Doctorate Dissertation

博士論文

Measurement of the branching fraction of tau lepton decay to the final state of pion, lepton, lepton and neutrino at Belle

(Belle実験でのタウレプトンのパイオン＋レプトン＋レプトン＋ニュートリノへの崩壊の測定)

A Dissertation Submitted for Degree of Doctor of Philosophy

July 2018

平成30年7月博士(理学)申請

Department of Physics, Graduate School of Science,
The University of Tokyo

東京大学大学院理学系研究科物理学専攻

Yifan Jin

金 憶凡

Abstract

Using a 562 fb^{-1} dataset collected at the $\Upsilon(4S)$ resonance with Belle detector at the KEKB asymmetric-energy e^+e^- collider, the branching fractions for rare tau decays $\tau^\pm \to \pi^\pm l^+ l^- \nu_\tau$ are measured, where l is an electron or a muon.

The branching fraction of $\tau^\pm \to \pi^\pm e^+ e^- \nu_\tau$ is measured to be $\mathcal{B}(\tau^\pm \to \pi^\pm e^+ e^- \nu_\tau) = (2.33 \pm 0.19 \pm 0.30) \times 10^{-5}$, where the first uncertainty is statistical and the second is systematic. This result is the first measurement on this decay mode. In the case of $\tau^\pm \to \pi^\pm \mu^+ \mu^- \nu_\tau$, an upper limit on branching fraction is obtained, $\mathcal{B}(\tau^\pm \to \pi^\pm \mu^+ \mu^- \nu_\tau) < 0.55 \times 10^{-5}$, at 90% confidence level. This result is the first upper limit on this decay mode.

The measured results are consistent with the theoretical prediction from the Standard Model. The result of this measurement would contribute to future studies of physics beyond the Standard Model, such as lepton flavour violation searches.

Acknowledgements

First, I would like to express my deep gratitude towards Prof.Hiroaki Aihara and Dr.Denis Epifanov for their instruction and supervision upon this study. During the course of PhD training, one will continuously feel isolated and hopeless. Fortunately, every time I go out on a limb, their instruction successfully leads me through it. Without my professors' supervision, this study cannot be closed.

I am also grateful to members of Tau-2photon Group in Belle collaboration, especially Prof.Hayasaka and Prof.Uehara. I benefit tremendously from their suggestions and comments. Meanwhile, I would like to thank everyone in Aihara-Yokoyama Lab, Prof.Masashi Yokoyama, Prof.Yoshiyuki Onuki, Kuniko Kono, Dr.Nobuhiro Shimizu, Dr.Yusuke Suda, Taichiro Koga, Junya Sasaki, Hiroko Niikura and Kun Wan, for their help in my study and daily life. Besides, I would like to take this opportunity to express my sincere thanks to the Ministry of Education, Culture, Sports, Science and Technology of Japan for offering me the MEXT scholarship that covers my living expense and tuition fee here.

Finally, I would like to stress my love to my parents for their unceasing encouragement and support. Family is everything.

The harder thing to do and the right
thing to do are usually the same thing.
Nothing that has meaning is easy.
"Easy" doesn't enter into grown-up life.
≪The Weather Man≫

Contents

1 Introduction

Particle physics is a subject of science that studies the fundamental interactions amongst elementary particles in the universe. So far, four fundamental interactions are known to exist: gravitational, electromagnetic, strong and weak interactions. The latter three have been well accounted for by the Standard Model (SM), where nature is described in terms of quantum fields and interactions are described as an effect of interchanging quanta of fields (force carriers). Interchanging a specific type of bosons gives rise to the corresponding force, such as photons to electromagnetic interaction, gluons to strong interaction, W bosons and Z bosons to weak interaction. In addition, matter is composed of several types of fermions, which is further categorised into leptons: $\{e, \nu_e\}$, $\{\mu, \nu_\mu\}$, $\{\tau, \nu_\tau\}$, and quarks: $\{u, d\}$, $\{c, s\}$, $\{t, b\}$, in ascending order of mass[1].

The intensity of interactions is quantified by the running coupling constant, which is a function of the energy scale. Quantum electrodynamics (QED), a theory in the Standard Model describing the electromagnetic interaction, has unprecedentedly given a variety of the most accurate theoretical predictions in human history due to its tiny coupling constant that reduces the uncertainty substantially in the calculation using a perturbation theory. More intriguingly, an electroweak theory has been proposed where the electromagnetic and weak interactions are unified via an SM Lagrangian of SU(2)×U(1) gauge symmetry. Quantum chromodynamics (QCD), the counterpart of QED dealing with the strong interaction in the Standard Model, is enacted with a relatively large coupling constant, especially in low energy regimes, where the perturbation theory appears to be invalid. Hence, Chiral Perturbation Theory [1], an effective field theory constructed with a Lagrangian consistent with the chiral symmetry, is developed to study the low-energy (< 1 GeV) dynamics of strong interactions. To tackle increasing energy scale (≥ 1 GeV), various phenomenological approaches have been proposed, e.g., Resonance Chiral Theory [2], an extension of Chiral Perturbation Theory including the lightest multiplet of spin-1 resonances as active degrees of freedom.

In spite of the existence of three generations in the Standard Model, heavy fermions rapidly decay into their lighter replica via weak interactions, which leads to the abundance of only the first generation in ordinary matter. τ lepton is the only lepton massive enough to decay both leptonically and hadronically. Its hadronic final states provide an extremely clean laboratory to study the dynamics of strong interaction in the energy region below τ mass. Therefore, τ decays can serve as a perfect candidate to test Resonance Chiral Theory. The precision measurement of τ decay was stimulated by ALEPH, CLEO and OPAL at the end of last century. Afterwards, BaBar and Belle, two experiments dedicated to studying CP-violation in B mesons, also accumulated a remarkable amount of $\tau^+\tau^-$ pairs, which make it possible to measure the rare decays of τ lepton precisely. Numerous decay modes of τ have been measured by these two B-factories and offered crucial information to test the extrapolation from theory. In this dissertation, a 562 fb^{-1} dataset collected with Belle SVD2 at the $\Upsilon(4S)$ resonance using the KEKB asymmetric-energy e^+e^- collider is exploited to measure the branching fraction of a rare decay of τ, $\tau^\pm \to \pi^\pm l^+ l^- \nu_\tau$ ($l = e$ or μ), via a blind analysis of counting number. The branching fraction is so small that previous experiments have no access to explore it.

Physics motivations of the decay mode under consideration are highlighted in Chapter 2.

[1]The order of mass of ν is not determined yet.

Details of the experimental apparatus, including the accelerator KEKB, detector Belle and experimental and Monte Carlo simulation data sets are given in Chapter 3. Reconstructions and selections of the events collected by Belle are addressed in Chapter 4, followed by background study in Chapter 5. Finally, the extraction of branching fractions is carried out in Chapter 6. A conclusion is given at the end.

2 Motivations

Unlike e lepton and μ lepton, which have been studied extensively and with high precision, our knowledge of the τ lepton is relatively poor. However, decays of τ lepton have enormous potential to study the dynamics of strong interaction in the energy region below τ mass, to extract some fundamental parameters of the Standard Model and to test various phenomenological approaches, e.g., Resonance Chiral Theory. Moreover, rare decays of τ often serve as a probe of the physics beyond the Standard Model, since contributions of New Physics could give rise to observable deviations from the SM expectations.

2.1 Motivations within the Standard Model

Hadronic decays, or the dynamics of strong interactions that underlies it, has drawn persistent attention in particle physics. To tame the non-perturbative feature of strong interactions in low energy regime and carry through the model building in line with a sense of symmetry, an effective theory, Chiral Perturbation Theory, is constructed by writing down a Lagrangian of terms that are consistent with the chiral symmetry of QCD, symmetry of parity and charge conjugation. However, there is no doubt that light mesons do own non-zero mass, albeit light. As a result, the chiral symmetry is just an approximate symmetry and Chiral Perturbation Theory can only give reliable predictions at the energy region of several hundred MeV. To deal with the energy regime of ~ 1 GeV, the Lagrangian is extended by including resonances as active degrees of freedom, which is referred to as Resonance Chiral Theory. In order to determine the free couplings of these active fields, the short distance constraints of QCD in the limit of the infinite number of colours (N_c) are imposed [3].

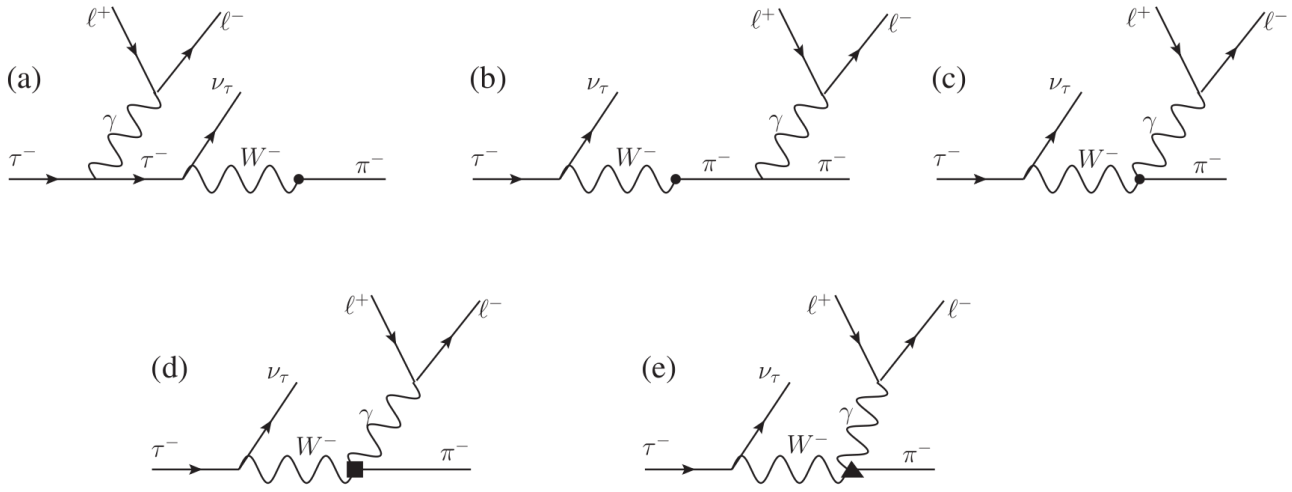


Figure 1: Feynman Diagrams of the decay modes under consideration: $\tau^- \to \pi^- l^+ l^- \nu_\tau$, $l = e$ or μ [4]. Diagrams (a), (b) and (c) indicate the internal bremsstrahlung due to the radiative γ off the τ, π and vertex. (d) and (e) indicate the contribution of γ emitted off the vector and axial-vector current mediating the hadronization.

The Feynman diagrams of the decay mode $\tau^- \to \pi^- l^+ l^- \nu_\tau$ are demonstrated in the Figure 1,

where diagrams (a) and (b) indicate the model independent contribution due to the internal bremsstrahlung of a radiative γ emitted off the τ and the π, (c) indicates the same internal bremsstrahlung effect of a γ emitted off the $\gamma^* W^{*\mp}\pi^{\pm}$ vertex assuming a point-like pion, (d) and (e) indicate the model dependent contribution of γ emitted off the vector and axial-vector current mediating the hadronization. In other words, these two correspond to the non-point-like part of the interaction. The decay amplitude can be composed of several gauge-invariant contributions [4].

$$
\begin{aligned}
\mathcal{M}[\tau^-(p_\tau) \to \pi^-(p)\nu_\tau(q)l^+(p_+)l^-(p_-)] &= \mathcal{M}_{IB} + \mathcal{M}_V + \mathcal{M}_A, \\
\mathcal{M}_{IB} =& -iG_F V_{ud}\frac{e^2}{k^2}F_\pi M_\tau \bar{u}(p_-)\gamma_\mu v(p_+)\bar{u}(q)(1+\gamma_5)[\frac{2p^\mu}{2p\cdot k + k^2} + \frac{2p_\tau^\mu - \not{k}\gamma^\mu}{-2p_\tau\cdot k + k^2}]u(p_\tau), \\
\mathcal{M}_V =& -G_F V_{ud}\frac{e^2}{k^2}\bar{u}(p_-)\gamma^\nu v(p_+)F_V(p\cdot k, k^2)\epsilon_{\mu\nu\rho\sigma}k^\rho p^\sigma \bar{u}(q)\gamma^\mu(1-\gamma_5)u(p_\tau), \\
\mathcal{M}_A =& iG_F V_{ud}\frac{2e^2}{k^2}\bar{u}(p_-)\gamma_\nu v(p_+)\{F_A(p\cdot k, k^2)[(k^2 + p\cdot k)g^{\mu\nu} - k^\mu p^\nu] - \frac{1}{2}A_2(k^2)k^2 g^{\mu\nu} \\
&+ \frac{1}{2}A_4(k^2)k^2(p+k)^\mu p^\nu\}\bar{u}(q)\gamma_\mu(1-\gamma_5)u(p_\tau),
\end{aligned}
$$

where $k = p_+ + p_-$, G_F is the fermi constant, V_{ud} is the first element in CKM matrix, and F_π is pion decay constant.

The model dependent contributions are described by one vector ($F_V(p\cdot k, k^2)$) and three axial-vector Lorentz-invariant form factor ($F_A(p\cdot k, k^2)$, $A_2(k^2)$ and $A_4(k^2)$). These form factors are calculated in the framework of the Resonance Chiral Theory [4].

As a result, it is convenient to express the decay rate in terms of six parts that correspond to three moduli squared terms (IB, VV, AA) and three interference terms (IB$-$V, IB$-$A, V$-$A). $\Gamma_{\rm total} = \Gamma_{\rm IB} + \Gamma_{\rm VV} + \Gamma_{\rm AA} + \Gamma_{\rm IB-V} + \Gamma_{\rm IB-A} + \Gamma_{\rm V-A}$. The central values of different contributions to the branching fraction of the $\tau^\pm \to \pi^\pm l^+ l^- \nu_\tau$, ($l = e,\ \mu$) decays is demonstrated in the Table 1, in the second column and third column for $l = e$ and $l = \mu$ respectively. The error of the branching fractions is presented as well in the forth and fifth columns. Therefore, a precise measurement of the branching fractions can be used to test the numerical result from Resonance Chiral Theory.

Table 1: Table of different contributions to the branching fraction of $\tau^- \to \pi^- l^+ l^- \nu_\tau$. "IB" indicates the sum of internal bremsstrahlung γ emitted off the τ lepton, π meson and $\gamma^* W^{*\mp} \pi^\pm$ vertex, whose Feynman Diagrams are shown in Figure 1, (a), (b) and (c) respectively. "V" indicates the model-dependent vector contribution, as shown in Figure 1 (d). "A" indicates the model-dependent axial-vector contribution, as shown in Figure 1 (e). "−" refers to interference terms [4].

	Central value of BR ($l = e$)	Central value of BR ($l = \mu$)	Error of BR ($l = e$)	Error of BR ($l = \mu$)
IB	1.461×10^{-5}	1.600×10^{-7}	$\pm 0.006 \times 10^{-5}$	$\pm 0.007 \times 10^{-6}$
IB − V	-2×10^{-8}	1.4×10^{-8}	$[-1 \times 10^{-7}, 1 \times 10^{-7}]$	$[-4 \times 10^{-9}, 4 \times 10^{-8}]$
IB − A	-9×10^{-7}	1.01×10^{-7}	$[-3 \times 10^{-6}, 2 \times 10^{-6}]$	$[-2 \times 10^{-7}, 6 \times 10^{-7}]$
VV	1.16×10^{-6}	6.30×10^{-7}	$[4 \times 10^{-7}, 4 \times 10^{-6}]$	$[1 \times 10^{-7}, 3 \times 10^{-6}]$
AA	2.20×10^{-6}	1.033×10^{-6}	$[1 \times 10^{-6}, 9 \times 10^{-6}]$	$[2 \times 10^{-7}, 6 \times 10^{-6}]$
V − A	2×10^{-10}	-5×10^{-11}	$\sim 10^{-10}$	$\sim 10^{-10}$
Total	1.710×10^{-5}	1.938×10^{-6}	$1.7^{+1.1}_{-0.3} \times 10^{-5}$	$[3 \times 10^{-7}, 1 \times 10^{-5}]$

2.2 Motivations beyond the Standard Model

Although the Standard Model has been regarded as a great triumph in modern physics, there are still a few issues that cannot be accounted for by this theory, such as, neutrino mass and oscillation, dark matter, baryon asymmetry of the universe, etc. Besides, the theory does not incorporate the gravitational interaction. Theorists are committed to develop New Physics models that extend the Standard Model to interpret these issues.

Rare decays of τ are also expected as probes to shed light on the physics beyond the Standard Model. Previously, in order to elucidate the excess of signal obtained in MiniBooNE experiment [5, 6], heavy long-lived sterile neutrinos [7], have been introduced with following parameter ranges: mass $\in [400, 600]$ MeV, lifetime $< 1 \times 10^{-9}$ s and mixing strength $\in [1 \times 10^{-3}, 4 \times 10^{-3}]$. The effect of such a sterile neutrino in the decay $\tau^- \to \pi^- \mu^+ \mu^- \nu_\tau$ has been studied [8], which gives an upper limit $\leq 1.3 \times 10^{-5}$ on the branching fraction of $\tau^- \to \pi^- \mu^+ \mu^- \nu_\tau$. In other words, if such a sterile neutrino does enter the diagram, as shown in Figure 2, it would result in an observable enhancement to the branching ratio of $\tau^- \to \pi^- l^+ l^- \nu_\tau$. Therefore the measured result could be used to constrain the parameters of the proposed sterile neutrino.

