.-X. Xu, *Unparticle-Induced Lepton Flavor Violating Decays $\tau \rightarrow \ell(V^0, P^0)$* , *Phys. Lett. B* **677** (2009) 150 [[0901.3266](#)].
- [47] A. Ilakovac, *Lepton flavor violation in the standard model extended by heavy singlet Dirac neutrinos*, *Phys. Rev. D* **62** (2000) 036010 [[hep-ph/9910213](#)].
- [48] S. Kurokawa and E. Kikutani, *Overview of the KEKB accelerators*, *Nucl. Instrum. Meth. A* **499** (2003) 1.
- [49] BELLE collaboration, *Physics Achievements from the Belle Experiment*, *PTEP* **2012** (2012) 04D001 [[1212.5342](#)].
- [50] T. Abe et al., *Achievements of KEKB*, *PTEP* **2013** (2013) 03A001.
- [51] BELLE collaboration, *The Belle Detector*, *Nucl. Instrum. Meth. A* **479** (2002) 117.
- [52] G. Alimonti, H. Aihara, J. Alexander, Y. Asano, A. Bakich, A. Božek et al., *The BELLE silicon vertex detector*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **453** (2000) 71.

- [53] Z. Natkaniec, H. Aihara, Y. Asano, T. Aso, A. Bakich, T. Browder et al., *Status of the Belle silicon vertex detector*, [*Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **560** \(2006\) 1.](#)
- [54] M. Akatsu, K. Fujimoto, Y. Fujita, H. Hirano, M. Hirose, K. Inami et al., *Cathode image readout in the BELLE central drift chamber*, [*Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **454** \(2000\) 322.](#)
- [55] E. Nakano, *Belle PID*, [*Nucl. Instrum. Meth. A* **494** \(2002\) 402.](#)
- [56] A.J. Bevan, B. Golob, T. Mannel, S. Prell, B.D. Yabsley, H. Aihara et al., *The Physics of the B Factories*, [*Eur. Phys. J. C* **74** \(2014\) .](#)
- [57] T. Iijima, M. Amami, I. Adachi, R. Enomoto, R. Itoh, S. Kobayashi et al., *Study on fine-mesh PMTs for detection of aerogel Cherenkov light*, [*Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **387** \(1997\) 64.](#)
- [58] T. Iijima, I. Adachi, R. Enomoto, R. Suda, T. Sumiyoshi, C. Leonidopoulos et al., *Aerogel Cherenkov counter for the BELLE detector*, [*Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **453** \(2000\) 321.](#)
- [59] H. Kichimi, H. Sagawa, Y. Yoshimura, T. Kishida, N. Tamura, H. Suzuki et al., *Timing characteristics of a micro-channel plate and fine mesh photomultiplier tubes in a 1 T field*, [*Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **325** \(1993\) 451.](#)
- [60] BELLE collaboration, *A detailed Monte Carlo simulation for the Belle TOF system*, [*Nucl. Instrum. Meth. A* **491** \(2002\) 54 \[\[hep-ex/0204030\]\(#\)\].](#)
- [61] K. Miyabayashi, *Belle electromagnetic calorimeter*, [*Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **494** \(2002\) 298.](#)
- [62] H. Ikeda, A. Satpathy, B. Ahn, V. Aulchenko, A. Bondar, B. Cheon et al., *A detailed test of the CsI(Tl) calorimeter for BELLE with photon beams of energy between 20 MeV and 5.4 GeV*, [*Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **441** \(2000\) 401.](#)
- [63] A. Abashian, K. Abe, K. Abe, S. Azuchi, P. Behera, S. Chidzik et al., *The K_L/μ detector subsystem for the BELLE experiment at the KEK B-factory*, [*Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **449** \(2000\) 112.](#)
- [64] Y. Ushiroda, A. Mohapatra, H. Sakamoto, Y. Sakai, M. Nakao, Q. An et al., *Development of the central trigger system for the BELLE detector at the KEK B-factory*, [*Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **438** \(1999\) 460.](#)
- [65] E. Nygard, P. Aspell, P. Jarron, P. Weilhammer and K. Yoshioka, *CMOS low noise amplifier for microstrip readout: Design and results*, [*Nucl. Instrum. Meth. A* **301** \(1991\) 506.](#)
- [66] O. Toker, S. Masciocchi, E. Nygard, A. Rudge and P. Weilhammer, *VIKING: A CMOS low noise monolithic 128-channel front end for Si strip detector readout*, [*Nucl. Instrum. Meth. A* **340** \(1994\) 572.](#)

- [67] BELLE collaboration, *Data acquisition system for the BELLE experiment*, *IEEE Trans. Nucl. Sci.* **47** (2000) 56.
- [68] S.Y. Suzuki, R. Itoh, H.W. Kim, H.J. Kim, H.O. Kim, M. Nakao et al., *Belle DAQ system upgrade at 2001*, *Nucl. Instrum. Meth. A* **494** (2002) 535.
- [69] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima and T. Tsukamoto, *Electron identification in Belle*, *Nucl. Instrum. Meth. A* **485** (2002) 490 [[hep-ex/0108044](#)].
- [70] A. Abashian et al., *Muon identification in the Belle experiment at KEKB*, *Nucl. Instrum. Meth. A* **491** (2002) 69.
- [71] S. Jadach, B.F.L. Ward and Z. Was, *The Precision Monte Carlo event generator K K for two fermion final states in e^+e^- collisions*, *Comput. Phys. Commun.* **130** (2000) 260 [[hep-ph/9912214](#)].
- [72] D.J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152.
- [73] S. Jadach, E. Richter-Was, B.F.L. Ward and Z. Was, *Monte Carlo program BHLUMI-2.01 for Bhabha scattering at low angles with Yennie-Frautschi-Suura exponentiation*, *Comput. Phys. Commun.* **70** (1992) 305.
- [74] F.A. Berends, P.H. Daverveldt and R. Kleiss, *Monte Carlo Simulation of Two Photon Processes. 2. Complete Lowest Order Calculations for Four Lepton Production Processes in electron Positron Collisions*, *Comput. Phys. Commun.* **40** (1986) 285.
- [75] R. Brun, F. Bruyant, M. Maire, A.C. McPherson and P. Zancarini, *GEANT3*, .
- [76] S. Brandt, C. Peyrou, R. Sosnowski and A. Wroblewski, *The principal axis of jets — an attempt to analyse high-energy collisions as two-body processes*, *Physics Letters* **12** (1964) 57.
- [77] E. Farhi, *A QCD Test for Jets*, *Phys. Rev. Lett.* **39** (1977) 1587.
- [78] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma et al., *LightGBM: A Highly Efficient Gradient Boosting Decision Tree*, in *Advances in Neural Information Processing Systems*, I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan et al., eds., vol. 30, Curran Associates, Inc., 2017, <https://proceedings.neurips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf>.
- [79] G. Punzi, *Sensitivity of searches for new signals and its optimization*, *eConf C030908* (2003) MODT002 [[physics/0308063](#)].
- [80] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, *Comput. Phys. Commun.* **10** (1975) 343.
- [81] S. Jadach, *Studies of μ -Pair and π -Pair Production at the Electron-Positron Low Energy Colliders*, *Acta Phys. Pol. B* **36** (2005) 2387 [[hep-ph/0506180](#)].
- [82] S. Jadach, B.F.L. Ward and Z. Was, *Coherent exclusive exponentiation for precision Monte Carlo calculations*, *Phys. Rev. D* **63** (2001) 113009 [[hep-ph/0006359](#)].
- [83] A. Caldwell, D. Kollar and K. Kroninger, *BAT: The Bayesian Analysis Toolkit*, *Comput. Phys. Commun.* **180** (2009) 2197 [[0808.2552](#)].

- [84] S. Banerjee, B. Pietrzyk, J.M. Roney and Z. Was, *Tau and muon pair production cross-sections in electron-positron annihilations at $\sqrt{s} = 10.58$ GeV*, [*Phys. Rev. D* **77** \(2008\) 054012](#) [[0706.3235](#)].
- [85] G.J. Feldman and R.D. Cousins, *Unified approach to the classical statistical analysis of small signals*, [*Phys. Rev. D* **57** \(1998\) 3873](#).