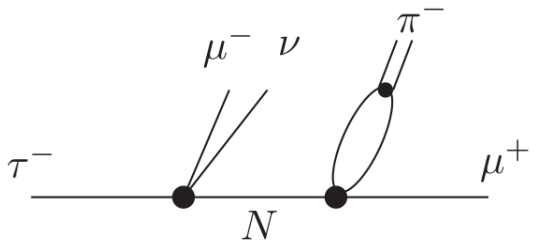


Figure 2: τ lepton decays into a μ, a μ neutrino and a sterile neutrino N, which further decays into a π and a μ [8]. The loop connecting with N and π indicates the $\bar{u}d$ pair. The final state is the same as that of the decay mode under consideration.

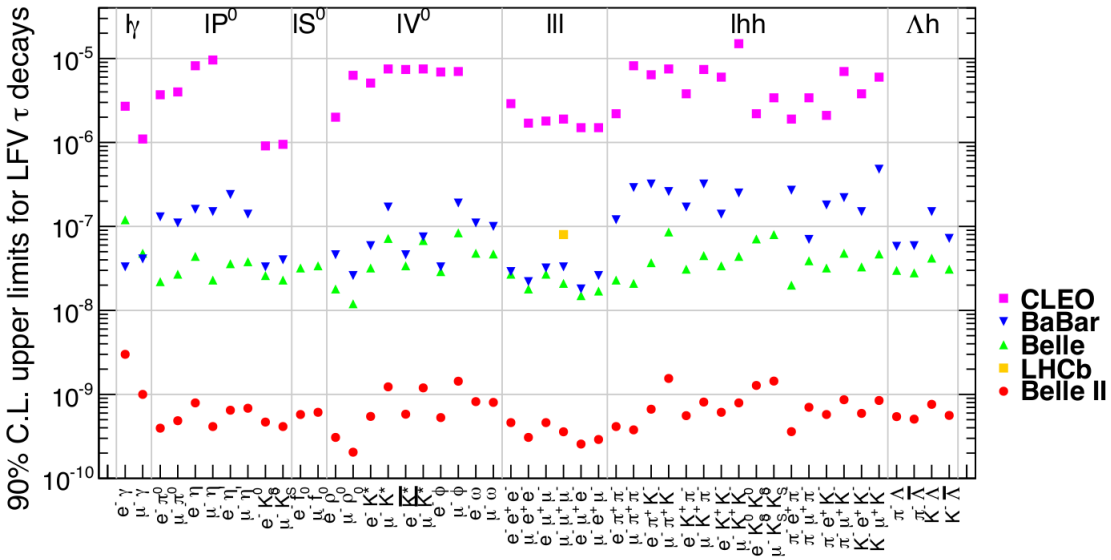


Figure 3: τ LFV summary and prospects, overlaid with Belle II extrapolation (red dots) to 50 ab^{-1} assuming zero background [9]. The top four types of marks are experimental results from CLEO (pink squares), BaBar (blue triangles), Belle (green triangles) and LHCb (yellow square).

Moreover, a precise measurement of $\mathcal{B}(\tau^{\pm} \to \pi^{\pm} l^{+} l^{-} \nu_{\tau})$ also facilitates searches for lepton flavour violated (LFV) modes. For example, if the π is misidentified as a μ (the fake rate is at the order of 10^{-2}), the decay mode under consideration could mimic the signature of LFV decays, $\tau \to \mu\ l^{+} l^{-}$, or oppositely, the μ is misidentified as a π, then it could mimic $\tau \to \mu\ \pi^{-} \pi^{+}$, and $\tau \to \pi\ l^{+} l^{-}$ when the ν_{τ} is carrying exceedingly little momentum. As shown in Figure 3, up to now, from Belle (the green triangles) and BaBar (the blue triangles) experiments, the upper limits of these LFV modes have already been suppressed to the order of 10^{-8}. With the coming Belle II data, the upper limits of LFV modes are expected to be further reduced to the order of 10^{-9}. The $\tau^{\pm} \to \pi^{\pm} l^{+} l^{-} \nu_{\tau}$ decay mode, whose branching fractions are at the order of 10^{-5}, would hinder the LFV searches as a notable background. For this purpose, a precise measurement of branching fractions of $\tau^{\pm} \to \pi^{\pm} l^{+} l^{-} \nu_{\tau}$ is indispensable.

Last, in this analysis, the charge conjugated modes $\tau^{-} \to \pi^{-} l^{+} l^{-} \nu_{\tau}$ and $\tau^{+} \to \pi^{+} l^{+} l^{-} \bar{\nu}_{\tau}$ are measured separately. As the decay modes of concern only involve the first element in CKM matrix, V_{ud}, there is supposed to be no CP violation between the charge conjugated modes. As a result, if any difference is observed on the branching fractions between the charge conjugated modes, it would be a clear sign for New Physics effect.

3 Experimental apparatus and data sets

Since the establishment of the Standard Model, various predictions given by it have been verified experimentally. Among these experimental apparatus, colliders have been playing important roles in particle physics for decades. Fruitful achievements have emerged out of colliders.

The development of colliders can be categorised into two directions. One is to increase the energy of the beam, known as energy frontier. The other is to increase the luminosity, namely intensity frontier. At energy frontier (represented by LHC), physicists aim to directly discover new particles that could lead to an answer to questions like the origin of dark matter and imbalance between matter and anti-matter. At intensity frontier (as in this dissertation, Belle), physicists are looking for extremely rare processes to understand some mysterious properties of the elementary particles and to reveal new fundamental principles of physics. Two of them function as complementary tools in collider physics. The output from either of these two frontiers gives insight into the other.

In this dissertation, the decay mode of concern is analysed by using the data set collected at Belle detector, operated at KEKB accelerator in Tsukuba, Japan. Thanks to the great effort of Belle collaboration, Belle detector achieved the world's highest instantaneous luminosity during its data taking. The integrated luminosity is displayed in Figure 4.

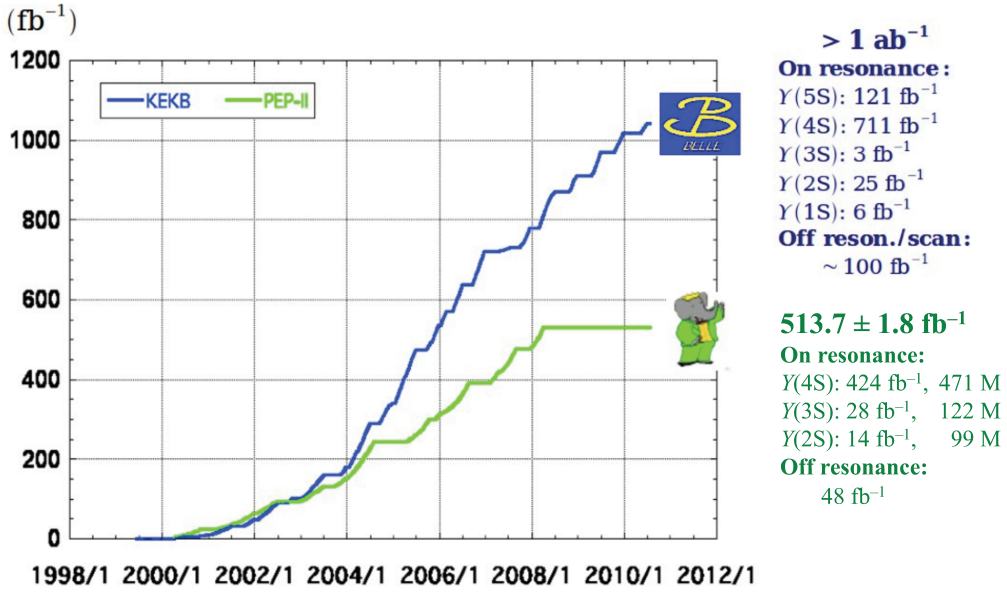


Figure 4: Belle integrated luminosity (blue line) as a function of time in comparison to that of BaBar (green line) [12]. The integrated luminosity collected at different center of mass energies are also displayed on right side.

3.1 KEKB accelerator

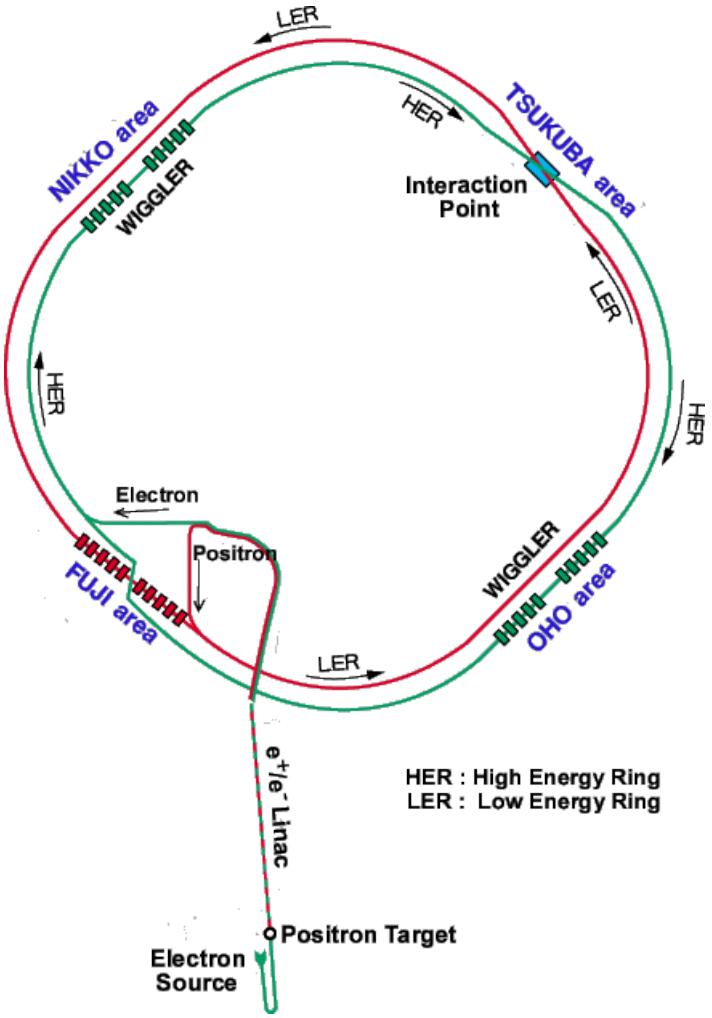


Figure 5: Schematic view of KEKB Accelerator [10]. The Belle detector is installed on the interaction point in Tsukuba area.

KEKB is an accelerator with asymmetric energy beams (8 GeV e^- on 3.5 GeV e^+) in a 3-kilometre long tunnel [14], equipped with a 508 MHz RF system [15]. The injection linear accelerator (Linac) provides a beam of 8 GeV electrons to the High Energy Ring and a beam of 3.5 GeV positrons to the Low Energy Ring. These two rings are aligned side-by-side and intersect in the Tsukuba Hall where Belle detector is located, as shown in Figure 5. Two beams collide with a finite crossing angle (21 mrad) so as to meet the short bunch spacing and offer more available space for inner components of Belle detector near interaction point of beams. Superconducting crab cavities (see Figure 6) were installed to increase the luminosity by rotating the beams to accomplish a head-on collision [16]. In addition, the rotation cancels off the unwanted transverse-longitudinal couplings induced by crossing angle between betatron and synchrotron motions. Eventually, KEKB reached a peak luminosity of 2.1×10^{34} cm^{-2}s^{-1}, which is twice of the originally designed luminosity.

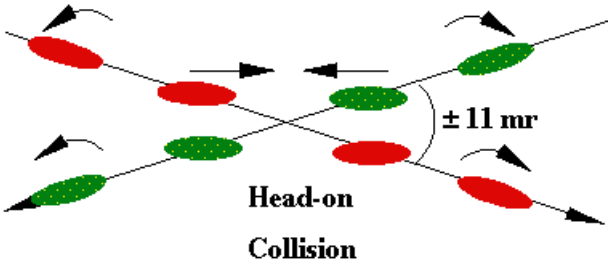


Figure 6: Demonstration of the Crab technic [10].

3.2 Belle detector

The Belle detector is a general purpose 4π spectrometer that consists of several sub-detectors. From innermost to outmost, there is a silicon vertex detector (SVD), a central drift chamber (CDC), an aerogel threshold Cherenkov counter (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL) located inside a superconducting solenoid coil that provides a 1.5 T magnetic field, as shown in Figure 7. An iron flux return located outside of the coil is instrumented to detect K_L^0 mesons and to identify muons (KLM). A detailed introduction of Belle detector can be found elsewhere [13]. Here, only brief descriptions of the key components are presented.

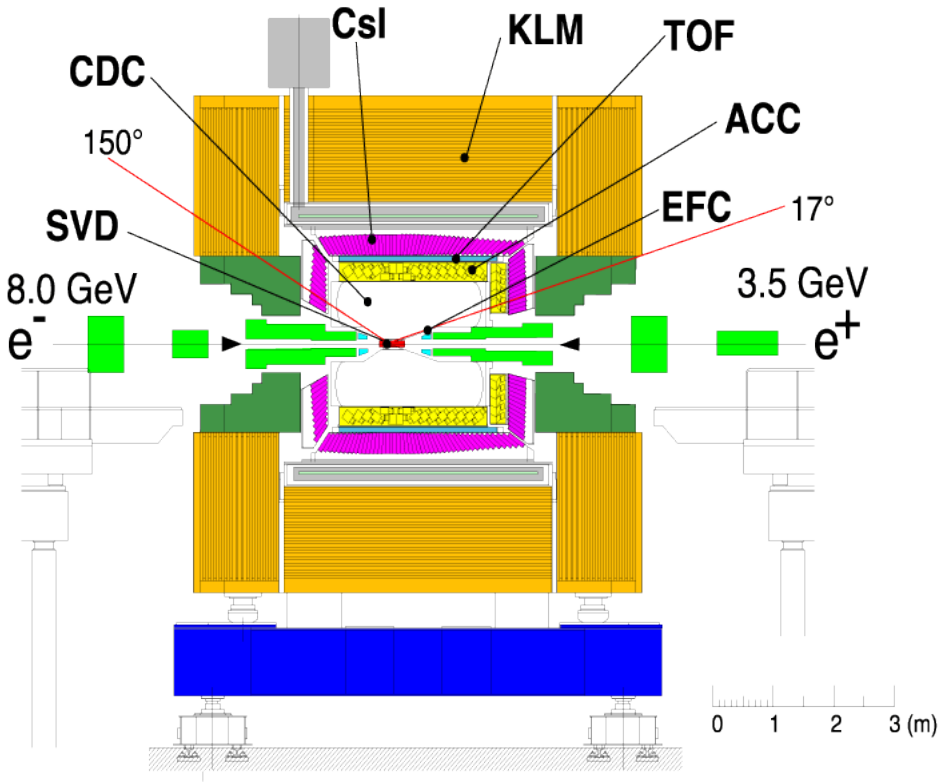


Figure 7: Side view of Belle detector [10]. The z-axis is in the opposite direction of incoming positron beam, the x-axis is horizontal and y-axis is vertical away from the center of KEKB ring.

Beam Pipe

The beam pipe [17] is a concentric double-wall beryllium cylinder surrounding the e^+e^- beams with an inner diameter only 30 mm to make room for the SVD2 inner detector so as to optimise vertex resolution. For the purpose of cooling, helium gas is ventilated through a 2.5 mm gap between the inner and outer walls. In order to prevent synchrotron radiation photons penetrating into the detector, a 10 μm-thick layer of gold is sputtered on the surface of the beryllium wall.

Silicon Vertex Detector (SVD)

Since October 2003, SVD2 is employed for data taking. This upgraded Silicon Vertex Detector consists of 4 layers of double-sided silicon-strip detectors (DSSD) and covers the full polar angle acceptance from $17° - 150°$. The radii of the SVD layers are 20, 44, 70 and 88 mm, respectively. The readout system of the DSSD is based on VA1TA chips that have a 0.8 μsec peaking time and a 20 Mrad radiation tolerance [18]. The main function of SVD is to interpolate tracks reconstructed in Central Drift Chamber to the decay vertices within the range of beam pipe. The resolution of SVD for a single hit is 12 μm in the transverse direction and 19 μm in the longitudinal direction [19]. The resolution of a track's impact parameter reaches to $\sigma_r = 21.9 \oplus 35.5/p$ μm and $\sigma_z = 27.8 \oplus 31.9/p$ μm [20], respectively, where p indicates the momentum of track in units of GeV/c and $\oplus$ indicates summation in quadrature. The resolution of vertex reaches to ~ 100 μm so as to measure the distance between two B decays.

Central Drift Chamber (CDC)

The Central Drift Chamber has an asymmetric structure in the longitudinal direction so as to cover the full polar angle acceptance ($17° - 150°$). The chamber is composed of 50 cylindrical layers, each containing $3 - 6$ either axial or small-angle stereo layers, and three cathode strip layers. In total, CDC has 8400 drift cells. For the two innermost super-layers, each is composed of three layers while for the three outer stereo super-layers, each is composed of four layers. A gas mixture of 50% Helium and 50% Ethane (C_2H_6) [21] is used in the chamber, owing to its low-Z nature that can minimise multiple scattering. Radeka-type pre-amplifiers are employed to readout signals from CDC and transmit via 30 m long twisted pair cables to the electronics modules for shaping, discrimination and charge to time conversion. CDC provides several functions. First, it measures the hit coordinates of charged tracks precisely in the detector volume, which offers information for tracks' momenta measurement. Meanwhile, the specific ionisation energy loss dE/dx from CDC can offer information to identify the particle. The spatial resolution of a single hit in CDC reaches 130 μm in the transverse direction and 200 $\sim$ 1400 μm in the longitudinal direction, which give rise to an excellent resolution on the transverse momentum: $\sigma_{P_t}/P_t = 0.0019P_t \oplus 0.003/\beta$, where P_t is in units of GeV/c. The resolution on dE/dx obtains 7% for minimum ionising particles.

Aerogel Cherenkov Counters (ACC)

The Aerogel Cherenkov Counters are composed of two parts, a barrel section covering $33.3° - 127.9°$ and a forward endcap section covering $13.6° - 33.4°$. The barrel part is segmented into 60 identical sectors, each containing 16 modules, in the ϕ direction. The forward endcap part consists of 5 layers with 228 modules in total. Based on their polar angle, the modules are filled with aerogel radiators of different refractive index, varying from 1.01 to 1.03, and wrapped with a white reflector whose reflectivity is higher than 93% [22]. In order to maximise the collection of Cherenkov light and obtain a uniform response, one or two fine mesh-type photomultiplier tubes (FM-PMTs) of different size (2" or 2.5" or 3") are attached to the aerogel module depending on the refractive index of the corresponding aerogel module. The output of PMTs is amplified by a pre-amp base then sent to a charge to time converter and digitised by a LeCroy 1877 TDC. ACC contributes substantially in the particle identification, especially in K/π separation for charged particles with momenta from 1.2 to 3.5 GeV/c.

Time of Flight Counter (TOF)

The time-of-flight system consists of a barrel-like arrangement of 64 TOF/TSC modules, each containing 2 plastic scintillator counters and 1 thin trigger scintillation counters (TSC), located at a radius of 1.2 m away from the interaction point of beams with a polar angle acceptance $33° - 121°$. The TOF and TSC scintillators (BC408, Bicron) are wrapped by a 45 μm thick layer of polyvinyl film. Hamamatsu (HPK) type R6680 fine-mesh photomultipliers, with a 2-inch diameter, are attached to both ends of the TOF counter with an air gap of 0.1 mm. The main function of TOF is to distinguish between kaons and pions that hit the TOF with momenta less than 1.2 GeV/c by its outstanding time resolution ($\sigma_t = 100$ ps for μ tracks) [23].

Electromagnetic Calorimeter (ECL)

The Electromagnetic Calorimeter is composed of a barrel section and two endcap sections, containing 6624 (barrel), 1152 (forward) and 960 (backward) CsI(Tl) crystals respectively. ECL covers the full acceptance of Belle ($12.4° - 31.4°, 32.2° - 128.7°, 130.7° - 155.1°$). Scintillation photons from CsI(Tl) crystals of approximate dimension $\sim 5.5 \times 5.5 \times 30$ cm^3 are collected by PIN photodiodes, converted into charge pulses and fed to charge sensitive pre-amplifiers [24]. ECL provides the information for electron/hadron separation. Photons are reconstructed from ECL clusters from those not matching with tracks. For a 5×5 matrix, the energy resolution of ECL can be expressed as $\sigma_E/E = 0.066\%/E \oplus 0.81\%/\sqrt[1/4]{E} \oplus 1.34\%$. The spatial resolution reaches to $\sigma_{\text{position}} = 3.4/\sqrt{E} \oplus 1.8/\sqrt[1/4]{E} \oplus 0.27$ mm (E is in units of GeV).

K_L^0 and Muon Detector (KLM)

The K_L^0 and Muon Detector [25] consists of alternating layers of charged particle detectors

and iron plates of 4.7 cm thickness. There are 15 detector layers and 14 iron layers in the octagonal barrel region, covering a polar range $45° - 125°$, and 14 detector layers and 14 iron layers in each end cap ($20° - 45°$, $125° - 155°$). The active elements in the charged particle detectors are glass-electrode resistive plate counters (RPCs), where a 2 mm gas gap (a non-flammable gas mixture of 62% HFC-134a, 30% argon and 8% butane-silver) is sandwiched between two layers of float glass. 2.4 mm thick float glass (73% SiO_2, 14% Na_2O, 9% CaO, and 4% trace elements) are used at barrel RPCs while 2.0 mm thick float glass (70-74% SiO_2, 12-16% Na_2O, 6-12% CaO, 0-2% Al_2O_3 and 0-4% MgO) are used at end cap RPCs. Discharge signals from RPCs, transmitted via external pickup strips, are read out by custom-made VME-based discriminator/time multiplexing boards. KLM identifies K_L^0 mesons and muons above 600 MeV/c with very high efficiency (μID efficiency larger than 90% with a fake rate less than 5% for particles with momenta over 1.5 GeV/c).

Extreme Forward Calorimeter (EFC)

The Extreme Forward Calorimeter covers the angular range from 6.4° to 11.5° in the forward direction and 163.3° to 171.2° in the backward direction. Pure BGO crystals of approximate dimension 2 cm × 1.5 cm × 12 cm with silicon photodiodes are used due to their great radiation hardness and excellent energy resolution $(0.3 \sim 1)\%/\sqrt{E}$. EFC serves as a beam monitor for the KEKB control and a luminosity monitor for the Belle experiment. Meanwhile, it can also function as a beam mask to reduce backgrounds on CDC.

Trigger

The trigger system of Belle consists of two levels, a Level-1 hardware trigger and a Level-3 software trigger. The Level-1 hardware trigger is composed of sub-triggers from sub-detectors (CDC, ECL, TOF, and KLM) and a global decision logic (GDL) [26]. The CDC and TOF output track triggers for charged particles. The ECL provides triggers based on total deposited energy and counting of crystal hits in a cluster. The KLM provides additional information on muons. The GDL, after combining all the information from sub-detectors, make a logical determination of the event to trigger on hadronic events, Bhabha and $\tau^+\tau^-$ pair events, etc. The Level-3 software trigger is implemented on the online computer farm.

Data Acquisition System (DAQ)

Initially, the front-end electronics of sub-detectors were read out by charge-to-time converters combined with the common FASTBUS multi-hit TDC (except for SVD, where flash ADC is employed), followed by the VME processor. A specially-designed event builder then collects data from these VMEs and pass data to the VME based online computer farm where L3 software trigger was performed. With the growth of luminosity, the FASTBUS based readout system first was upgraded into a pipelined version then replaced by a COPPER based TDC in order to reduce the dead time. The event builder and VME based online computer farm

were replaced by a set of Linux PC servers (EFARM) in 2001. In 2003, the luminosity was improved substantially by introducing the real-time reconstruction farm (RFARM) next to the event builder, which can feed back IP reconstruction information to accelerator control after its parallel processing of data receiving from the event builder. A schematic demonstration of final DAQ system is given in Figure 8.

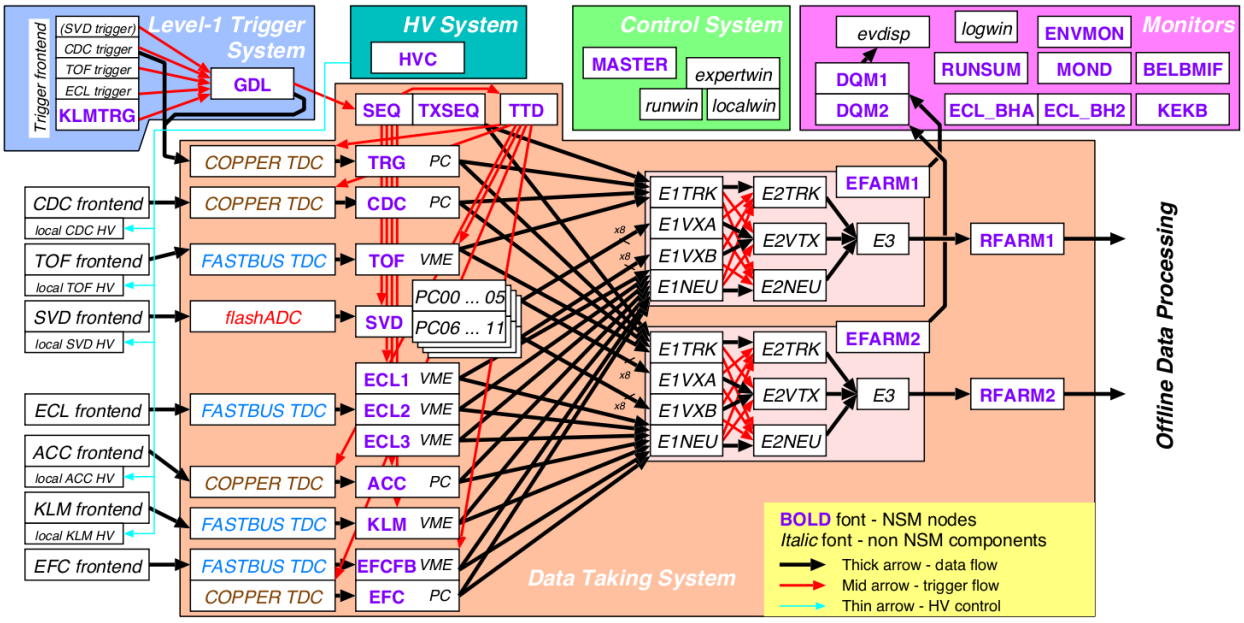


Figure 8: The configuration of Belle DAQ system at the end of data taking [11].

Particle Identification algorithm

At Belle, a likelihood-based selector is employed to compute the likelihood of a charged track into a variety of particle types based on the information from several sub-detectors: CDC, ACC, TOF, ECL and KLM. In the case of hadron identification, a likelihood of being type a is calculated in comparison to being type b, using informations such as specific ionisation energy loss dE/dx from CDC, time of flight from TOF and light yield in ACC, by the following formula: $L(a:b) = \frac{L_a^{\rm CDC} \cdot L_a^{\rm TOF} \cdot L_a^{\rm ACC}}{L_a^{\rm CDC} \cdot L_a^{\rm TOF} \cdot L_a^{\rm ACC} + L_b^{\rm CDC} \cdot L_b^{\rm TOF} \cdot L_b^{\rm ACC}}$. For example, kaon candidates are selected by a cut on $L(K:\pi)$. In the case of electron identification, a similar formula is constructed by $L_a^{\rm ACC}$, $L_a^{\rm CDC}$ and $L_a^{\rm ECL}$ for electron likelihood in comparison to being a hadron. Here ECL provides information such as the ratio of the shower energy to the track momentum (E/p), shower shape of the cluster (energy deposited in the central 9 crystals/energy deposited in the 25 crystals) and position matching of track in CDC and energy cluster in ECL. In the case of muon identification, the penetration length in KLM and matching between the track and KLM hits are used to calculate the likelihood ratio $\frac{L_\mu}{L_\mu + L_\pi + L_K}$.

The overview of Belle detector is summarised in Table 2.

Table 2: summary of Performance Parameters of the Belle Detector [10].

Detector	Type	Configuration	Readout	Performance
Beam pipe	Berylium double-wall	Cylindrical, r = 2.3 cm 0.5mm Be/2mm He/0.5mm Be		
SVD	Double sided silicon strip	4 layers Layer 1-3: $79.2 \times 28.4 \times 0.3$ cm^3 Layer 4: $76.4 \times 34.9 \times 0.3$ cm^3	ϕ: 41K θ: 41k	$\sigma_{\triangle z} \sim 100$ μm
CDC	Small cell drift chamber	Anode: 52 layers Cathode:3 layers r = 8.5 - 90 cm $-77 \leq$ z ≤ 160 cm	Anode: 8.4k Cathode: 1.5k	$\sigma_{r\phi} = 130$ μm $\sigma_z = 200 \sim 1400$ μm $\sigma_{P_t}/P_t = 0.3\% \cdot \sqrt{(P_t + 1)}$ $\sigma_{dE/dx} = 7\%$
ACC	n:1.01$\sim$1.03 Silica Aerogel	$\sim 12 \times 12 \times 12$ cm^3 blocks 960 barrel / 228 endcap FM - PMT readoput	1788 ch	$N_{p.e.} \geq 6$ K/π seperation: 1.2$\leq$ p $\leq$3.5 GeV/c
TOF	Plastic Scintillator	128 ϕ segmentation r = 120 cm 3 m long	128$\times$2 ch	σ_t =100ps K/π seperation: up to 1.2 GeV/c
ECL	CsI	Tower structure $\sim 5.5 \times 5.5 \times 30$ cm^3 crystals Barrel: r = 125 -162 cm Endcap:z = -102 and $+196$ cm	6624 (B) 1152 (FE) 960 (BE)	σ_E/E $=0.066\%/E \oplus 0.81\%/\sqrt[1/4]{E} \oplus 1.34\%$ $\sigma_{\rm pos} = 3.4/\sqrt{E} \oplus 1.8/\sqrt[1/4]{E} \oplus 0.27$ mm E in GeV
Magnet	Superconducting	inner radial = 170 cm		B = 1.5T
KLM	Resistive Plate Counter	14 layers (5cm Fe + 4cm gap) 2RPCs in each gap θ and ϕ strips	θ: 16k ϕ: 16k	$\triangle\phi = \triangle\theta = 30$ mrad for K_L σ_t =1 ns 1% hadron fakes
EFC	BGO	2 cm$\times$1.5 cm$\times$12 cm	θ: 5 ϕ: 32	$\sigma_E/E = (0.3 \sim 1)\%/\sqrt{E}$

3.3 Data set

Belle started data taking on the 1^{st} of June 1999 and ended on the 30^{th} of June 2010. "Runs" carried out between two major shutdowns are named "Experiments", which is enumerated from 7 to 73 (odd numbers only). Data taken in experiment number $7 - 27$ is collected by SVD1 and the rest is collected by SVD2. Most data are taken at the energy of $\Upsilon(4S)$ resonance to study B meson, with a 60 MeV lower off-resonance energy spot to estimate background from non-B meson processes. Besides, non-$\Upsilon(4S)$ data was also taken to supplement physics programs at Belle. 15% of the entire data is taken by SVD1, and 85% is taken by SVD2. In this analysis, the data taken by SVD2 at the Υ(4S) on-resonance is exploited.

Owing to the huge amount of data collected by the detector experimentally and simulated by Monte Carlo, event skimming is introduced so as to make analyses more efficient. Instead of reading the full data, analyses could be carried out with certain relevant skims that are made of events passing corresponding specific criteria. In Belle, after data processing, events are classified into several categories: Hadronic events skim (HadronBJ), Tau pair skim (TauSkim), low-multiplicity skim (LowMult), Bhabha events, muon pair events, γ events and so on. The first three are used for physics analyses while the latter three are used to detector calibration. As indicated by the name, HadronBJ is established mainly for facilitating physics analysis related to B and charm mesons, similarly, TauSkim for τ and LowMult for two-photon processes.

3.3.1 Experimental data (EXP data)

The Υ(4S) on-resonance data taken with SVD2 and non-zero `trigger type` (see below) is exploited in this dissertation. Experiment numbers are from 31 to 65. Two skims, HadronBJ and TauSkim, are exploited. The integrated luminosity is 562 fb^{-1}, corresponding to 517 million $\tau\tau$ pair events.

3.3.2 Generic Monte Carlo data

MC generators employed by Belle include KKMC [27] for fermion pairs, AAFH [28] for two-photon production of fermion pairs, CTOY [29] for cosmic ray muon, and EVTGEN [30] for hadronic processes. The decay of τ lepton is carried out by TAUOLA [31], where the spin polarisation effect of the τ lepton is taken into account via the spin density matrix and the spin correlation between the $\tau\tau$ pair are also considered. The final-state radiation of charged tracks is taken into account by using PHOTOS [32]. The response of the detector is simulated by a GEANT3-based program [33]. The trigger is simulated by the standard Belle module TSIM [34], which outputs `trigger type`. In this analysis, `trigger type` is not specified, only trigger efficiency is employed (trigger efficiency = number of events that have non-zero `trigger type` / number of all events, see Section 6.2).

In order to study the background contamination, generic $\tau^+\tau^-$, Bhabha, di-muon, two-photon and $q\bar{q}$ MC samples are exploited. Numbers of events of specific processes are enumerated in Table 3. The amounts of MC samples for all the processes are at least three times larger

than the experimental data (except Bhabha process). Generic $\tau^+\tau^-$, which is the dominant background in this analysis, is studied by a MC sample equivalent to 10 times the integrated luminosity. Detailed processes included in generic $\tau^+\tau^-$ MC are also listed in Table 4. In order to elaborate on the background from $\tau\tau$ processes at Belle, the MC of one additional rare decay of τ, $\tau^- \to l^-l^+l^-\nu_\tau\nu_l$ ($l = e$ or μ) (not included in generic $\tau^+\tau^-$ MC) is also exploited.

During the development of Belle experiment, the MC samples have been upgraded by sophisticated corrections. In this analysis, two well-established corrections on particle identification [36] and tracking [37] effects have been applied. In addition, two corrections concerning the π^0 reconstruction (Section 4.2.1) and $\pi \to \mu$ mis-identification (Section 4.3.1) have been carried out and applied to the MC samples.

3.3.3 Monte Carlo generator for signal mode

Based on the recently published analytical formalism for the $\tau^- \to \pi^-l^+l^-\nu_\tau$ decays [4], a Monte Carlo event generator was developed and embedded in the TAUOLA generator. Three million events of $e^+e^- \to \tau^+\tau^-$, $\tau^+\tau^- \to$ (all known decays of $\tau^\pm; \tau^\mp \to \pi^\mp l^+l^-\nu_\tau$) were generated by the aforementioned KK-TAUOLA based program, where one τ decays into the mode of concern, the other τ decays into all known modes with the branching fractions listed in latest PDG [38]. The detector response is simulated by the standard Belle GEANT3-based program.

All MC samples are reconstructed with the same reconstruction programs (tracking, particle identification etc) as the one used for the experimental data sets.

Table 3: List of background processes with cross section and the number of events of corresponding MC samples [35].

Process	cross section at $\Upsilon(4S)$	number of events
$\tau^+\tau^-$	0.919 ± 0.003 nb	5.0 Billion
Bhabha	123.5 ± 0.2 nb	15.4 Billion
di-muon	1.005 ± 0.001 nb	2.8 Billion
2 photon-eeee	40.9 nb	74.2 Billion
2 photon-ee$\mu\mu$	18.9 nb	34.3 Billion
2 photon-eeuu/eedd	12.498 nb	66.0 Billion
2 photon-eecc	0.03 nb	169 Million
2 photon-eess	0.227 nb	1.3 Billion
B^+B^-	0.525 nb	1.8 Billion
B^0B^0	0.525 nb	1.8 Billion
uds	2.09 nb	7.0 Billion
charm	1.3 nb	4.3 Billion

Table 4: τ lepton with $\mathcal{B}$ in generic $\tau^+\tau^-$ MC.