Appendix A

Validation of N_{BG} estimation method using background MC

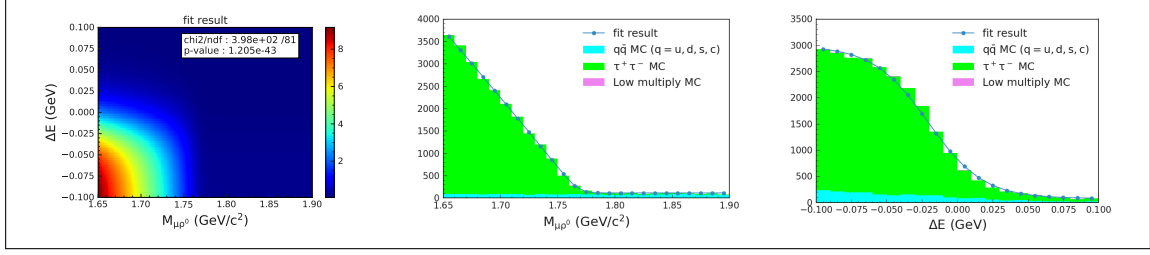
I check our N_{BG} estimation method using the background MC events. Figures A.1 and A.2 are the fit results (see section 5.2.2 for the fit functions) to the control samples in the MC. I use those parameters to define the shape of functions except for the scale factors (c_1 , c_0 and c_0^{lm}) in the function. Figures A.3 and A.4 are the fit results to the MC observed sideband regions with cut-base selection. Figures A.5 and A.6 are that of BDT selection. The blind regions are removed from the fitting as well as the fitting to the data in the sideband. The expected and the observed N_{BG} in MC samples are shown in tables A.1 (cut-base) and A.2 (BDT). The number of events in the validation region, where is inside the blind region but outside the signal region, is also shown. They are scaled to 980 fb^{-1} . The N_{BG} in the signal region is consistent with the observed number of background events in the MC within 2σ . The N_{BG} in the validation region is also consistent with the observed one for most of the ℓV^0 modes, except for the eK^{*0} mode.

Table A.1: Comparison of expected and observed N_{BG} of the MC samples (cut-base selection). They are scaled to 980 fb^{-1} .

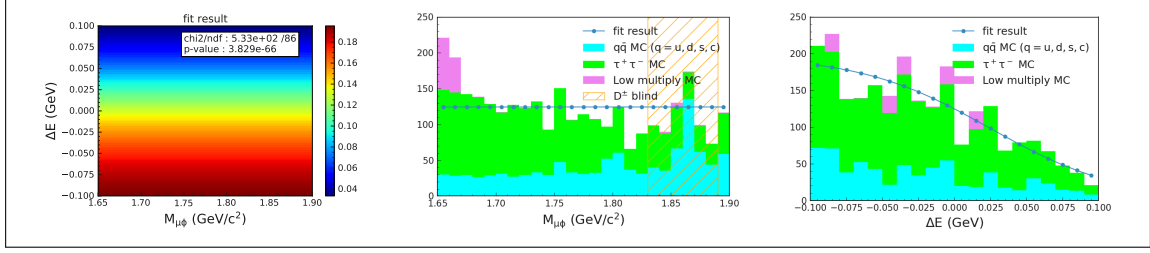
Mode	signal region		validation region	
	expected N_{BG}^{MC}	N_{obs}^{MC}	expected N_{BG}^{MC}	N_{obs}^{MC}
$\tau^- \rightarrow \mu^- \rho^0$	$0.35 \pm 0.05(\text{stat}) \pm 0.02(\text{syst})$	0.55 ± 0.25	$2.5 \pm 0.3(\text{stat}) \pm 0.2(\text{syst})$	2.4 ± 0.5
$\tau^- \rightarrow e^- \rho^0$	$0.31 \pm 0.07(\text{stat}) \pm 0.00(\text{syst})$	0.38 ± 0.22	$2.1 \pm 0.5(\text{stat}) \pm 0.0(\text{syst})$	0.9 ± 0.3
$\tau^- \rightarrow \mu^- \phi$	$0.19 \pm 0.04(\text{stat}) \pm 0.00(\text{syst})$	$0 (<0.15)$	$1.1 \pm 0.2(\text{stat}) \pm 0.0(\text{syst})$	0.8 ± 0.5
$\tau^- \rightarrow e^- \phi$	$0.06 \pm 0.03(\text{stat}) \pm 0.00(\text{syst})$	0.08 ± 0.08	$0.3 \pm 0.2(\text{stat}) \pm 0.0(\text{syst})$	0.3 ± 0.2
$\tau^- \rightarrow \mu^- \omega$	$0.68 \pm 0.13(\text{stat}) \pm 0.03(\text{syst})$	0.45 ± 0.23	$2.7 \pm 0.5(\text{stat}) \pm 0.1(\text{syst})$	1.6 ± 0.5
$\tau^- \rightarrow e^- \omega$	$1.00 \pm 0.16(\text{stat}) \pm 0.00(\text{syst})$	0.68 ± 0.35	$2.2 \pm 0.4(\text{stat}) \pm 0.0(\text{syst})$	2.4 ± 0.6
$\tau^- \rightarrow \mu^- K^{*0}$	$0.72 \pm 0.10(\text{stat}) \pm 0.01(\text{syst})$	0.47 ± 0.33	$2.7 \pm 0.4(\text{stat}) \pm 0.0(\text{syst})$	2.7 ± 0.7
$\tau^- \rightarrow e^- K^{*0}$	$0.66 \pm 0.12(\text{stat}) \pm 0.01(\text{syst})$	0.24 ± 0.24	$4.1 \pm 0.7(\text{stat}) \pm 0.0(\text{syst})$	2.1 ± 0.6
$\tau^- \rightarrow \mu^- \bar{K}^{*0}$	$0.70 \pm 0.07(\text{stat}) \pm 0.01(\text{syst})$	0.51 ± 0.30	$2.6 \pm 0.3(\text{stat}) \pm 0.0(\text{syst})$	3.2 ± 0.8
$\tau^- \rightarrow e^- \bar{K}^{*0}$	$0.83 \pm 0.13(\text{stat}) \pm 0.02(\text{syst})$	0.54 ± 0.28	$3.3 \pm 0.5(\text{stat}) \pm 0.0(\text{syst})$	2.2 ± 0.6

Table A.2: Comparison of expected and observed N_{BG} of the MC samples (BDT selection). They are scaled to 980 fb^{-1} .

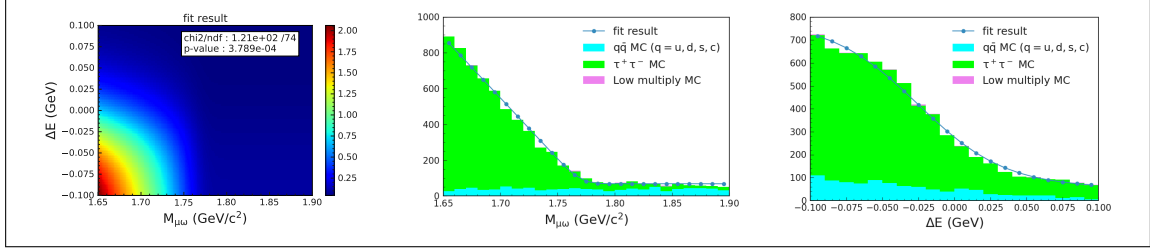
Mode	signal region		validation region	
	expected N_{BG}^{MC}	N_{obs}^{MC}	expected N_{BG}^{MC}	N_{obs}^{MC}
$\tau^- \rightarrow \mu^- \rho^0$	$0.31 \pm 0.05(\text{stat}) \pm 0.01(\text{syst})$	0.54 ± 0.24	$2.1 \pm 0.3(\text{stat}) \pm 0.1(\text{syst})$	1.6 ± 0.4
$\tau^- \rightarrow e^- \rho^0$	$0.28 \pm 0.07(\text{stat}) \pm 0.01(\text{syst})$	0.11 ± 0.11	$1.8 \pm 0.4(\text{stat}) \pm 0.0(\text{syst})$	1.4 ± 0.4
$\tau^- \rightarrow \mu^- \phi$	$0.31 \pm 0.06(\text{stat}) \pm 0.01(\text{syst})$	0.11 ± 0.11	$1.8 \pm 0.3(\text{stat}) \pm 0.0(\text{syst})$	1.2 ± 0.6
$\tau^- \rightarrow e^- \phi$	$0.08 \pm 0.04(\text{stat}) \pm 0.00(\text{syst})$	0.08 ± 0.08	$0.4 \pm 0.2(\text{stat}) \pm 0.0(\text{syst})$	0.3 ± 0.2
$\tau^- \rightarrow \mu^- \omega$	$0.29 \pm 0.08(\text{stat}) \pm 0.02(\text{syst})$	0.55 ± 0.28	$1.2 \pm 0.3(\text{stat}) \pm 0.0(\text{syst})$	1.2 ± 0.4
$\tau^- \rightarrow e^- \omega$	$0.73 \pm 0.15(\text{stat}) \pm 0.00(\text{syst})$	0.34 ± 0.24	$1.7 \pm 0.3(\text{stat}) \pm 0.0(\text{syst})$	1.5 ± 0.5
$\tau^- \rightarrow \mu^- K^{*0}$	$0.61 \pm 0.09(\text{stat}) \pm 0.03(\text{syst})$	0.91 ± 0.47	$2.2 \pm 0.3(\text{stat}) \pm 0.1(\text{syst})$	2.6 ± 0.7
$\tau^- \rightarrow e^- K^{*0}$	$0.50 \pm 0.10(\text{stat}) \pm 0.01(\text{syst})$	0.71 ± 0.38	$3.2 \pm 0.6(\text{stat}) \pm 0.0(\text{syst})$	1.2 ± 0.4
$\tau^- \rightarrow \mu^- \bar{K}^{*0}$	$0.52 \pm 0.06(\text{stat}) \pm 0.00(\text{syst})$	0.59 ± 0.34	$2.0 \pm 0.2(\text{stat}) \pm 0.0(\text{syst})$	2.1 ± 0.7
$\tau^- \rightarrow e^- \bar{K}^{*0}$	$0.48 \pm 0.08(\text{stat}) \pm 0.01(\text{syst})$	0.39 ± 0.23	$1.8 \pm 0.3(\text{stat}) \pm 0.0(\text{syst})$	1.4 ± 0.5



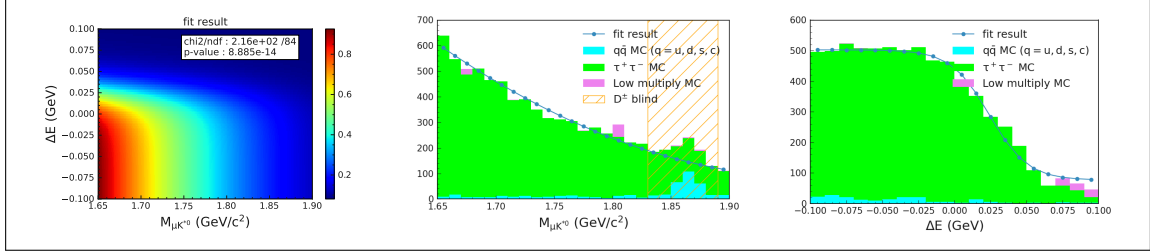
(a) $\tau \rightarrow \mu\rho^0$



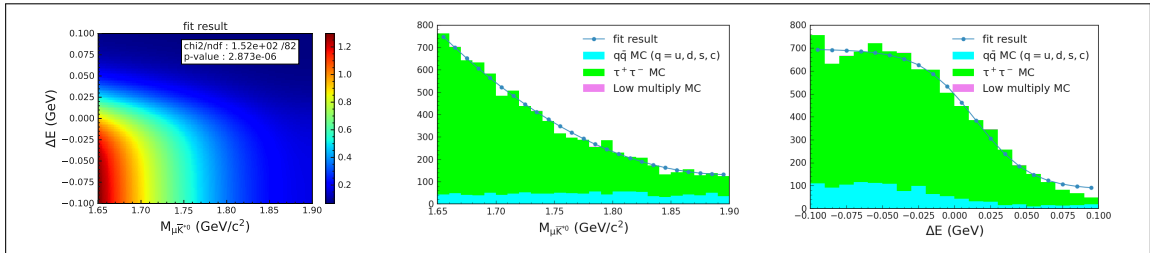
(b) $\tau \rightarrow \mu\phi$



(c) $\tau \rightarrow \mu\omega$

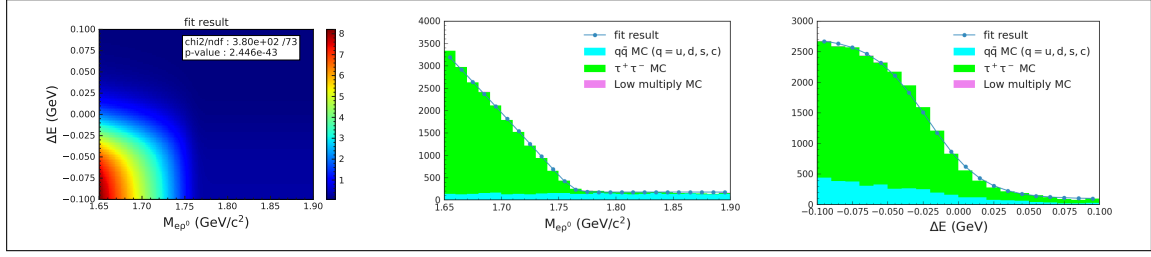


(d) $\tau \rightarrow \mu K^{*0}$

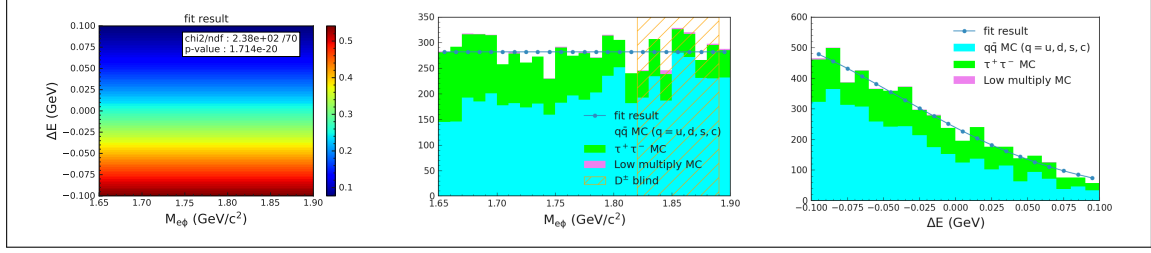


(e) $\tau \rightarrow \mu \bar{K}^{*0}$

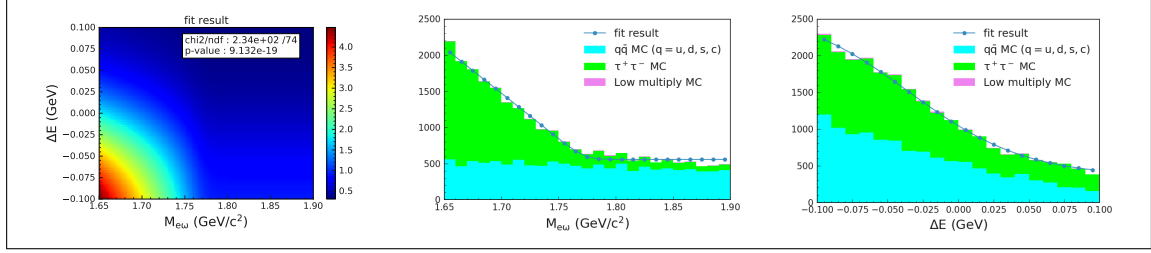
Figure A.1: The fitting results to the μV^0 control samples in the MC (left). The luminosity is scaled to that of $q\bar{q}$ MC (5720 fb^{-1}). Projection to the $M_{\ell V^0}$ (middle) and ΔE (right). The orange hatched regions in $\mu\phi$ and μK^{*0} modes are blinded to avoid $D^\pm$ meson event peaks. They are removed from the fitting and the plot of ΔE projection.