Decay mode	$\mathcal{B}$(TAUOLA)	Decay mode	$\mathcal{B}$(TAUOLA)
$\bar{\nu_e}e^-\nu_\tau$	0.1785	$\eta\pi^-\pi^0\nu_\tau$	0.00139
$\bar{\nu_\mu}\mu^-\nu_\tau$	0.1736	$\pi^-\pi^0\gamma\nu_\tau$	0.0013
$\pi^-\nu_\tau$	0.1091	$K^-K^0(K^0_S)\nu_\tau$	0.00159×0.5
$\rho^-\nu_\tau$	0.2551	$K^-K^0(K^0_L)\nu_\tau$	0.00159×0.5
$a_1^-(\pi^+\pi^-\pi^-)\nu_\tau$	0.1832×0.496	$\pi^-\pi^0\pi^0\pi^0\pi^0\nu_\tau$	0.0011
$a_1^-(\pi^-\pi^0\pi^0)\nu_\tau$	0.1832×0.504	$\pi^-\omega\pi^0\nu_\tau$	0.0041
$K^-\nu_\tau$	0.00696	$\pi^-\pi^+\pi^-\eta\nu_\tau$	0.000164
$K^{*-}(K^0_S\pi^-)\nu_\tau$	0.01269×0.333	$\pi^-\pi^0\pi^0\eta\nu_\tau$	0.00015
$K^{*-}(K^0_L\pi^-)\nu_\tau$	0.01269×0.333	$f_1\pi^-\nu_\tau$	0.00036
$K^{*-}(K^-\pi^0)\nu_\tau$	0.01269×0.0334	$K^-\omega\nu_\tau$	0.00041
$\pi^-\pi^-\pi^+\pi^0\nu_\tau$	0.0448	$K^-\eta\nu_\tau$	0.000161
$\pi^-\pi^0\pi^0\pi^0\nu_\tau$	0.0104	$K^{*-}\eta\nu_\tau$	0.000138
$\pi^-\pi^-\pi^+\pi^0\pi^0\nu_\tau$	0.001	$K^-\pi^+\pi^-\pi^0\nu_\tau$	0.00037
$\pi^-\pi^-\pi^-\pi^+\pi^+\nu_\tau$	0.000839	$K^-\pi^0\pi^0\pi^0\nu_\tau$	0.00049
$\pi^-\pi^-\pi^+\pi^0\pi^0\pi^0\nu_\tau$	0.00023	$\bar{K}(K^0_S)\pi^-\pi^+\pi^-\nu_\tau$	0.00023×0.5
$\pi^-\pi^-\pi^-\pi^+\pi^+\pi^0\nu_\tau$	0.000178	$\bar{K}(K^0_L)\pi^-\pi^+\pi^-\nu_\tau$	0.00023×0.5
$K^-\pi^-K^+\nu_\tau$	0.0014	$\bar{K}(K^0_S)\pi^-\pi^0\pi^0\nu_\tau$	0.00026×0.5
$K^0\bar{K}(K^0_SK^0_L)\pi^-\nu_\tau$	0.0017	$\bar{K}(K^0_L)\pi^-\pi^0\pi^0\nu_\tau$	0.00026×0.5
$K^-K^0(K^0_S)\pi^0\nu_\tau$	0.00159×0.5	$\pi^-K^+K^-\pi^0\nu_\tau$	0.000061
$K^-K^0(K^0_L)\pi^0\nu_\tau$	0.00159×0.5	$\pi^-K^0\bar{K}(K^0_S)\pi^0\nu_\tau$	0.00031×0.5
$\pi^0\pi^0K^-\nu_\tau$	0.00065	$\pi^-K^0\bar{K}(K^0_L)\pi^0\nu_\tau$	0.00031×0.5
$K^-\pi^+\pi^-\nu_\tau$	0.00342	$\pi^-\omega\pi^-\pi^+\nu_\tau$	0.00012
$\pi^-\bar{K}(K^0_S)\pi^0\nu_\tau$	0.004×0.5	$\pi^-\omega\pi^0\pi^0\nu_\tau$	0.00014
$\pi^-\bar{K}(K^0_L)\pi^0\nu_\tau$	0.004×0.5		

4 Event selection

In this chapter, the reconstruction and selection of events are discussed. Using MC samples that have been corrected by particle identification and tracking effects established in previous Belle studies, selection criteria are optimised by maximising a Figure-of-Merit (FOM), which is defined as $\frac{S}{\sqrt{S+B}}$, where S stands for the number of signal events and B for the number of background events. The standard $\tau\tau$ pre-selection of Belle is employed first to select $\tau\tau$ events then followed by additional selections to select the events in which one τ decays into the mode of concern ($\tau^{\pm} \to \pi^{\pm} l^{+} l^{-} \nu_{\tau}$) while the other τ decays into one prong, as shown in Figure 9. During optimisation, the branching fraction of $\tau^{\pm} \to \pi^{\pm} e^{+} e^{-} \nu_{\tau}$ is set as 1.7×10^{-5}, the central value of theoretical prediction, and the branching fraction of $\tau^{\pm} \to \pi^{\pm} \mu^{+} \mu^{-} \nu_{\tau}$ is set as 1×10^{-5}, the maximum of theoretical prediction, due to the huge uncertainty of theoretical range.

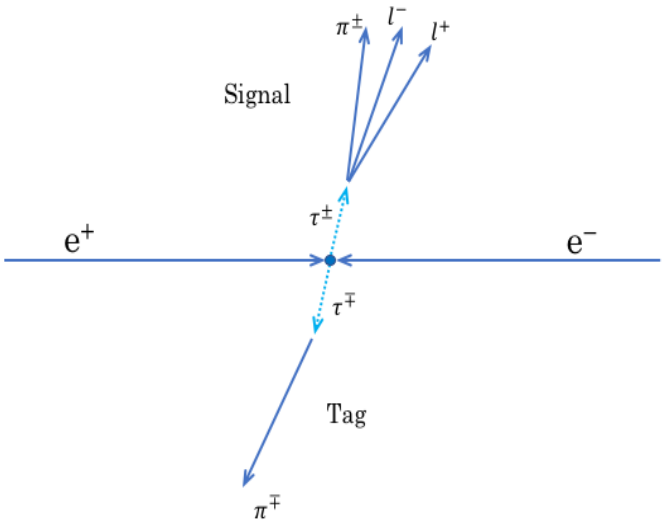


Figure 9: The schematic demonstration of event topology.

4.1 Pre-selection of $\tau\tau$ pair

As shown in Table 3, numerous processes occur at $\Upsilon(4S)$. At Belle, a standard pre-selection has been developed [39] to select $\tau\tau$ pair samples and to suppress other background processes as much as possible.

- The number of charged tracks is required to be more than 2 while less than 8 and the net charge should be no more than 2.
- The energy deposited in the ECL (charged tracks + neutral clusters) should be less than 10 GeV.
- In the laboratory frame, at least one charged track should have transverse momentum (P_{t}) larger than 0.5 GeV/c.

- The primary vertex (obtained by fitting all tracks event by event) should be close to the interaction point (IP) (obtained by averaging the primary vertex of multi-hadron events in one entire run), the distance in $x - y$ plane is required to be less than 0.5 cm and less than 3 cm in longitudinal direction.

- In the centre of mass frame of the e^+e^- beams (CM always refers to the centre of mass frame of the e^+e^- beams), the sum of momentum moduli of all charged tracks ($\sum |P^{\rm CM}|$) should be less than 10 GeV/c.

- The sum of momentum moduli of all charged tracks ($\sum |P^{\rm CM}|$) plus the sum of energy of gamma ($\sum E_\gamma^{\rm CM}$) should be larger than 3 GeV or the maximal transverse momentum ($P_{\rm t\_max}$ in laboratory frame) of charged tracks to be larger than 1 GeV/c.

- For events having track number from 2 to 4, two more conditions need to be satisfied. First, the sum of momentum moduli of all charged tracks ($\sum |P^{\rm CM}|$) plus the sum of energy of gamma ($\sum E_\gamma^{\rm CM}$) plus the module of the missing momentum ($P_{\rm miss}^{\rm CM} = P_{\rm beams}^{\rm CM} - P_{\rm obs}^{\rm CM}$, where $P_{\rm obs}^{\rm CM}$ includes both charged tracks and neutral clusters) needs to be less than 9 GeV/c or the maximal opening angle between tracks needs to be less than 175°. Second, the number of tracks in the barrel region ($30° < \theta < 130°$) needs to be equal or more than 2 or the energy deposited in ECL by charged tracks needs to be less than 5.3 GeV.

In addition to the standard pre-selection, extra terms are set especially for decay modes under consideration during this pre-selection.

- The number of events is required to be exactly equal to 4 with 0 net charge.

- The invariant mass of the missing momentum needs to be larger than 1 GeV/c^2 and less than 7 GeV/c^2.

- The polar angle of the missing momentum needs to be larger than 30° and less than 150°.

- The primary vertex needs to be close to the IP within 2.5 cm in longitudinal direction.

In the $\tau\tau$ pair selection, the missing momentum serves as the most effective cut. A 2D plot of the missing momentum is shown in Figure 10, where x axis is the polar angle of missing momentum and y axis is the invariant mass of missing momentum. Distinguishable difference between $\tau\tau$ and other processes are demonstrated.

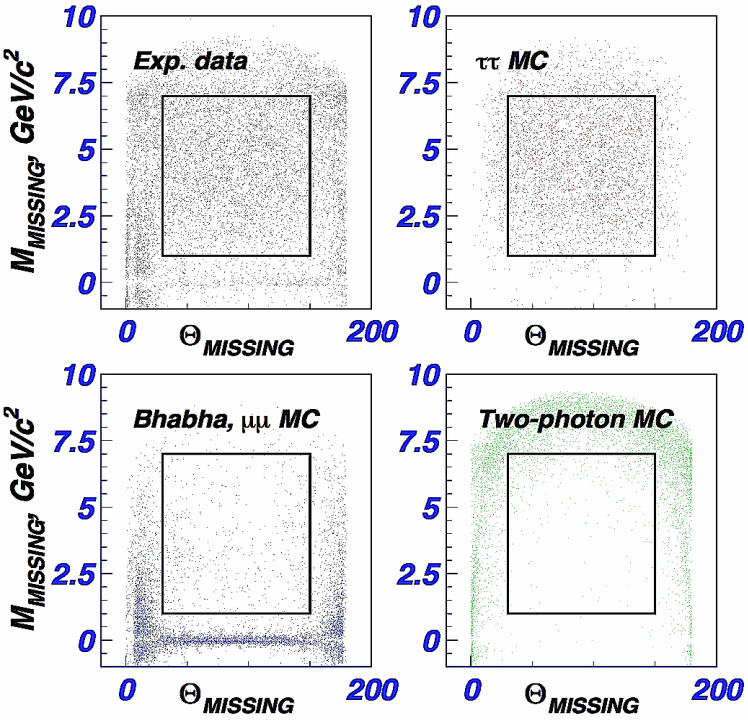


Figure 10: A 2D plot of the invariant mass of the missing momentum verses the polar angle of the missing momentum of events after the standard pre-selection. Non-$\tau\tau$ events are rejected effectively [40].

4.2 Selections for $\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}$

In the analysis of $\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}$, charged tracks with transverse momentum larger than 0.1 GeV/c and within the acceptance of the detector are selected. The closest approach of each track to the beam axis should be less than 1 cm in the $x-y$ plane while the distance from a track to the interaction point along the $z-$axis should be less than 5 cm. Four such charged tracks are required exactly in one event, among which, two opposite charged tracks with electron identification (eID) likelihood higher than 0.5 should be able to reconstruct a converted gamma vertex (expressed as $e^{+}e^{-}$ vertex in the following). In case of multiple $e^{+}e^{-}$ vertex candidates, the one with the least chi-square is chosen. Then a tau candidate is reconstructed by the $e^{+}e^{-}$ vertex candidate and a pion candidate whose pion identification (πID) likelihood (in π/K separation) is higher than 0.4 and its charge is the same as the charge of the $\tau-$lepton considered. The momentum of pion candidate in both CMS and LAB frame are required to be larger than 0.2 GeV/c and its muon identification likelihood (μID) less than 0.95 and electron identification likelihood (eID) less than 0.2. The fourth charged track is regarded as the daughter of the accompanying τ. Photons are selected with energy threshold 50 MeV in barrel ($32^{\circ} < \theta < 130^{\circ}$) and 100 MeV in end cap. To take into account bremsstrahlung effect, photon candidates that associate $e^{\pm}$ tracks with a cross angle < 0.05 radian and energy lower than that of $e^{\pm}$ in CMS frame are combined with the corresponding $e^{\pm}$ candidates. Meanwhile, the invariant mass of new $e^{+}e^{-}$ should be less than 0.05 GeV/c^2.

A $3-1$ topology is required for a signal event candidate, by a thrust perpendicularly dividing the space into two hemispheres as shown in Figure 11. The thrust is defined by two steps. First, we search for a unit vector that maximising the projection of 4 tracks' momenta upon it, $\overrightarrow{V_{\rm Thrust}}\cdot\overrightarrow{P_{e^+}}+\overrightarrow{V_{\rm Thrust}}\cdot\overrightarrow{P_{e^-}}+\overrightarrow{V_{\rm Thrust}}\cdot\overrightarrow{P_{\pi}}-\overrightarrow{V_{\rm Thrust}}\cdot\overrightarrow{P_{\rm tag}}$. The hemisphere containing 3 tracks is regarded as the signal hemisphere and the other one containing 1 track is regarded as tag hemisphere. In both hemispheres, the genuine τ should be located within a cone [41] around the sum momenta of all particles observed in each hemisphere. Therefore, the second step is reflecting the momentum vector of tag cone with respect to the origin and summing two momenta vectors (tag and signal cones) with a weight of their corresponding cone's opening angle (the corrected vector is referred to as $\overrightarrow{n_T}$).

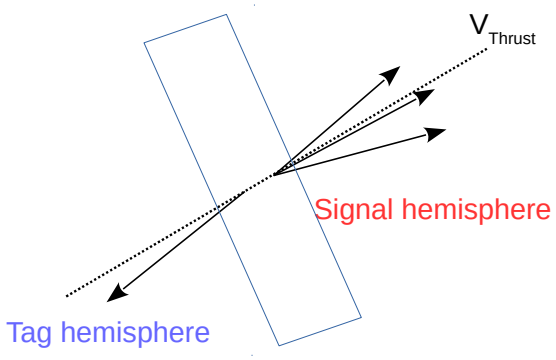


Figure 11: The schematic demonstration of the thrust. The thrust divides the space into two hemispheres. The hemisphere contains 3 tracks is regarded as the signal hemisphere. The other one is regarded as tag hemisphere.

In order to suppress external gamma conversion, cuts on the position of e^+e^- vertex candidate are employed, transverse position $|R_{xy}|<1.2$ cm, and longitudinal position $-1<Z<1.5$ cm. To suppress hadronic background, the event is required to have a back to back topology via thrust magnitude, $\frac{\sum_i|\overrightarrow{p_i}\cdot\overrightarrow{n_T}|}{\sum_i|\overrightarrow{p_i}|}$, in the range of [0.85, 0.99] and having an angle between signal τ and 3 prongs located in physical region, i.e. $|\cos(\tau-3{\rm prongs})|\equiv|\frac{2E_{\rm beam}E_{\rm 3prongs}-m_\tau^2-M_{\rm 3prongs}^2}{2|P_\tau|\cdot|P_{\rm 3prongs}|}|\leq 1$ (the cosine of angle between the momentum of signal τ and its daughter 3 prongs system in the rest frame of $\Upsilon(4S)$ under the assumption that the massless neutrino is not reconstructed). To suppress $\tau^-\to\pi^-\pi^0\nu_\tau$, γ veto and π^0 veto are employed. In the signal hemisphere, at most one gamma (excluding those have been combined with $e^\pm$ tracks) is allowed with energy less than 300 MeV. π^0 is reconstructed by two gamma ($110<M_{\gamma\gamma}<160$ MeV/c^2) or one gamma and a pair of $e^\pm$ ($110<M_{\rm e^+e^-\gamma}<165$ MeV/c^2). Also, to further suppress the $\rho^-(\pi^-\pi^0)$ resonance, the invariant mass of 3 prongs ($M_{\rm 3prongs}$) is required to be larger than 1.05 GeV and less than 1.8 GeV.

All selection cuts are listed in Table 5. The last cut is used to define the sideband, details of which will be introduced in next chapter.

Table 5: Selection Criteria of the $\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}$

Descriptions	Relevant cuts
Energy-Momentum	$P_t^{\rm LAB} > 0.1$ GeV/c, $P_{\pi}^{\rm CMS/LAB} > 0.2$ GeV/c
PID	eID($e^{\pm}$)>0.5, μID(π)<0.95, eID(π)<0.2, πID(π) > 0.4
Secondary $e^{+}e^{-}$ vertex	$\lvert R_{xy}\rvert < 1.2$ cm, $-1 < Z < 1.5$ cm
Angular distribution	$\lvert\cos(\tau - 3\text{prongs})\rvert \leq 1$
Thrust magnitude	$0.85 < \frac{\sum_i \lvert\vec{p_i}\cdot\vec{n_T}\rvert}{\sum_i \lvert\vec{p_i}\rvert} < 0.99$ (in CMS)
Gamma	$N_{\rm signal\_hemisphere\_\gamma} \leq 1$, $E_{\rm signal\_hemisphere\_\gamma\_max} \leq 300$ MeV
π^0 veto	$110 < M_{\gamma\gamma} < 160$ MeV/c^2, $110 < M_{\rm e^+e^-\gamma} < 165$ MeV/c^2
Invariant mass	$1.05 < M_{\rm 3prongs} < 1.8$ GeV/c^2

4.2.1 π^0 reconstruction correction using a reference mode $\tau^- \to \pi^- e^+ e^- \gamma \nu_\tau$

In the sideband study, it is observed that there exists discrepancy on the π^0 reconstruction between MC and EXP data (the π^0 candidate is reconstructed by $e^+e^-\gamma$, a candidate with invariant mass $110 < M_{e^+e^-\gamma} < 165$ MeV/c^2 is regarded as a genuine π^0). In order to study this discrepancy of π^0 veto, a reference mode $\tau^- \to \pi^-\nu_\tau e^+e^-\gamma$ is studied to check the $\pi^0(e^-e^+\gamma)$ reconstruction efficiency and we introduce the correction on π^0 veto in MC data.

The selection on the reference mode is the same as that applied upon the signal mode, except two points. First, here existence of $\pi^0(e^+e^-\gamma)$ is required, rather than vetoed in the signal mode study. Second, the cut on $M(\pi\ e^+e^-)$ $(= M_{\rm 3prongs})$ is not applied. There are several considerations for the second points (removing the cut on $M(\pi\ e^+e^-)$). It is the range $[1.05, 1.8]$ GeV/c^2 of M($\pi\ e^+e^-$) that is of concern, instead of the sideband range $[0, 1]$ GeV/c^2. At the same time removing this cut could increase the statistics so as to decrease the uncertainty.

Figure 12 and Figure 13 show the distributions of $M(e^+e^-\gamma)$ and $M(\pi\ e^+e^-)(= M_{\rm 3prongs})$, respectively, with the selection on the reference mode. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data. Discrepancies exist between the EXP distribution and MC distribution of the $M(e^+e^-\gamma)$ on both the number of events and the shape. In the case of $M(\pi\ e^+e^-)$, discrepancy exists on the number of events only. Then by removing the π^0 requirement, which correspond to the selection of the signal mode without π^0 veto, the EXP data and MC data agree quite well both on the number of events and the shape, as shown in Figure 14. As a result, the correction of π^0 veto on the MC data is introduced.

In total, 6732 EXP events and 7364.56 (BKG) + 6.29 (signal) MC events pass the selection of the reference mode. The number of EXP events is less than the MC events, and EXP/MC $= 0.913 \pm 0.011$, relatively uncertainty = 1.2%. Here error is calculated via Poisson variance, $\sqrt{\rm EXP}/{\rm MC}$.

Based on the events number ratio EXP/MC of the reference mode, we introduce the correction on the $\pi^0(e^-e^+\gamma)$ reconstruction efficiency in MC samples for the signal mode study. When such an event is found in EXP data, the event is vetoed completely. While in MC, in order to compensate the over-veto effect, such an event is assigned a weight 1−EXP/MC, rather than

completely vetoed. For all the events having this π^0 veto, an uncertainty of this correction is assigned.

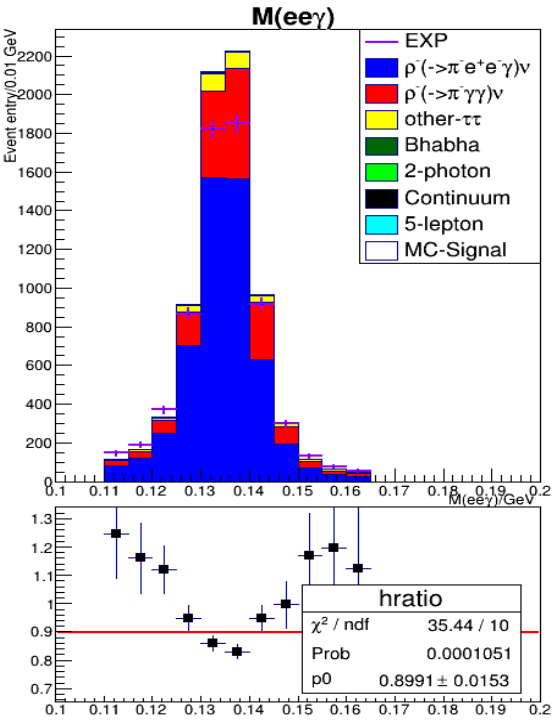


Figure 12: The distribution of invariant mass of $\pi^0(e^+e^-\gamma)$ of the reference decay mode. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

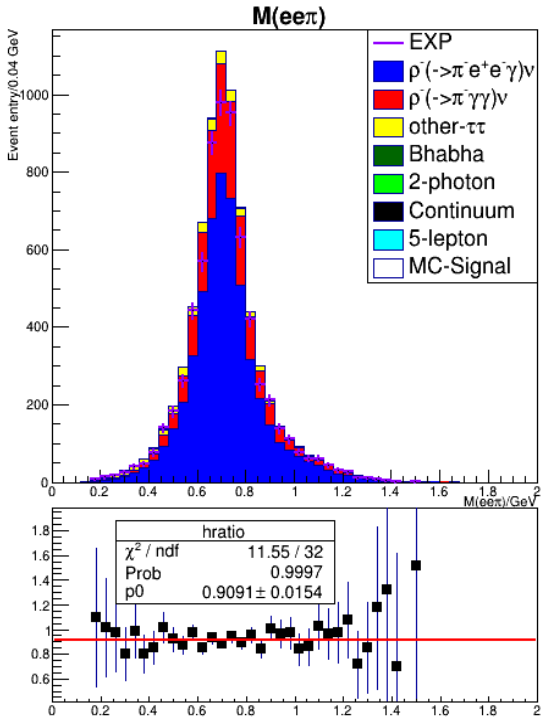


Figure 13: The distribution of invariant mass of $e^+e^-\pi^-$ of the reference decay mode $\tau^- \to \pi^-\nu_\tau e^+e^-\gamma$. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

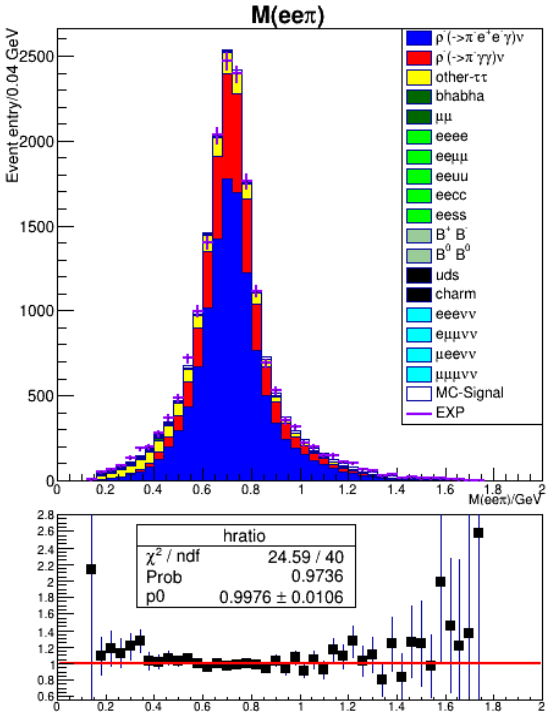
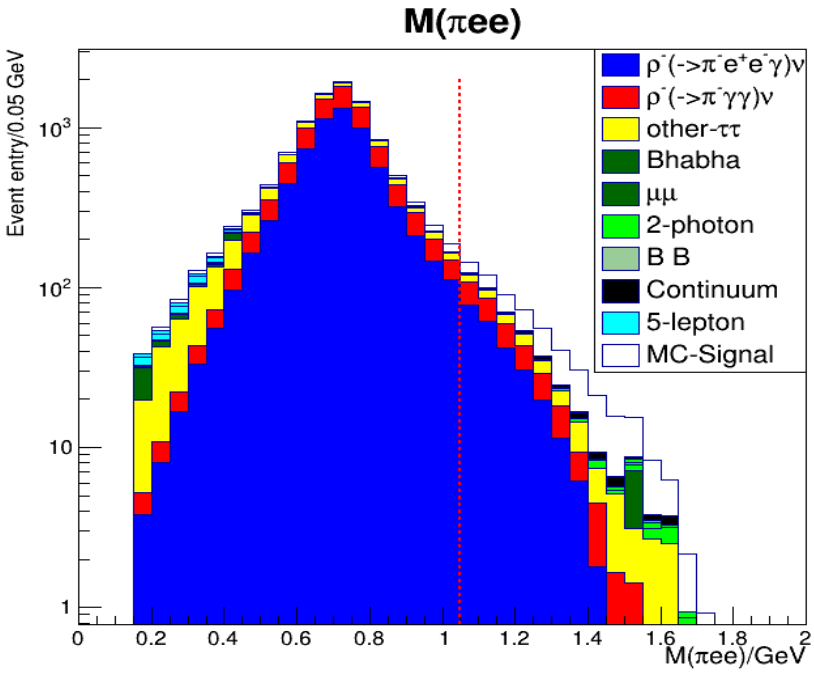


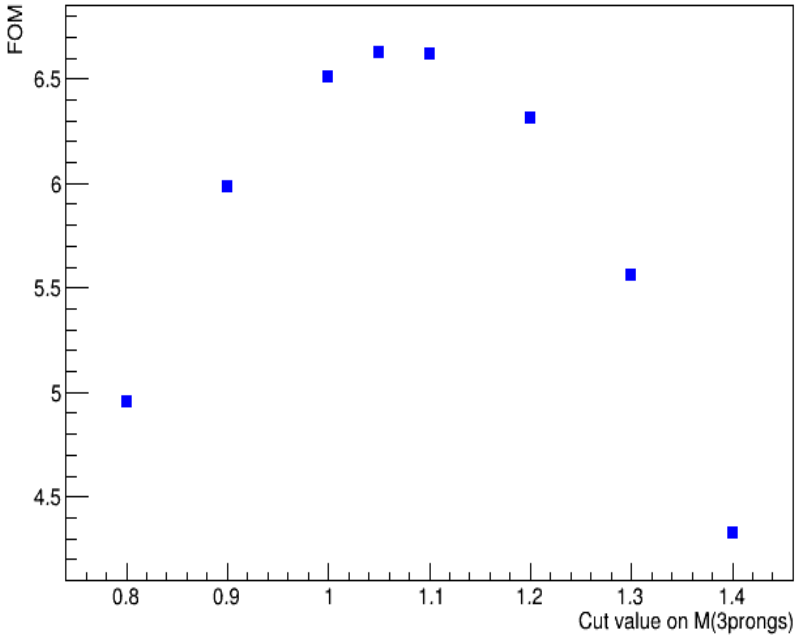
Figure 14: The distribution of invariant mass of $e^+e^-\pi^-$ of the events without π^0 veto. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

4.2.2 Distributions of cuts

After applying all corrections, distributions of selection cuts from MC data are shown below. The distribution of the invariant mass of 3 prongs are shown in Figure 15, together with its FOM curve plot. The optimisation of selections is implemented by setting cut value that corresponds to the maximal FOM. Similarly, the distribution of transverse position of the secondary e^+e^- vertex is shown in Figure 16. The distribution of longitudinal position of the secondary e^+e^- vertex is shown in Figure 17. The distribution of $\cos(\tau{-}3\text{prongs})$ is shown in Figure 18. The distribution of thrust magnitude is shown in Figure 19. The distribution of number of γ in signal hemisphere is shown in Figure 20. The distribution of energy of γ in signal hemisphere is shown in Figure 21. The distribution of invariant mass of $\pi^0(ee\gamma)$ candidates is shown in Figure 22. All figures are drawn with cut on the specific variable removed while other cuts are still functioning. Colourful bars indicate BKG MC, while the white bars indicate signal MC. Table 6 shows the number of events for signal MC and BKG MC as a function of selection cuts.



(a) Invariant mass of 3 prongs



(b) FOM corresponding to different cut value

Figure 15: The distribution of invariant mass of 3 prongs from MC is shown in (a). Cut is set at 1.05 GeV/c^2. Cut value is determined by the maximal FOM, as shown in (b). Colourful bars indicate BKG MC, while the white bars indicate signal MC.

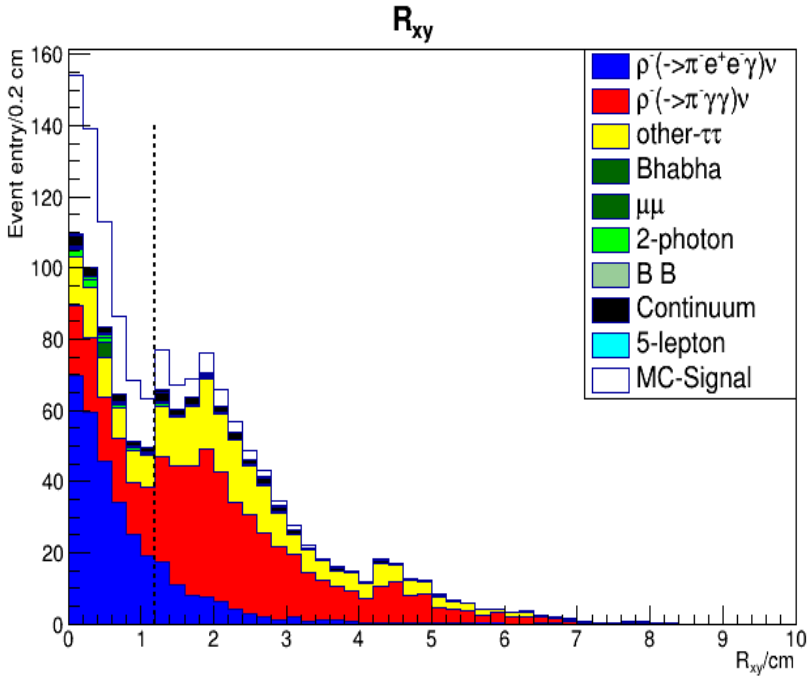


Figure 16: The distribution of transverse position of the secondary e^+e^- vertex. Cut is set at 1.2 cm. Colourful bars indicate BKG MC, while the white bars indicate signal MC.

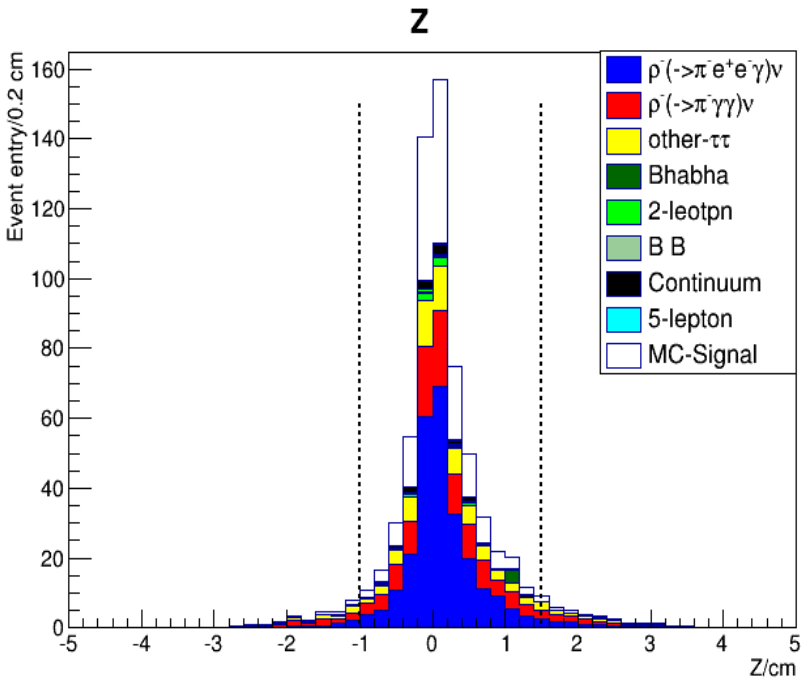


Figure 17: The distribution of longitudinal position of the secondary e^+e^- vertex. Cut is set at -1 cm and 1.5 cm. Colourful bars indicate BKG MC, while the white bars indicate signal MC.

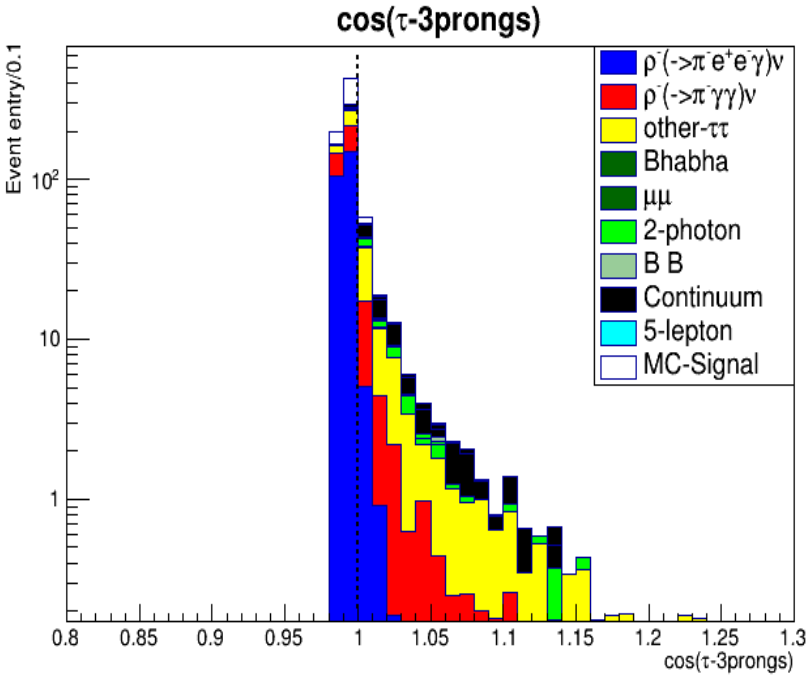


Figure 18: The distribution of cos($\tau-$3prongs). Cut is set at 1. Colourful bars indicate BKG MC, while the white bars indicate signal MC.