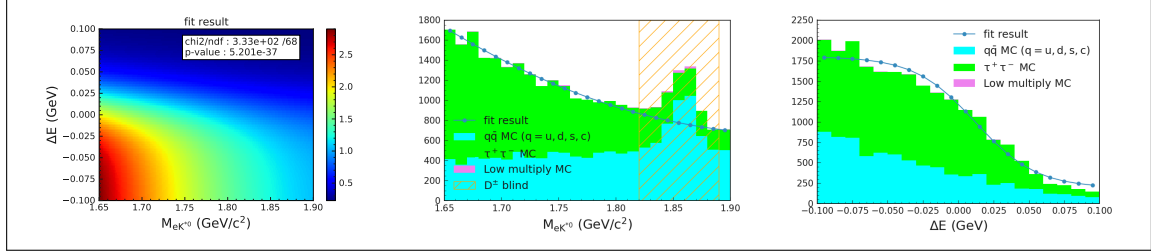
(a) $\tau \rightarrow e\rho^0$



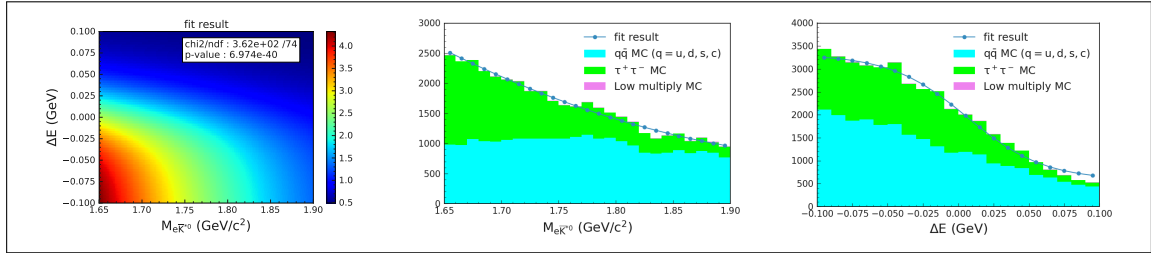
(b) $\tau \rightarrow e\phi$



(c) $\tau \rightarrow e\omega$

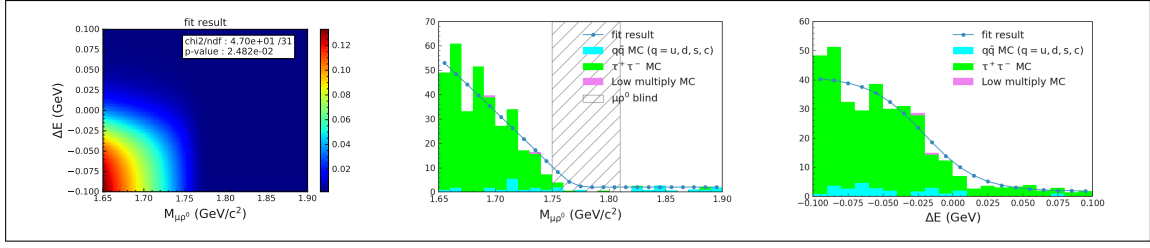


(d) $\tau \rightarrow eK^{*0}$

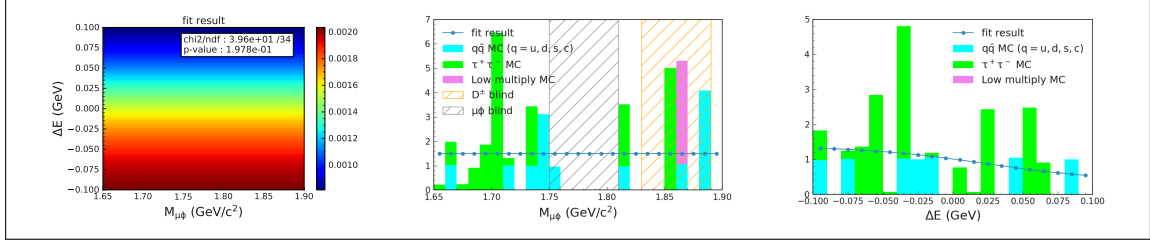


(e) $\tau \rightarrow e\bar{K}^{*0}$

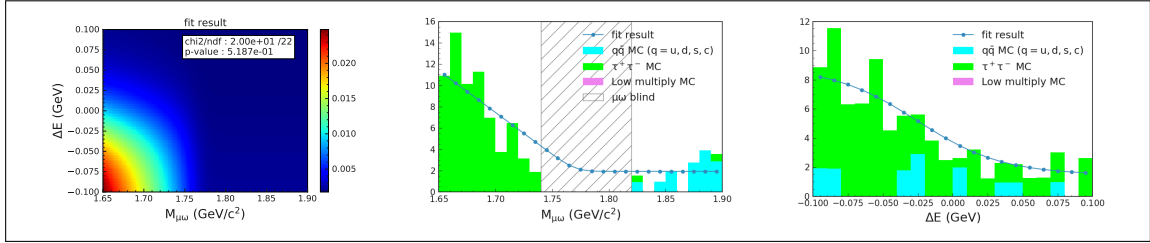
Figure A.2: The fitting results to the eV^0 control samples in the MC (left). The luminosity is scaled to that of $q\bar{q}$ MC (5720 fb^{-1}). Projection to the M_{eV^0} (middle) and ΔE (right). The orange hatched regions in $e\phi$ and eK^{*0} modes are blinded to avoid $D^\pm$ meson event peaks. They are removed from the fitting and the plot of ΔE projection.



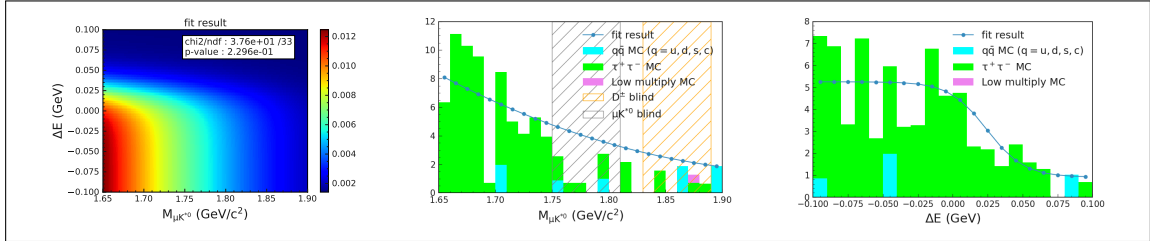
(a) $\tau \rightarrow \mu\rho^0$



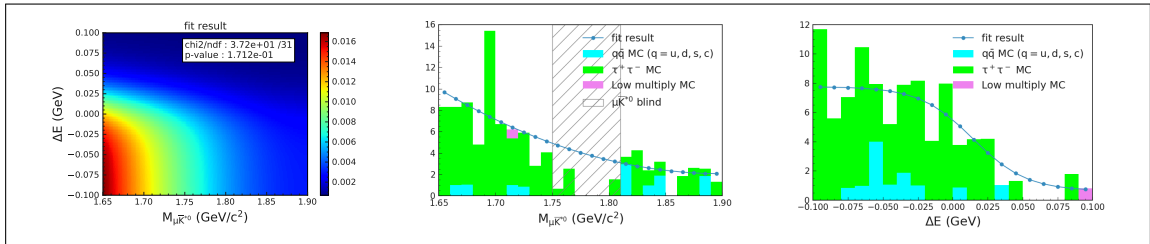
(b) $\tau \rightarrow \mu\phi$



(c) $\tau \rightarrow \mu\omega$

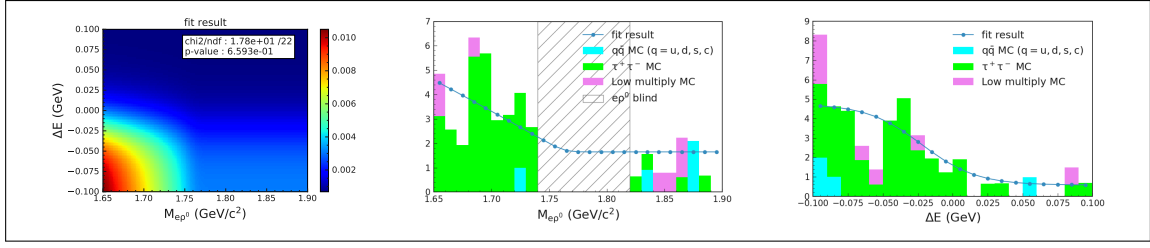


(d) $\tau \rightarrow \mu K^{*0}$

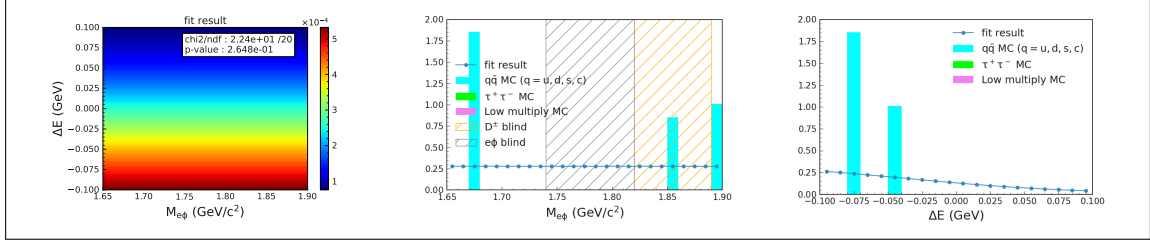


(e) $\tau \rightarrow \mu\bar{K}^{*0}$

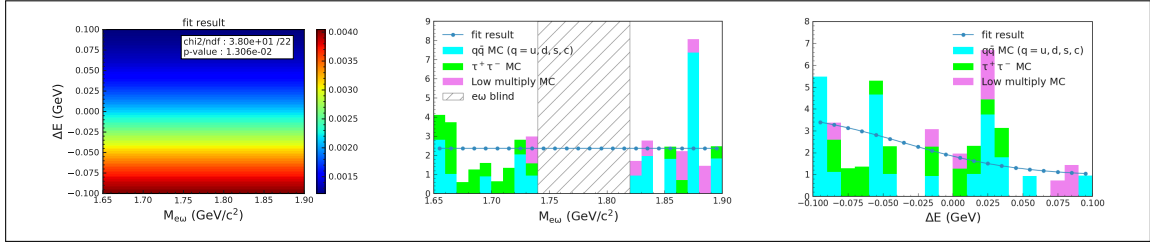
Figure A.3: The fitting results to the μV^0 sideband MC samples after the cut-base selection (left). The luminosity is scaled to that of $q\bar{q}$ MC (5720 fb^{-1}). Projection to the $M_{\ell V^0}$ (middle) and ΔE (right). The gray hatched regions are the blind regions. The orange hatched regions in $\mu\phi$ and μK^{*0} modes are blinded to avoid $D^\pm$ meson event peaks. The hatched regions are removed from the fitting and the plot of ΔE projection. The blind regions are not included for the fitting.