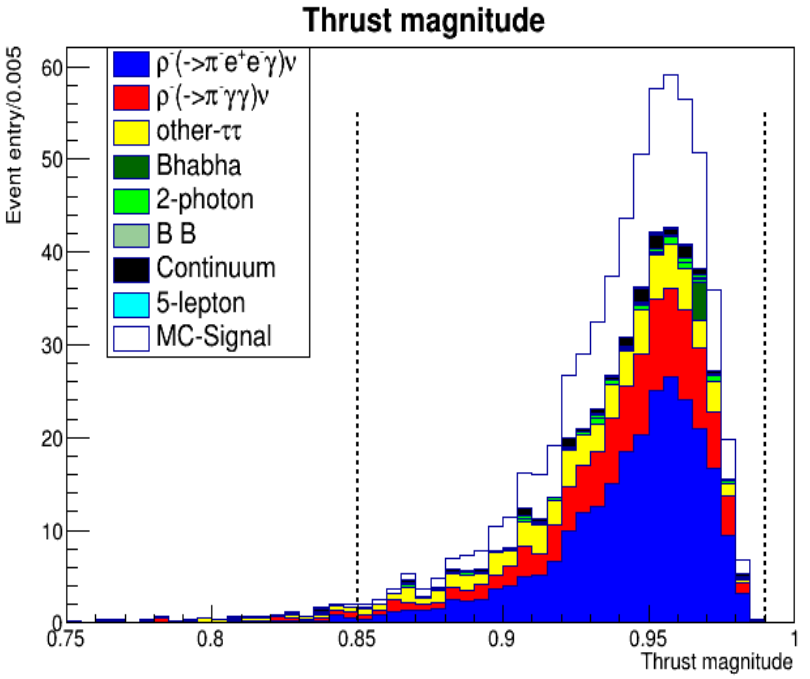


Figure 19: The distribution of thrust magnitude. Cuts are set at 0.85 and 0.99. Colourful bars indicate BKG MC, while the white bars indicate signal MC.

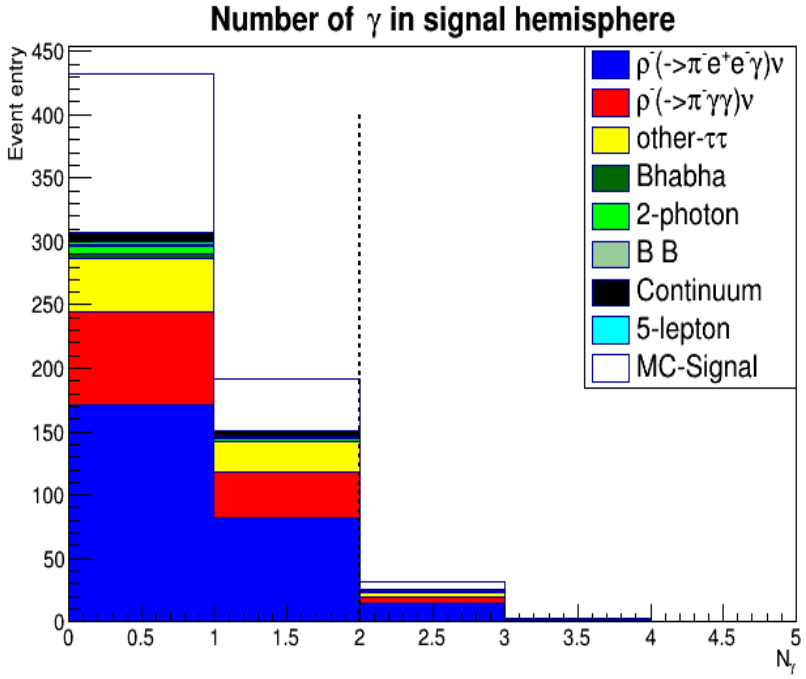


Figure 20: The distribution of number of γ in signal hemisphere. Cut is set at 1. Colourful bars indicate BKG MC, while the white bars indicate signal MC.

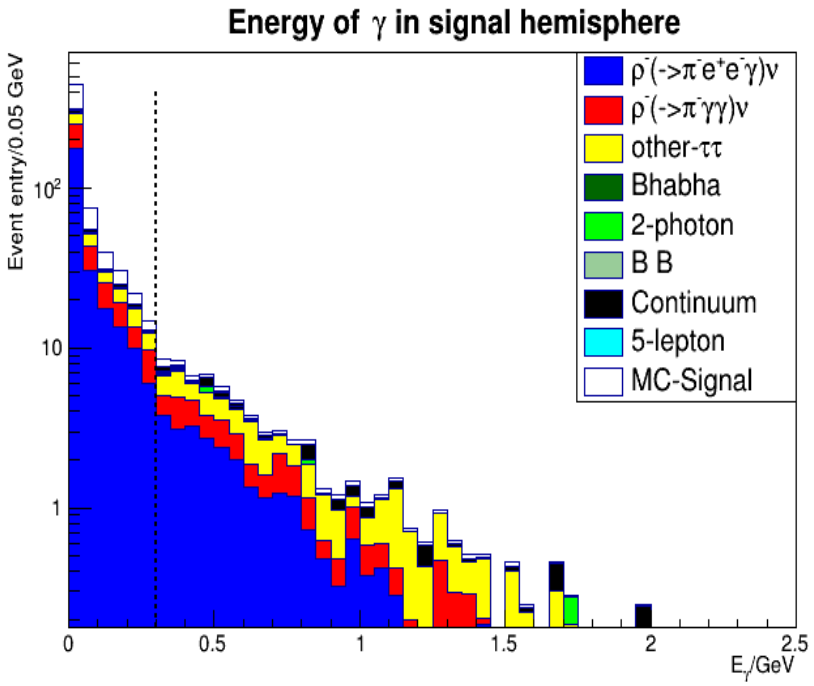
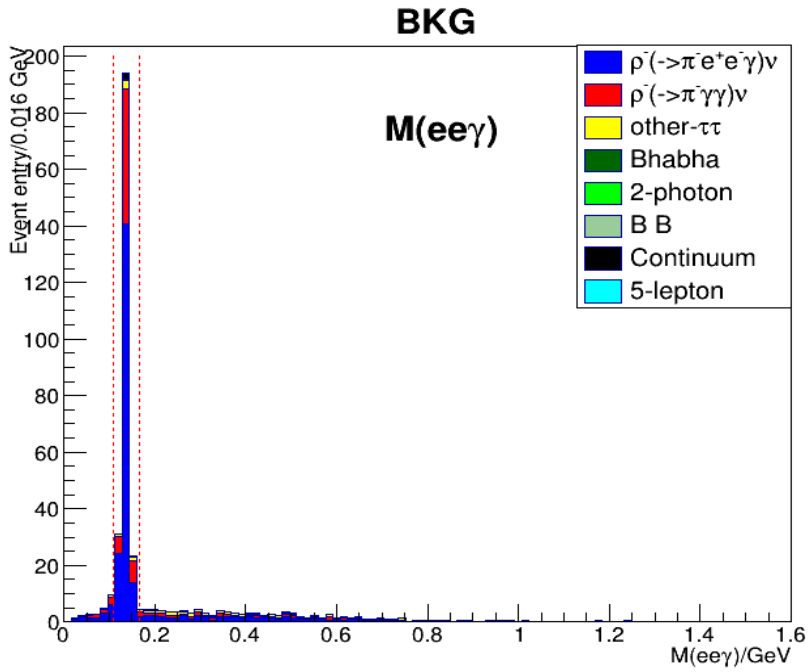
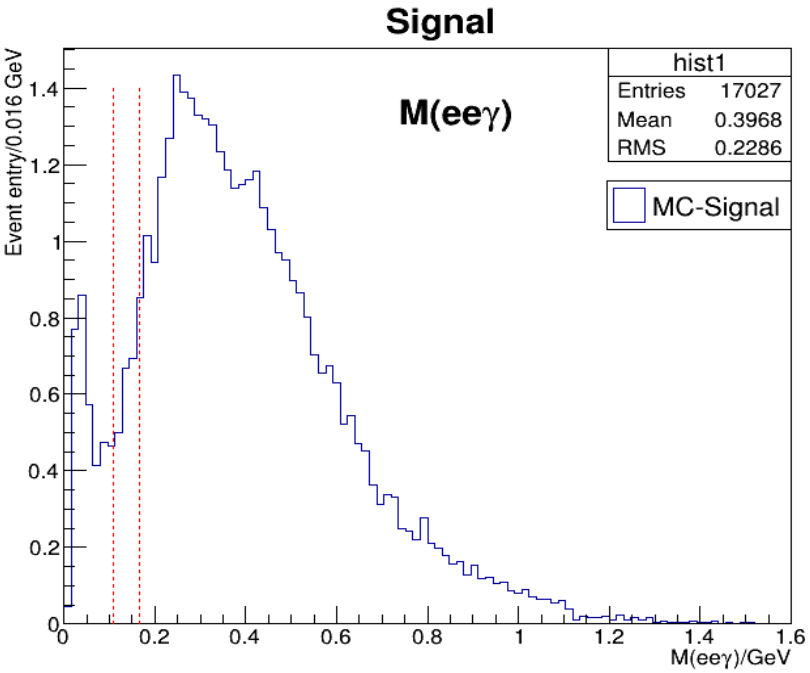


Figure 21: The distribution of energy of γ in signal hemisphere. Cut is set at 0.3 GeV. Colourful bars indicate BKG MC, while the white bars indicate signal MC.



(a) Invariant mass of π^0 reconstructed by $e^+e^-\gamma$ of BKG MC



(b) Invariant mass of π^0 reconstructed by $e^+e^-\gamma$ of signal MC

Figure 22: The mass spectrum of π^0 candidate reconstructed by $e^+e^-\gamma$ of BKG MC (a) and signal MC (b). Only candidates located between two red vertical (110 MeV/c^2 and 165 MeV/c^2) dotted lines are regarded as genuine π^0 candidates.

Table 6: Cut-flow table of number of signal MC events and BKG MC events for τ^- (one charge), normalised to 562 fb^{-1}.

After cuts	Number of signal MC events	Number of BKG MC events
Energy-Momentum	735	429548
Secondary e^+e^- vertex	479	102675
Angular distribution	469	101742
Thrust magnitude	451	98134
Gamma	397	17663
π^0 veto	388	10677
Invariant mass	165	458

4.3 Selections for $\tau^\pm \to \pi^\pm\mu^+\mu^-\nu_\tau$

In the analysis of $\tau^\pm \to \pi^\pm\mu^+\mu^-\nu_\tau$, charged tracks with transverse momentum larger than 0.1 GeV/c and within the acceptance of the detector are selected. The closest approach of each track to the beam axis should be less than 1 cm in the $x-y$ plane while the distance from a track to the interaction point along the $z-$axis should be less than 5 cm. Four such charged tracks are required exactly in one event, among which, two opposite charged tracks need to have a μID likelihood higher than 0.7 and one charged track have a πID likelihood higher than 0.4 as a pion candidate. Photons are selected with energy threshold 50 MeV in barrel ($32^\circ < \theta < 130^\circ$) and 100 MeV in end cap. Events with pion candidate having an eID > 0.8 are rejected. The tracks of μ^- and μ^+ candidates need to have $\chi^2(\mu^-) + \chi^2(\mu^+) < 175$ (χ^2 indicates track quality) and then get fitted again to seek for a vertex. In case of multiple $\mu^-\mu^+$ vertices, the vertex having the smallest angle between μ^- and μ^+ in CMS frame is chosen. The transverse radius of $\mu^-\mu^+$ vertex is required to be $|R_{xy}| < 0.15$ cm. The invariant mass of $\mu^-\mu^+$ pair is required to be less than 0.85 GeV. Then a tau candidate is reconstructed by a pion candidate and the $\mu^-\mu^+$ vertex candidate. The fourth charged track is regarded as the daughter of the accompanying τ. The $3-1$ topology in the analysis of $\tau^\pm \to \pi^\pm e^+e^-\nu_\tau$ is also required. Also, the thrust is defined by the aforementioned procedure of $\tau^\pm \to \pi^\pm e^+e^-\nu_\tau$.

In order to suppress $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ mode, stringent μID likelihood are imposed upon μ^- and μ^+ candidates, $\mu\text{ID}(\mu^\pm) > 0.97$. Meanwhile, it is required that $P_t(\mu^\pm) > 720$ MeV/c to ensure $\mu^\pm$ candidates hit the KLM detector. To suppress $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$ mode, γ veto is employed. In the signal hemisphere, the sum of the energy of all gammas is required to be less than 300 MeV. Moreover, the number of gammas in both hemispheres need to be no more than 5. To suppress $\tau^- \to K^-\pi^+\pi^-\nu_\tau$ mode, $\mu^\pm$ candidates are required to have Kaon ID likelihood less than 0.8. To suppress hadronic background, the event is required to have a back-to-back topology via thrust magnitude $(\frac{\sum_i |\vec{p_i}\cdot\vec{n_T}|}{\sum_i |\vec{p_i}|}) > 0.9$ in CMS frame. Meanwhile, a pseudo mass of 3 prongs is required to be less than 1.8 GeV/c^2, and the sum of the energy of 3 prongs (in CMS) less than beam energy. The pseudo mass of 3 prongs is defined as:

$m^* = \sqrt{2 \cdot E_{3\text{prongs}}(E_\tau - E_{3\text{prongs}}) + m^2_{3\text{prongs}} - 2 \cdot |p_{3\text{prongs}}|(E_\tau - E_{3\text{prongs}})}$, where E_τ is taken as half of the beam energy.

All selection cuts are listed in Table 7. The last cut is used to define the sideband, details of

which will be introduced in next chapter.

Table 7: Selection Criteria of the $\tau^{\pm} \to \pi^{\pm}\mu^{+}\mu^{-}\nu_{\tau}$

Descriptions	Relevant cuts
Energy-Momentum	$P_t^{\rm LAB} > 0.1$ GeV/c, $P_t(\mu^{\pm}) > 720$ MeV/c, $E_{\pi} + E_{\mu^+} + E_{\mu^-} < E_{beam}$
PID	πID(π)>0.4, eID(π)<0.8, μID($\mu^{\pm}$)>0.97, KID($\mu^{\pm}$)<0.8
Mass	$m^*(\pi\mu^+\mu^-) < 1.8$GeV/c^2, M($\mu^+\mu^-$) < 0.85GeV/c^2
Thrust magnitude	$\frac{\sum_i \lvert\vec{p_i}\cdot\vec{n_T}\rvert}{\sum_i \lvert\vec{p_i}\rvert} > 0.9$ (in CMS)
Gamma	$E_{\rm total\_signalside\_\gamma} < 300$ MeV, $N_{\gamma\_\rm bothsides} < 6$
$\mu^-\mu^+$ vertex	$\lvert R_{xy}\rvert < 0.15$ cm

4.3.1 $\pi \to \mu$ mis-identification correction using a reference mode $\tau^{\pm} \to \pi^{-}\pi^{+}\pi^{\pm}\nu_{\tau}$

In the analysis of $\tau^{\pm} \to \pi^{\pm}\nu_{\tau}\mu^{+}\mu^{-}$, a stringent cut on μID likelihood is applied. At Belle, the conventional studies concerning μID and mis-identification efficiency of $\pi \to \mu$ (or called π fake rate) are carried out through hadronic processes. However, in this analysis, the main contribution of background events arises from τ decays. The environments of these two are different especially in multiplicity. Also, pion fake rate is only studied at μID = 0.1, 0.9 and 0.95 [42]. In order to investigate the fake rate of π being mis-identified as μ under the environment of τ decays at a high μID cut (0.97 in the analysis for $\tau^{\pm} \to \pi^{\pm}\mu^{+}\mu^{-}\nu_{\tau}$), a reference decay mode $\tau^{\pm} \to \pi^{\pm}\pi^{+}\pi^{-}\nu_{\tau}$ is exploited. To study $\pi^{-} \to \mu^{-}$ mis-identification, $\tau^{+} \to \pi^{+}\pi^{+}\pi^{-}\bar{\nu}_{\tau}$ is employed, where μID cut can be imposed upon the unambiguous π^{-} candidate. A reverse procedure is carried out for the charge conjugated case ($\pi^{+} \to \mu^{+}$), i.e., μID cut is imposed on π^{+} in the $\tau^{-} \to \pi^{-}\pi^{+}\pi^{-}\nu_{\tau}$. Selections are set according to that of $\tau^{\pm} \to \pi^{\pm}\mu^{+}\mu^{-}\nu_{\tau}$. All selection cuts are listed in Table 8. One more thing worth of mention is that the two same sign pions are required to have μID(π) < 0.97, which ensures that all the data used here will not be able to enter the signal box, since signal box requires two charge particles with μID(π) > 0.97.

The μID likelihood is a function with strong dependence on momentum and polar angle. Using the data collected by SVD2 at the energy of $\Upsilon(4S)$ resonance, distributions of momentum and polar angle are drawn. Distributions of the polar angle of $\pi^{\pm}$ in the reference mode are shown in Figure 23. Distributions of the polar angle of $\mu^{\pm}$ candidates of BKG MC in the signal box are shown in Figure 24. Distributions of the momentum of $\pi^{\pm}$ in the laboratory frame in the reference mode are shown in Figure 25. Distributions of the momentum of $\mu^{\pm}$ candidates of BKG MC in laboratory frame in the signal box are shown in Figure 26. The distributions of $\pi^{\pm}$ in reference mode are quite similar to that of the $\mu^{\pm}$ candidates in the signal box. Therefore, we only tabulate the correction factor based on momenta and polar angles into nine blocks, $(0^{\circ}-45^{\circ}, 45^{\circ}-125^{\circ}$ and $125^{\circ}-180^{\circ})\otimes(0-1.1$ GeV, $1.1-1.6$ GeV and > 1.6 GeV). $45^{\circ}-125^{\circ}$ is the coverage of barrel KLM.

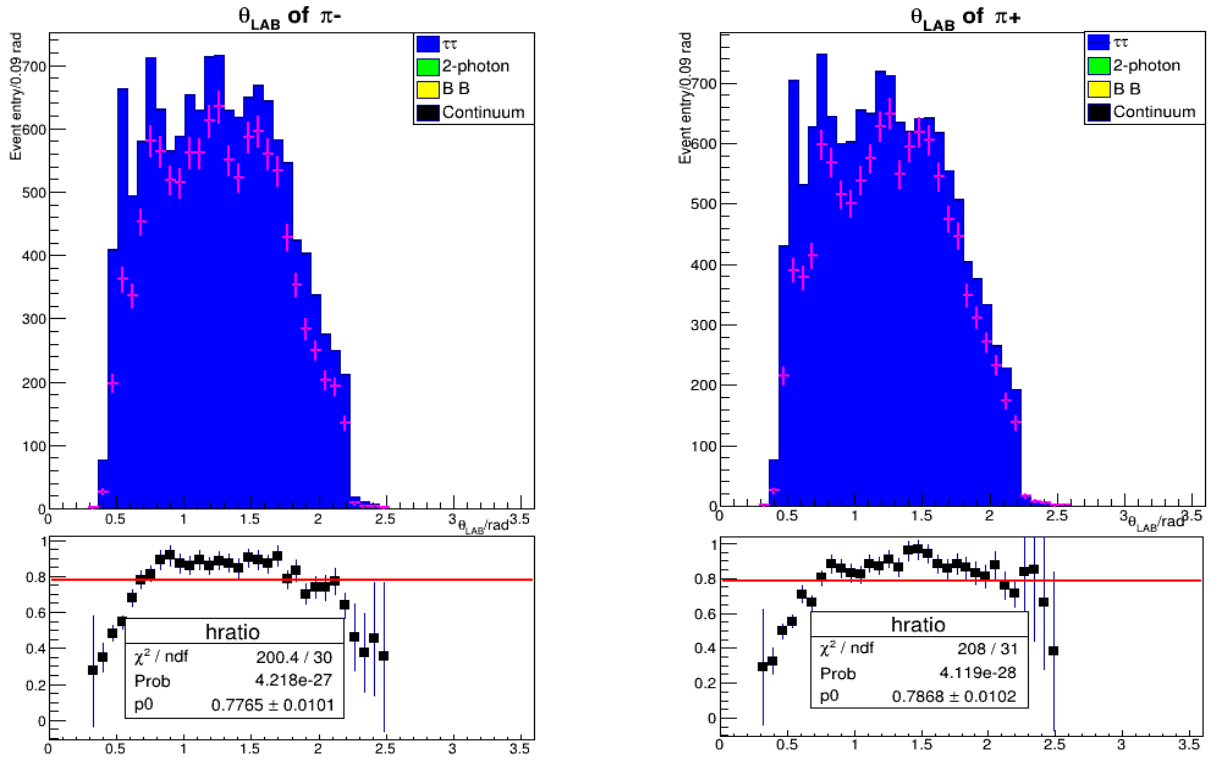
The mis-id efficiency is defined as: $\epsilon\ (\pi^{-} \to \mu^{-}) = \frac{N_1}{N_0}$, where N_0 is the number of event surviving the selections mentioned above and N_1 is the number of event surviving the same selections and additionally μID(π') > 0.97 (π' is the candidate of concern). The ratio of mis-id efficiency between EXP/MC is obtained, as shown in Table 9 for the case of π^{-} and Table 10

for the case of π^+.

A correction factor $R_{\pi\to\mu}(P_{\rm LAB},\theta_{\rm LAB})$ is applied to the generic $\tau\tau$ MC sample event by event. In other words, every MC event is weighted by the correction factor $R_{\pi\to\mu}(P_{\rm LAB},\theta_{\rm LAB}) = \frac{\epsilon_{\rm EXP}(\pi^-\to\mu^-)}{\epsilon_{\rm MC}(\pi^-\to\mu^-)} \cdot \frac{\epsilon_{\rm EXP}(\pi^+\to\mu^+)}{\epsilon_{\rm MC}(\pi^+\to\mu^+)}$ depending on the momentum and polar angle of the $\mu^\pm$ candidates.

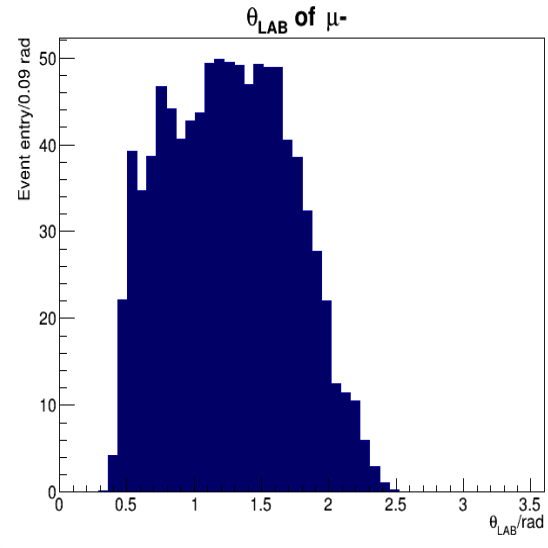
Table 8: Selection criteria of the reference mode, π' is the candidate of concern (i.e., the π^- in $\tau^+ \to \pi^+\pi^+\pi^-\bar{\nu}_\tau$ and the π^+ in $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$), π are the accompanying pions with same charge sign.

Descriptions	Relevant cuts
Energy-Momentum	$P_t^{\rm LAB} > 0.1$ GeV/c, $P_t(\pi') > 720$ MeV/c, $E_{\pi'} + E_\pi + E_\pi < E_{\rm beam}$
PID	πID(π')>0.9, πID(π)>0.9, eID(π)<0.8, μID(π)<0.97
Mass	m$^*(\pi'\pi\pi) < 1.8$ GeV/c^2, M($\mu^+\mu^-$) <0.85 GeV/c^2
Track quality	$\chi^2(\pi) + \chi^2(\pi') < 175$
Event shape	$\frac{\sum_i \lvert\vec{p_i}\cdot\vec{n_T}\rvert}{\sum_i \lvert\vec{p_i}\rvert} > 0.9$ (in CMS)
Gamma Veto	$E_{\rm total\_signalside\_\gamma} < 300$ MeV, $N_{\gamma\_\rm bothsides} < 6$

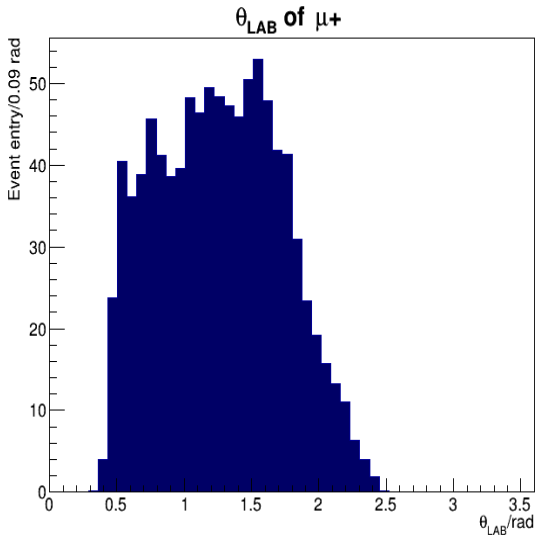


(a) π^- (b) π^+

Figure 23: The distribution of polar angles of $\pi^\pm$ in reference mode.

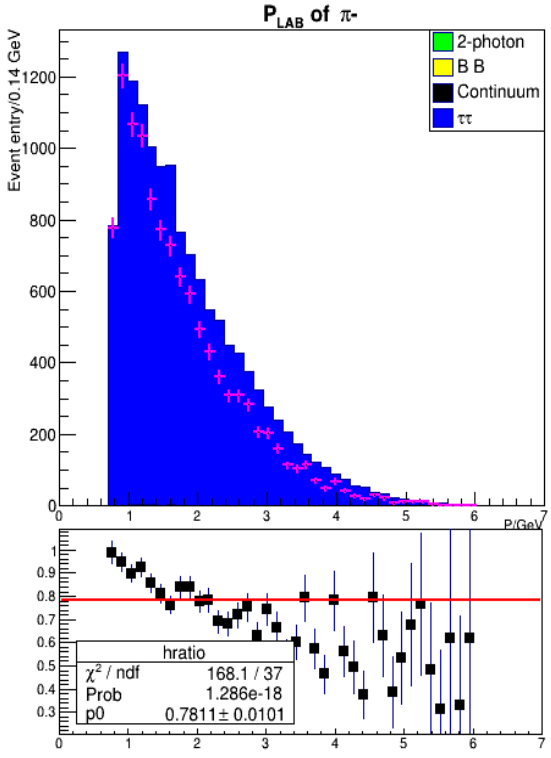


(a) μ^-

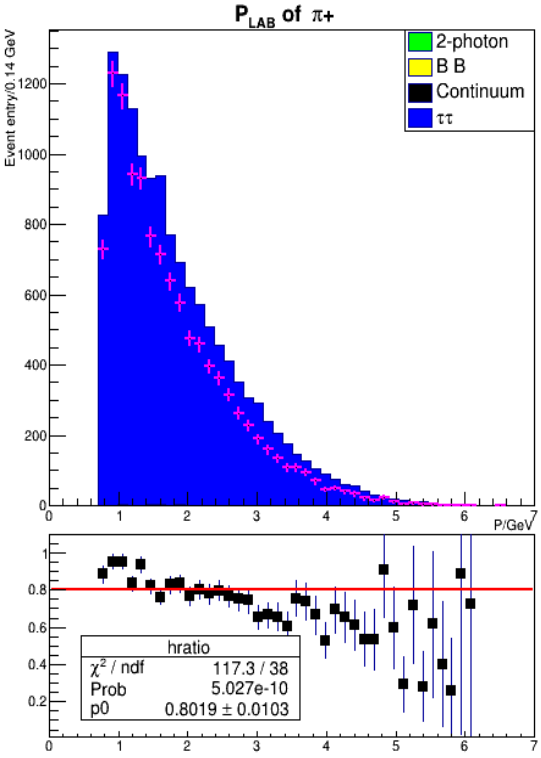


(b) μ^+

Figure 24: The distribution of polar angles of $\mu^{\pm}$ candidates of BKG MC in signal box.

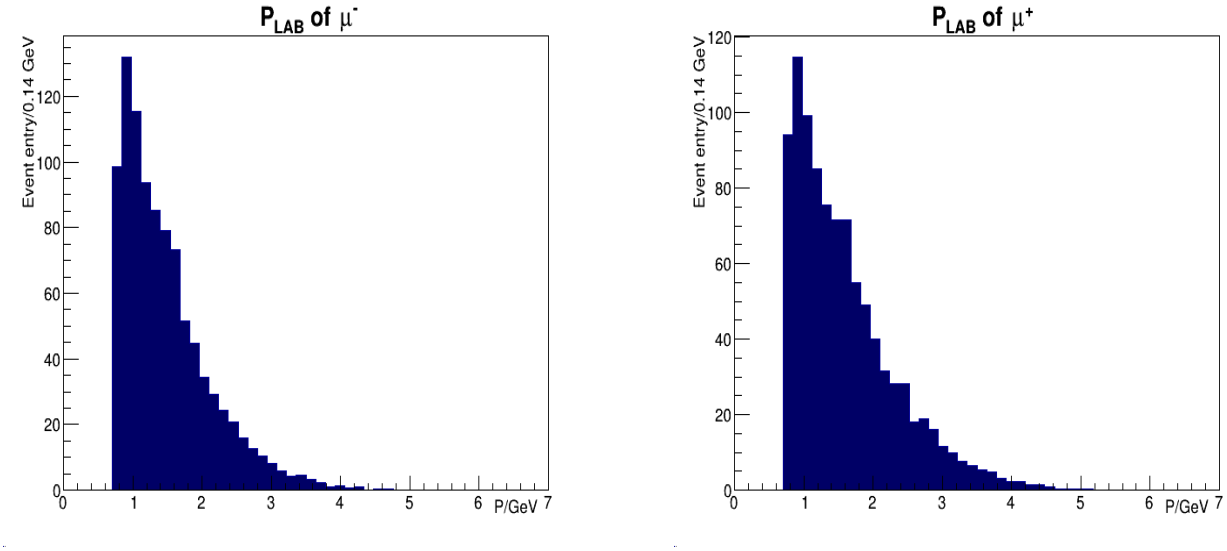


(a) π^-



(b) π^+

Figure 25: The distribution of laboratory momenta of $\pi^{\pm}$ in reference mode.



(a) μ^- (b) μ^+

Figure 26: The distribution of laboratory momenta of $\mu^\pm$ candidates of BKG MC in signal box.

Table 9: Result for $\frac{\epsilon_{\rm EXP}(\pi^-\to\mu^-)}{\epsilon_{\rm MC}(\pi^-\to\mu^-)}$ obtained from $\tau^+ \to \pi^+\pi^+\pi^-\bar{\nu}_\tau$.

$P_{\rm LAB} \setminus \theta_{\rm LAB}$	$\theta_{\rm LAB} < 45°$	$45° - 125°$	$\theta_{\rm LAB} > 125°$
$P_{\rm LAB} < 1.1$ GeV/c	1.23 ± 0.24	1.034 ± 0.019	0.80 ± 0.18
$1.1 - 1.6$ GeV/c	0.82 ± 0.04	0.990 ± 0.019	0.65 ± 0.12
$P_{\rm LAB} > 1.6$ GeV/c	0.71 ± 0.02	0.953 ± 0.016	0.74 ± 0.11

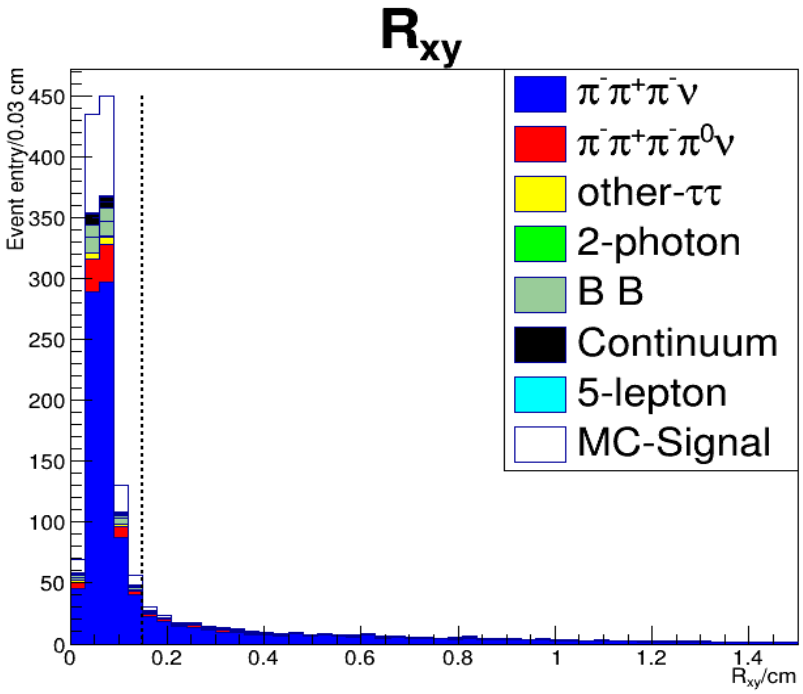
Table 10: Result for $\frac{\epsilon_{\rm EXP}(\pi^+\to\mu^+)}{\epsilon_{\rm MC}(\pi^+\to\mu^+)}$ obtained from $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$.

$P_{\rm LAB} \setminus \theta_{\rm LAB}$	$\theta_{\rm LAB} < 45°$	$45° - 125°$	$\theta_{\rm LAB} > 125°$
$P_{\rm LAB} < 1.1$ GeV/c	0.78 ± 0.18	1.059 ± 0.020	0.98 ± 0.22
$1.1 - 1.6$ GeV/c	0.76 ± 0.04	1.002 ± 0.019	0.97 ± 0.15
$P_{\rm LAB} > 1.6$ GeV/c	0.70 ± 0.02	0.961 ± 0.016	0.88 ± 0.12

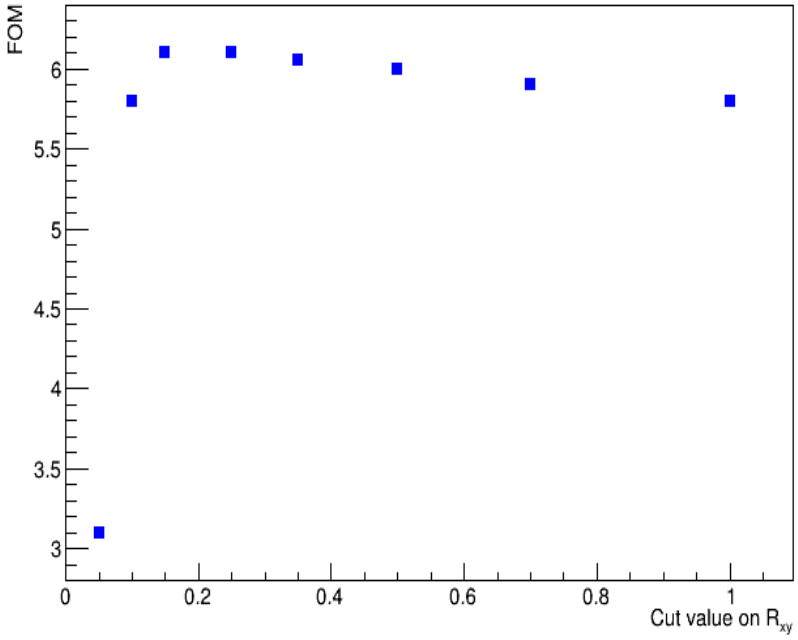
As π being mis-identified as μ is the largest contribution of BKG in the study of $\tau^\pm \to \pi^\pm\mu^+\mu^-\nu_\tau$, a cross-check is carried out to confirm the validity of this correction. The correction factor is applied to the generic $\tau\tau$ MC as an average value regardless of the division of momentum and polar angle. (The π^- fake rate is 1.307% in EXP, and 1.396% in MC. The π^+ fake rate is 1.303% in EXP, and 1.392% in MC.) Results obtained by two methods are comparable and within the uncertainty of this correction (1.5%).