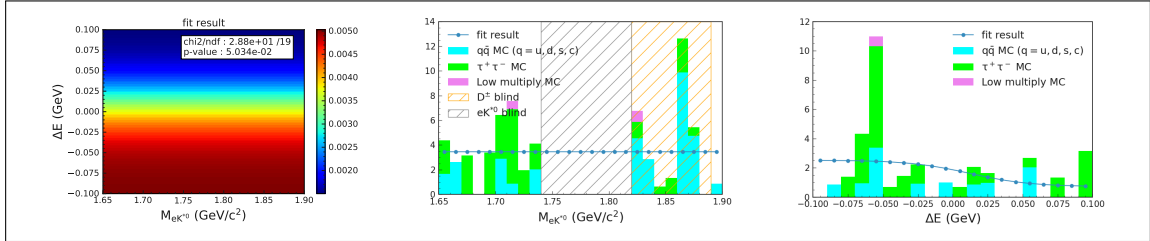
(a) $\tau \rightarrow e\rho^0$



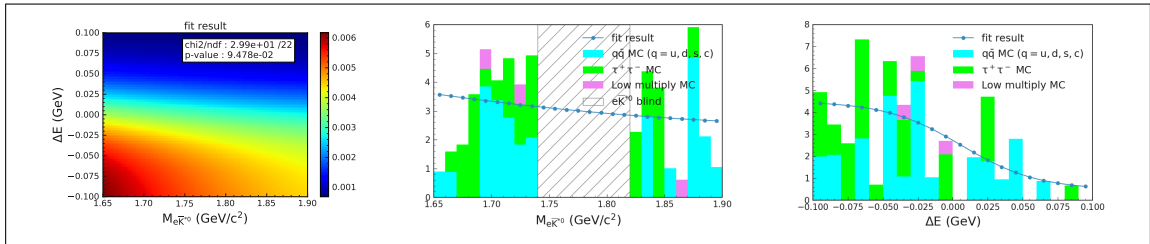
(b) $\tau \rightarrow e\phi$



(c) $\tau \rightarrow e\omega$

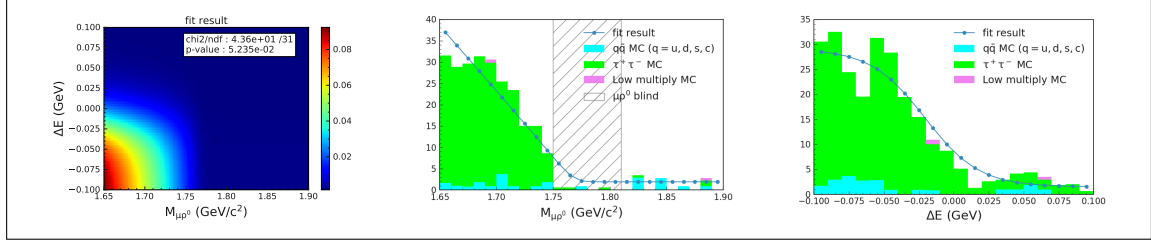


(d) $\tau \rightarrow eK^*0$

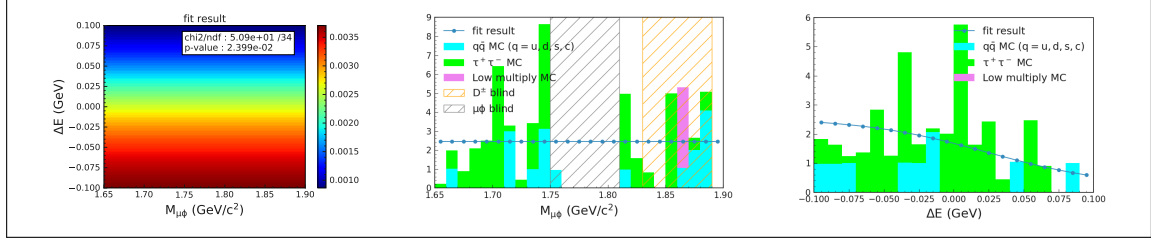


(e) $\tau \rightarrow e\bar{K}^*0$

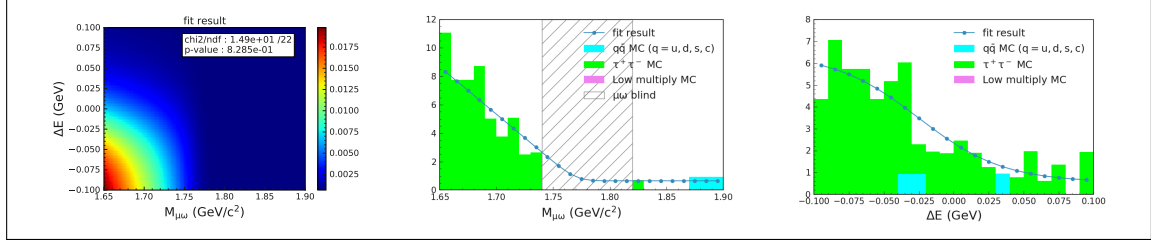
Figure A.4: The fitting results to the eV^0 sideband MC samples after the cut-base selection (left). The luminosity is scaled to that of $q\bar{q}$ MC (5720 fb^{-1}). Projection to the $M_{\ell V^0}$ (middle) and ΔE (right). The gray hatched regions are the blind regions. The orange hatched regions in $e\phi$ and eK^*0 modes are blinded to avoid $D^\pm$ meson event peaks. The hatched regions are removed from the fitting and the plot of ΔE projection. The blind regions are not included for the fitting.



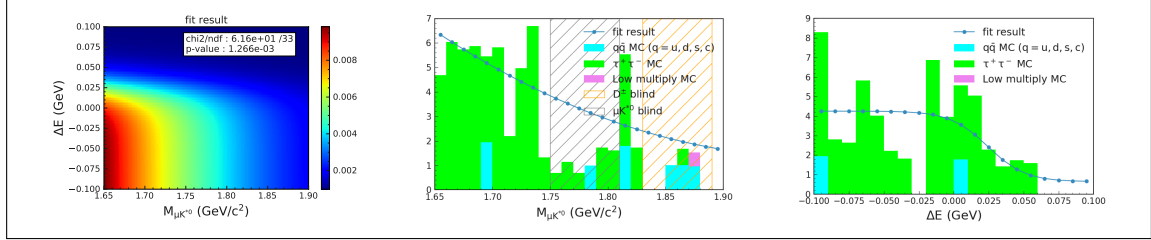
(a) $\tau \rightarrow \mu\rho^0$



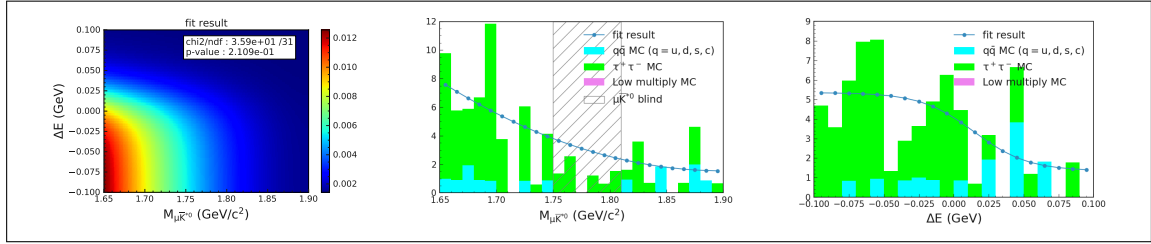
(b) $\tau \rightarrow \mu\phi$



(c) $\tau \rightarrow \mu\omega$

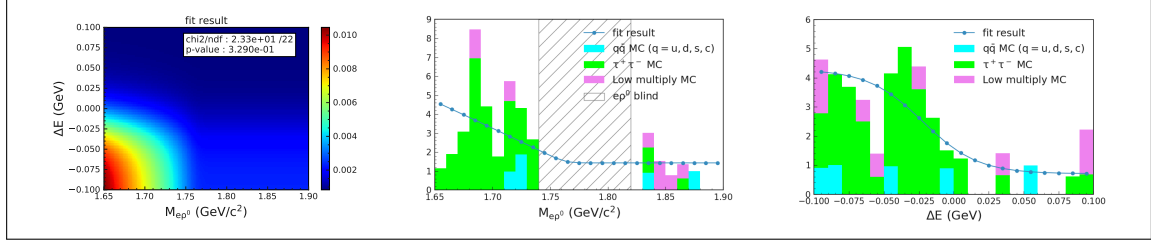


(d) $\tau \rightarrow \mu K^{*0}$

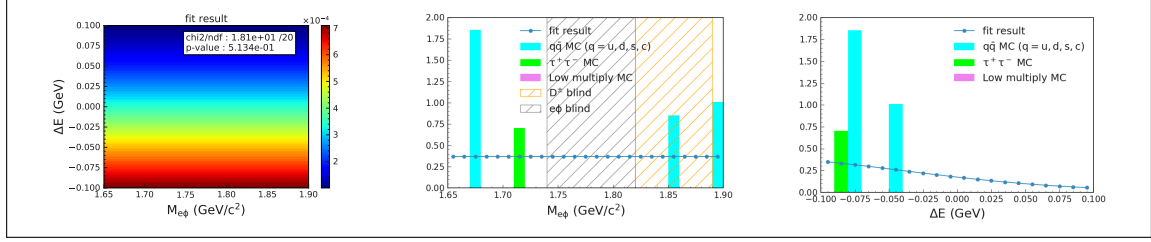


(e) $\tau \rightarrow \mu \bar{K}^{*0}$

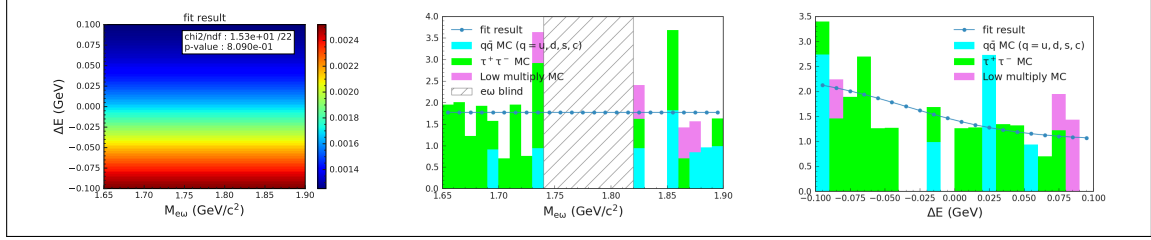
Figure A.5: The fitting results to the μV^0 sideband MC samples after the BDT selection (left). The luminosity is scaled to that of $q\bar{q}$ MC (5720 fb^{-1}). Projection to the $M_{\ell V^0}$ (middle) and ΔE (right). The gray hatched regions are the blind regions. The orange hatched regions in $\mu\phi$ and μK^{*0} modes are blinded to avoid $D^\pm$ meson event peaks. The hatched regions are removed from the fitting and the plot of ΔE projection. The blind regions are not included for the fitting.



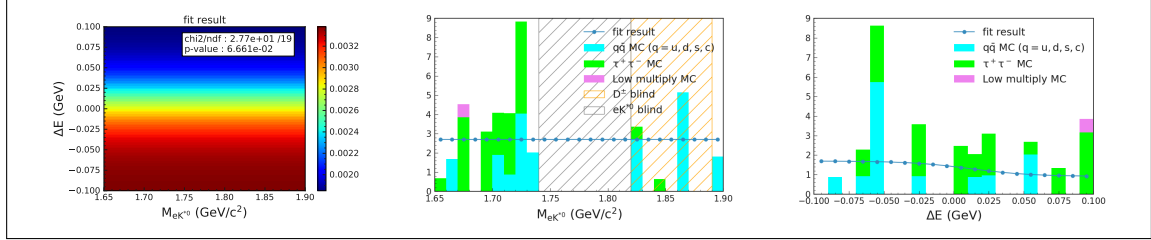
(a) $\tau \rightarrow e\rho^0$



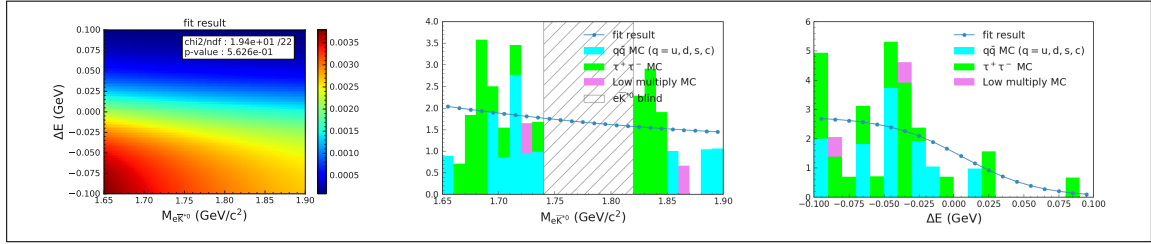
(b) $\tau \rightarrow e\phi$



(c) $\tau \rightarrow e\omega$



(d) $\tau \rightarrow eK^{*0}$



(e) $\tau \rightarrow e\bar{K}^{*0}$

Figure A.6: The fitting results to the eV^0 sideband MC samples after the BDT selection (left). The luminosity is scaled to that of $q\bar{q}$ MC (5720 fb^{-1}). Projection to the $M_{\ell V^0}$ (middle) and ΔE (right). The gray hatched regions are the blind regions. The orange hatched regions in $e\phi$ and eK^{*0} modes are blinded to avoid $D^\pm$ meson event peaks. The hatched regions are removed from the fitting and the plot of ΔE projection. The blind regions are not included for the fitting.

Appendix B

BDT training with MC control samples

Here, I discuss an additional study about the Boosted Decision Tree (BDT) training in section 4.4.2. The BDT performance basically improves as the number of training samples increases. In section 4.4.2, the BDT is trained using the control samples in the data. To increase the number of training samples, I add the control samples (section 4.4.1) in the background MC to the BDT training.

The BDT training is done with 40% of the signal MC events as signal samples, and 80% of the control samples in the data and background MC as background samples. The validation samples, which have 10% of the signal MC events as signal samples and 20% of the control samples in the data and background MC as background samples, are used for the AUC monitoring during the training. If the AUC is not improved for 50 rounds of the training, I stop training and choose the BDT with the best AUC score. The total weights of the control samples in the background MC are scaled to match the number of the control samples in the data. The total weights of the signal MC samples are scaled to match the sum of weights of the control samples in the data and in the background MC. The input variables and the hyper-parameters for the BDT are same as in section 4.4.2.

To compare the BDT performance with that in section 4.4.2, I calculate the AUC of the validation samples except for the background MC (Figures B.1 and B.2). The BDT trained with the data and background MC has better performance than the BDT in section 4.4.2. It means that the BDT can be improved in the future Belle II experiment that has 50 times larger integrated luminosity than the Belle experiment. Because this study is done after the unblind of the $\tau \rightarrow \ell V^0$ signal region, I do not apply this BDT for my final results.

Figures B.3 and B.4 shows the observed event distribution after the BDT selection trained with the control samples in the data and background MC. The $\ell\phi$ modes are not shown because I do not set the BDT output lower limits in the BDT selection for them. Table B.1 shows the number of background estimation and the number of observed events inside the signal region using this BDT at the same signal efficiency in table 6.2. The upper limits of the $\tau \rightarrow \ell V^0$ branching fractions are calculated with the same systematic uncertainty on the expected number of signal events in table 6.2. The observed upper limits are more dependent on the number of observed events than on the expected number of background events. Whereas the expected N_{BG} is decrease about 6.5% by the BDT with additional training samples, the sum of the observed events are same as my final results.

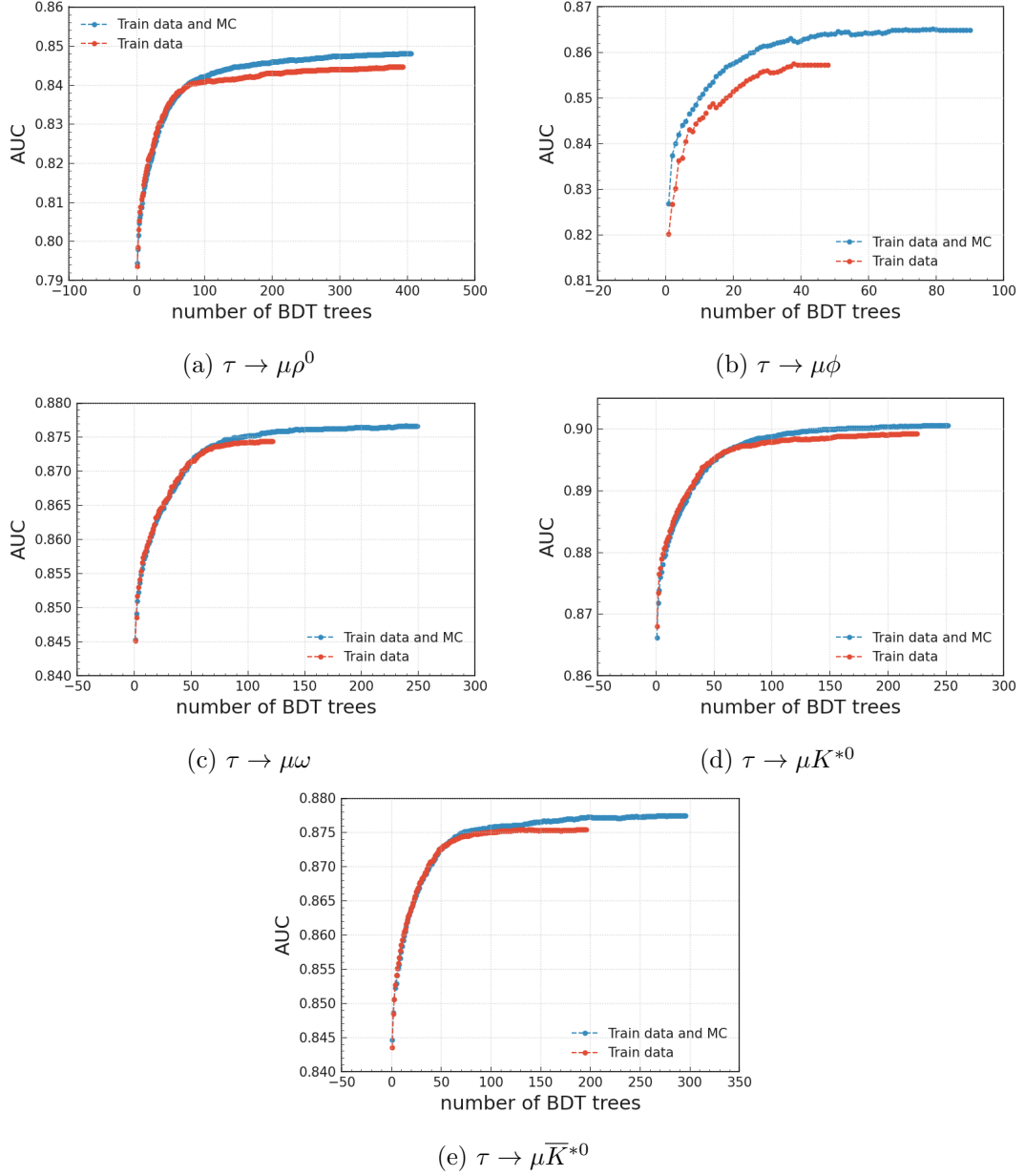


Figure B.1: The BDT performance indicator (AUC of the validation samples in the data) on each training stage (number of BDT tree) for the μV^0 modes. The filled blue circles are the performance of the new BDT that is trained with the control samples both in the data and background MC. The filled red circles are the performance of the BDT that is trained with the control samples only in the data. The additional training samples in the background MC improve the BDT performance. The training is stopped if there is no AUC improvement for 50 training stages.

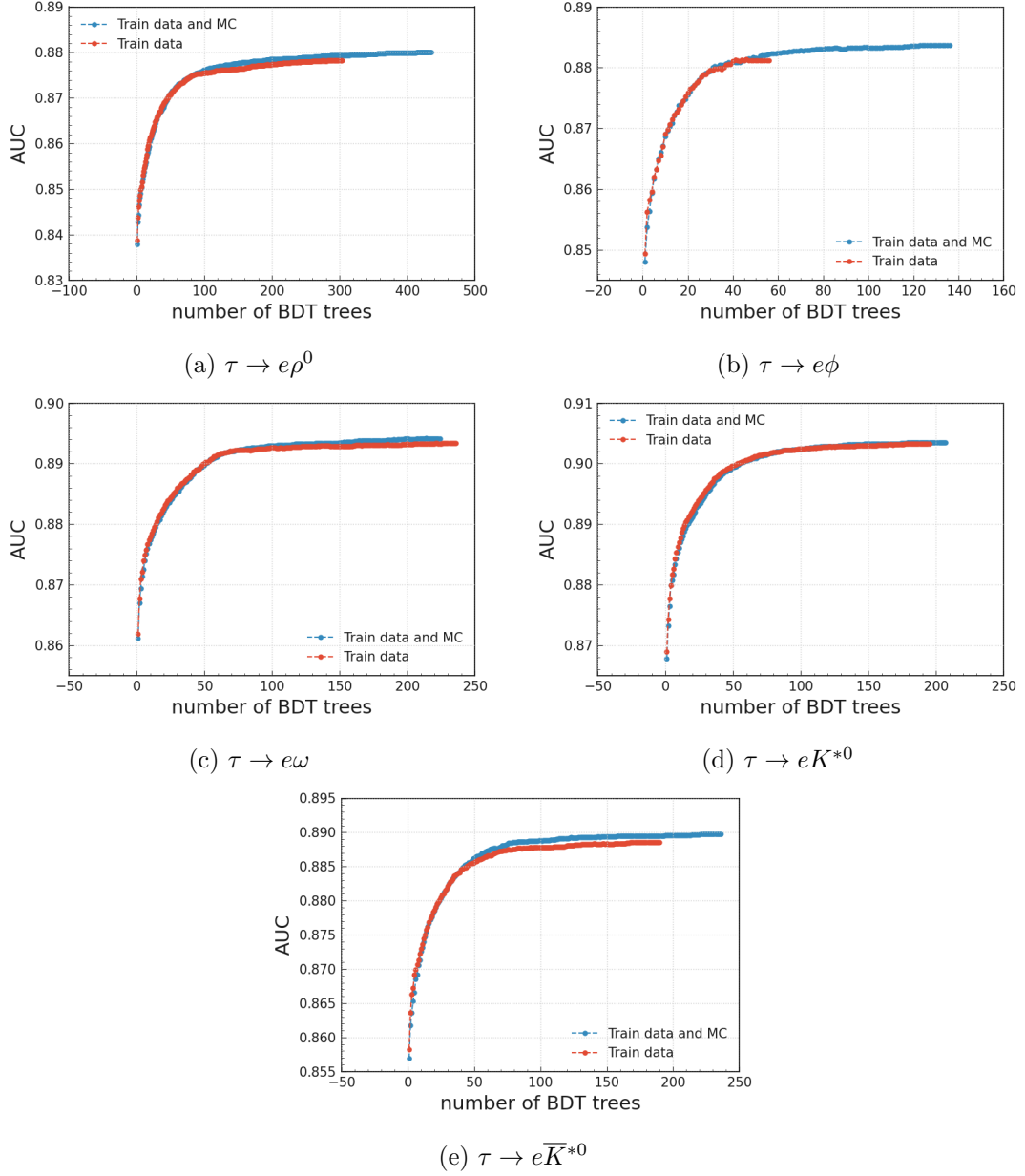


Figure B.2: The BDT performance indicator (AUC of the validation samples in the data) on each training stage (number of BDT tree) for the eV^0 modes. The filled blue circles are the performance of the new BDT that is trained with the control samples both in the data and background MC. The filled red circles are the performance of the BDT that is trained with the control samples only in the data. The additional training samples in the background MC improve the BDT performance. The training is stopped if there is no AUC improvement for 50 training stages.

Table B.1: Comparison of the BDT selection in section 4.4.2 (training with the data) and the BDT in this section (training with the data and background MC). The expected number of background events (N_{BG}), the expected 90% C.L. upper limits ($\mathcal{B}_{\text{exp}}$ (10^{-8})), the number of observed events in the signal region (N_{obs}), and the observed 90% C.L. upper limits ($\mathcal{B}_{\text{obs}}$ (10^{-8})). The signal efficiency is same as in table 6.2. The $\ell\phi$ modes are not here because I do not set the BDT selection for them.

Mode	training with data				training with data+background MC			
	N_{BG}	$\mathcal{B}_{\text{exp}}$	N_{obs}	$\mathcal{B}_{\text{obs}}$	N_{BG}	$\mathcal{B}_{\text{exp}}$	N_{obs}	$\mathcal{B}_{\text{obs}}$
$\tau^- \rightarrow \mu^- \rho^0$	$0.95 \pm 0.20(\text{stat})$	2.3	0	1.7	$0.98 \pm 0.21(\text{stat})$	2.4	0	1.7
$\tau^- \rightarrow e^- \rho^0$	$0.80 \pm 0.27(\text{stat})$	2.1	1	2.2	$0.40 \pm 0.18(\text{stat})$	1.8	1	2.4
$\tau^- \rightarrow \mu^- \omega$	$0.32 \pm 0.23(\text{stat})$	4.7	0	3.9	$0.51 \pm 0.30(\text{stat})$	5.0	0	4.0
$\tau^- \rightarrow e^- \omega$	$0.74 \pm 0.43(\text{stat})$	3.2	0	2.4	$0.74 \pm 0.43(\text{stat})$	3.2	0	2.4
$\tau^- \rightarrow \mu^- K^{*0}$	$0.84 \pm 0.25(\text{stat})$	4.0	0	2.9	$0.81 \pm 0.21(\text{stat})$	3.9	0	2.8
$\tau^- \rightarrow e^- K^{*0}$	$0.54 \pm 0.21(\text{stat})$	2.4	0	1.9	$0.54 \pm 0.20(\text{stat})$	2.4	1	2.8
$\tau^- \rightarrow \mu^- \bar{K}^{*0}$	$0.58 \pm 0.17(\text{stat})$	3.6	1	4.2	$0.49 \pm 0.16(\text{stat})$	3.5	0	2.8
$\tau^- \rightarrow e^- \bar{K}^{*0}$	$0.25 \pm 0.11(\text{stat})$	2.0	0	1.7	$0.17 \pm 0.12(\text{stat})$	1.9	0	1.7

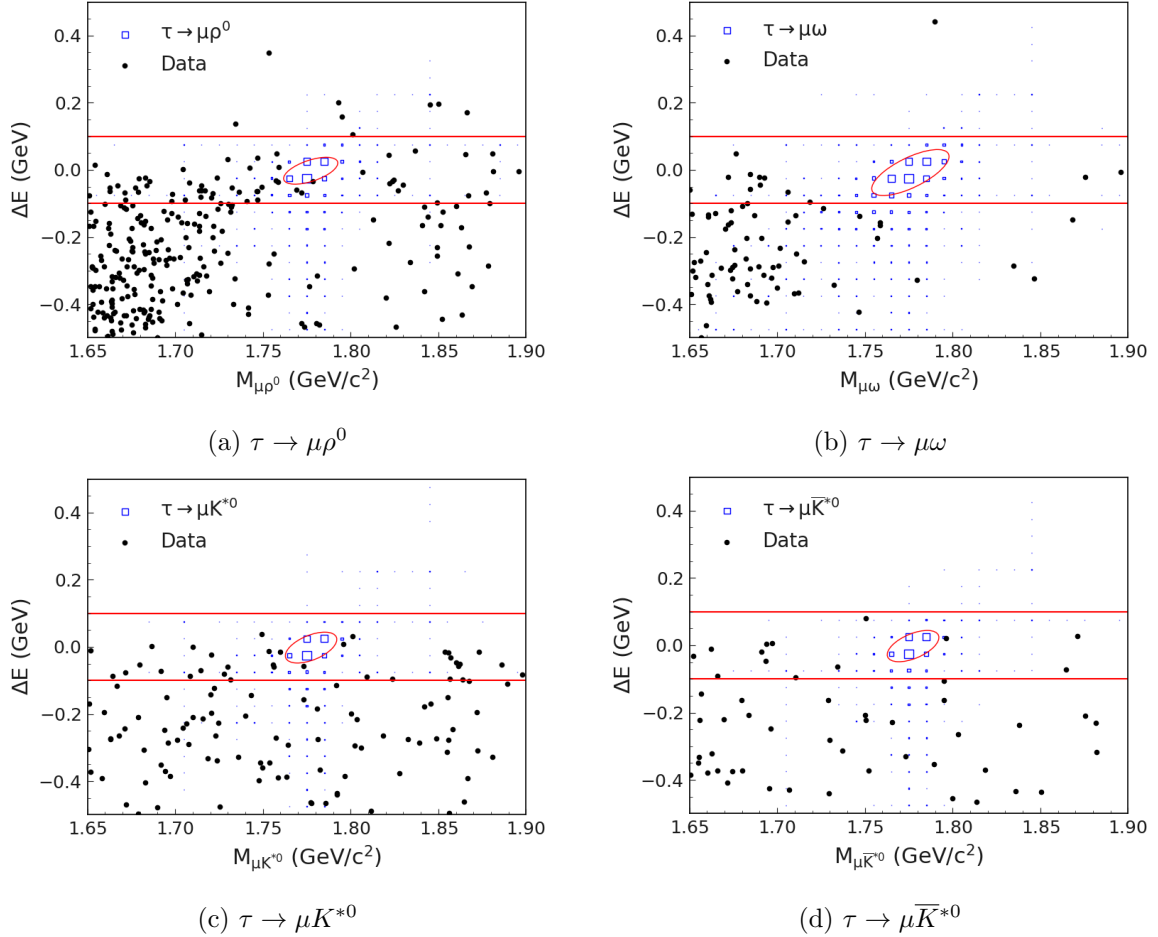
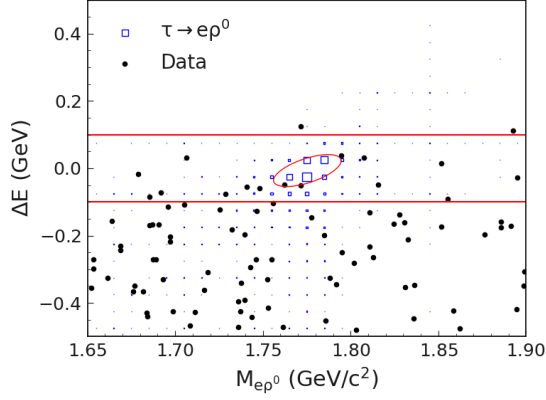
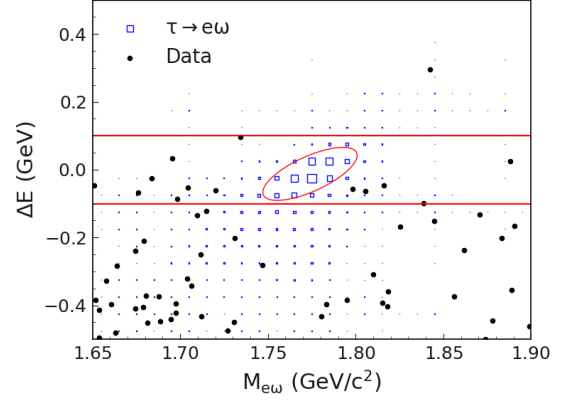


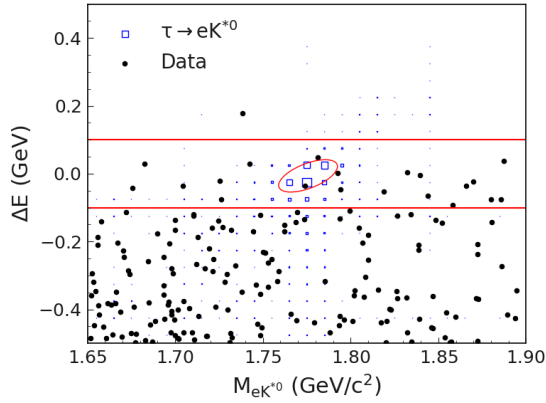
Figure B.3: Observed event distributions in $M_{\ell V^0} - \Delta E$ plane after the BDT selection trained with the control samples in the data and background MC. Black points are the data, blue squares show the signal MC distribution with an arbitrary normalization. The red ellipse is the signal region. The estimation of the number of background events is done using the data between the red horizontal lines except the blind region.



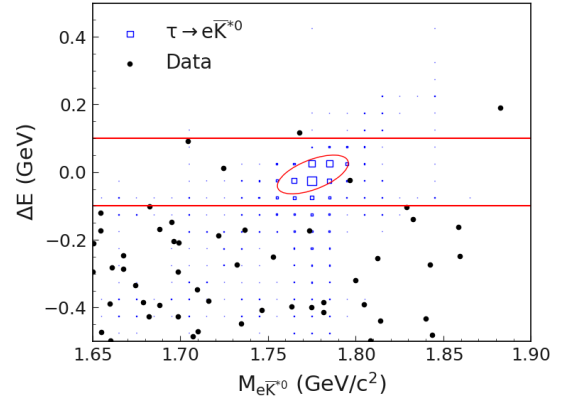
(a) $\tau \rightarrow e\rho^0$



(b) $\tau \rightarrow e\omega$



(c) $\tau \rightarrow eK^{*0}$



(d) $\tau \rightarrow e\bar{K}^{*0}$

Figure B.4: Observed event distributions in $M_{\ell V^0}-\Delta E$ plane after the BDT selection trained with the control samples in the data and background MC. Black points are the data, blue squares show the signal MC distribution with an arbitrary normalization. The red ellipse is the signal region. The estimation of the number of background events is done using the data between the red horizontal lines except the blind region.