4.3.2 Distributions of cuts

After applying all corrections, distributions of selection cuts from MC data are shown below. The distribution of the transverse position of the $\mu^+\mu^-$ vertex is shown in Figure 27, similarly, the cut value is optimised by FOM. The distribution of the pseudo mass of 3 prongs are shown in Figure 28. The distribution of the invariant mass of the $\mu^+\mu^-$ is shown in Figure 29. The distribution of the thrust magnitude is shown in Figure 30. The distribution of the number of photon is shown in Figure 31. The distribution of the energy of photon in signal hemisphere is shown in Figure 32. All figures are drawn with the specific cut on that variable removed while other cuts are still functioning. Table 11 shows the number of events for signal MC and BKG MC as a function of selection cuts.



(a) The transverse position of the $\mu^+\mu^-$ vertex



(b) FOM corresponding to different cut value

Figure 27: The distribution of transverse position of the $\mu^+\mu^-$ vertex from MC is shown in (a). Cut is set at 0.15 cm. Cut value is determined by the maximal FOM, as shown in (b). Colourful bars indicate BKG MC, while the white bars indicate signal MC.

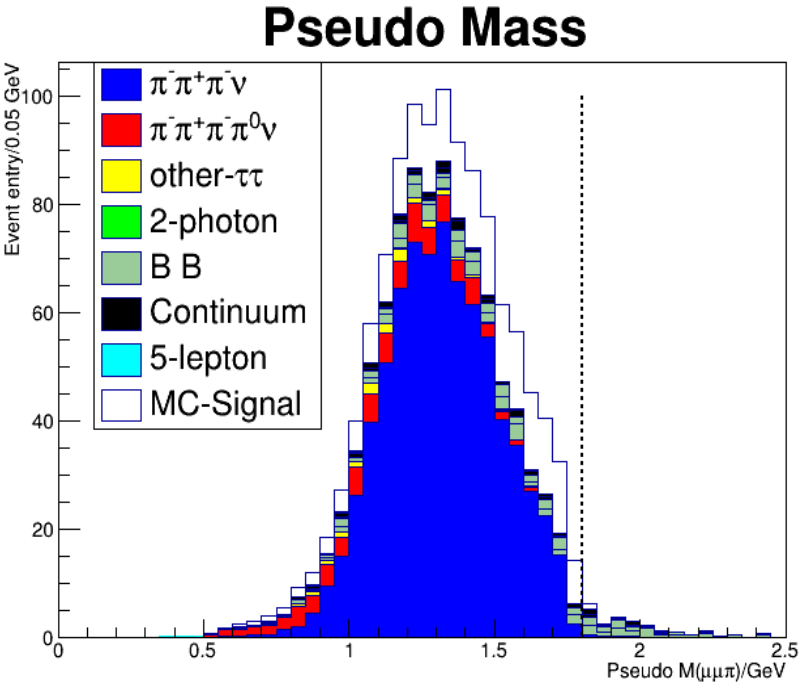


Figure 28: The distribution of pseudo mass of 3 prongs. Cut is set at 1.8 GeV/c^2. Colourful bars indicate BKG MC, while the white bars indicate signal MC.

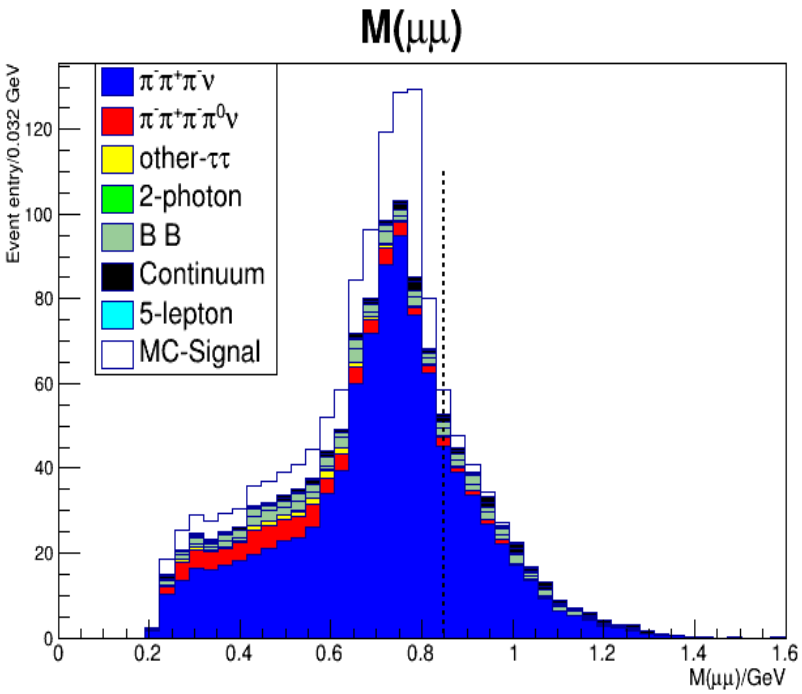


Figure 29: The distribution of invariant mass of $\mu^+\mu^-$. Cut is set at 0.85 GeV. Colourful bars indicate BKG MC, while the white bars indicate signal MC.

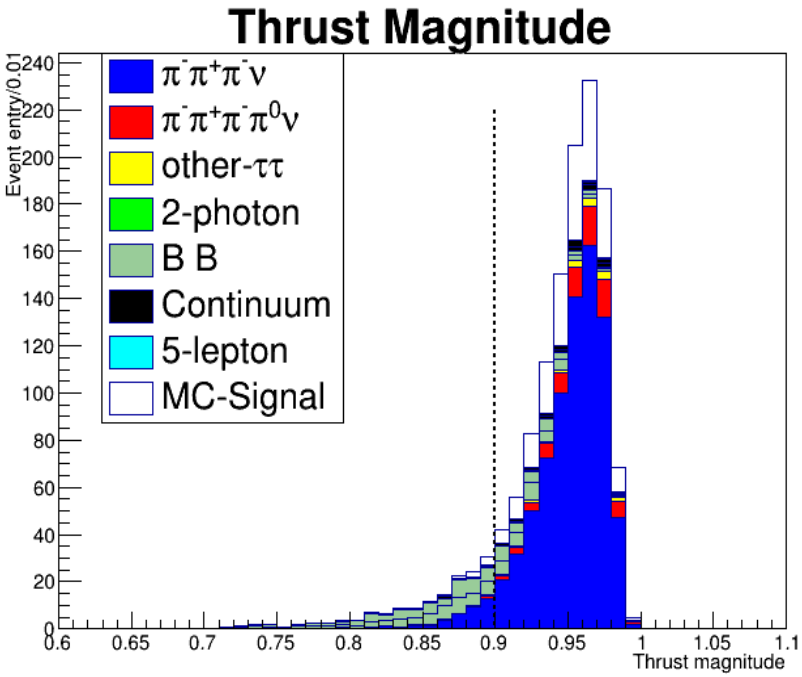


Figure 30: The distribution of thrust magnitude. Cut is set at 0.9. Colourful bars indicate BKG MC, while the white bars indicate signal MC.

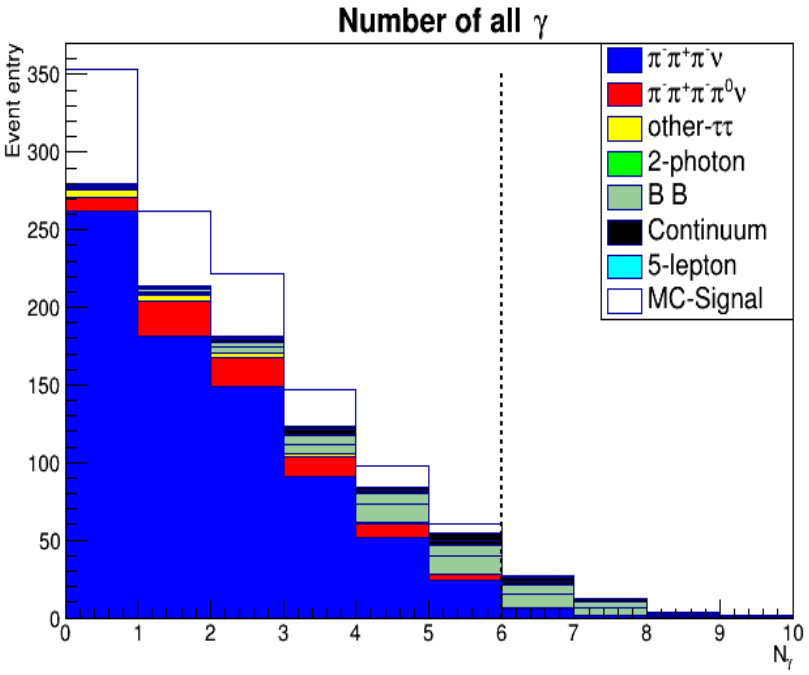


Figure 31: The distribution of number of all photons. Cut is set at 5. Colourful bars indicate BKG MC, while the white bars indicate signal MC.

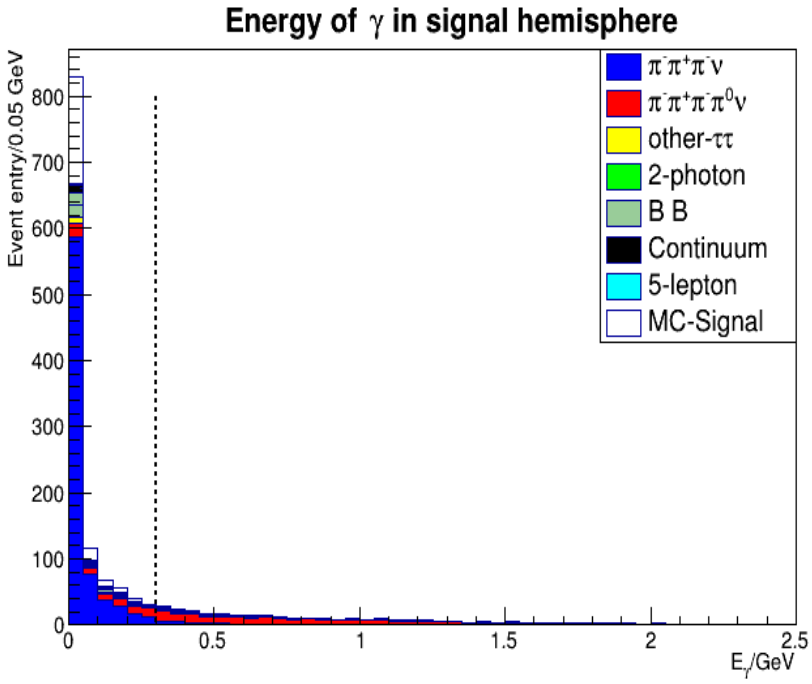


Figure 32: The distribution of energy of photon in signal hemisphere. Cut is set at 0.3 GeV. Colourful bars indicate BKG MC, while the white bars indicate signal MC.

Table 11: Cut-flow table of number of signal MC events and BKG MC events for τ^- (one charge), normalised to 562 fb^{-1}.

After cuts	Number of signal MC events	Number of BKG MC events
Energy-Momentun	340	35616
Mass	316	11347
Thrust magnitude	299	10324
Gamma	287	7333
μID	221	1294
$\mu^-\mu^+$ vertex	205	937

5 MC validation

In this study, the signal yield is obtained by observed events subtracting BKG events predicted by MC. In order to infer a reliable signal yield, it is necessary to implement sideband studies (using background controlled samples) for both decay modes to validate the background MC samples. MC samples are corrected by particle identification, tracking effects, π^0 and $\pi \to \mu$ corrections.

The sideband is defined by altering one cut in the selection explained in the last chapter. In the analysis of $\tau^- \to \pi^- e^+ e^- \nu_\tau$, the region of M($\pi^- e^+ e^-$) < 1 GeV/c^2 is chosen to define the sideband, meanwhile the rest of the selections still are kept as same (i.e., the region of M($\pi^- e^+ e^-$) in [1.05, 1.8] GeV/c^2 is regarded as the signal box). While in the case of $\tau^- \to \pi^- \mu^+ \mu^- \nu_\tau$ the transverse position of the $\mu\mu$ vertex is chosen to define the sideband. The cut used to define the sideband is chosen based on the considerations that the sideband should be dominated by the background processes and the constituents of background events are almost the same in the sideband and the signal box.

5.1 Sideband study for $\tau^- \to \pi^- e^+ e^- \nu_\tau$

5.1.1 Definition of the sideband

Among all the selection cuts on $\tau^- \to \pi^- e^+ e^- \nu_\tau$, the invariant mass of 3 prongs M($\pi^- e^+ e^-$) serves as the most powerful cut to reject background events as shown in Figure 33 and 34. Also background components located on both sides of this cut own similar scale ratio. Therefore, this variable is used to define the signal box. The range of invariant mass of 3 prongs < 1 GeV/c^2 is taken as sideband. In sideband, all selections introduced in the last chapter are applied, except that now M($\pi^- e^- e^+$) < 1 GeV/c^2 is required, rather than [1.05, 1.8] GeV/c^2. For the purpose of blinding, a narrow gap is kept between sideband and signal box. In order to validate the background MC, distributions of crucial cuts on sideband events are demonstrated below, together with a few other kinematic parameters of the particles. Both EXP data and MC data are shown in these histograms, and MC data is normalised with respect to the luminosity of experimental data. Daggers indicate the EXP data and colourful bars indicate the corrected MC data. The ratio between EXP events and MC events including backgrounds and signal are shown in the lower plot and the MC signal events is assumed of a branching fraction of 1.7×10^{-5}.

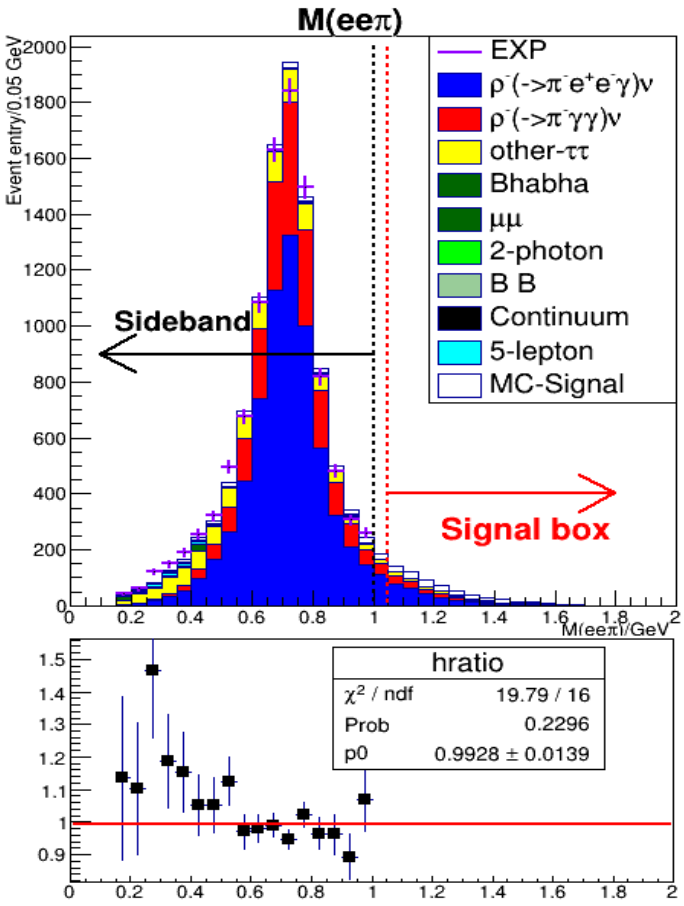


Figure 33: The distribution of invariant mass of $\pi^- e^+ e^-$. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data. The region (1.05 <M($\pi^- e^+ e^-$) < 1.8 GeV/c^2) on the right side of the red vertical dotted line is the signal box (a zoom-in picture is shown in next figure). The region (M($\pi^- e^+ e^-$) < 1 GeV/c^2) on the left side of the red vertical dotted line is the sideband, where both EXP and MC data are displayed. A narrow gap is kept between them for the purpose of blinding. In this chapter, in the lower plot, hratio=EXP/MC, MC includes BKG and Signal samples.

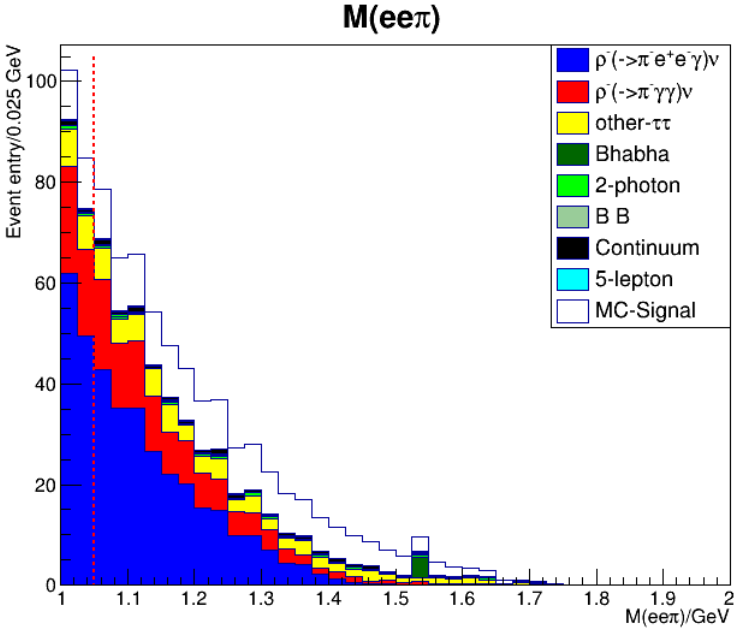
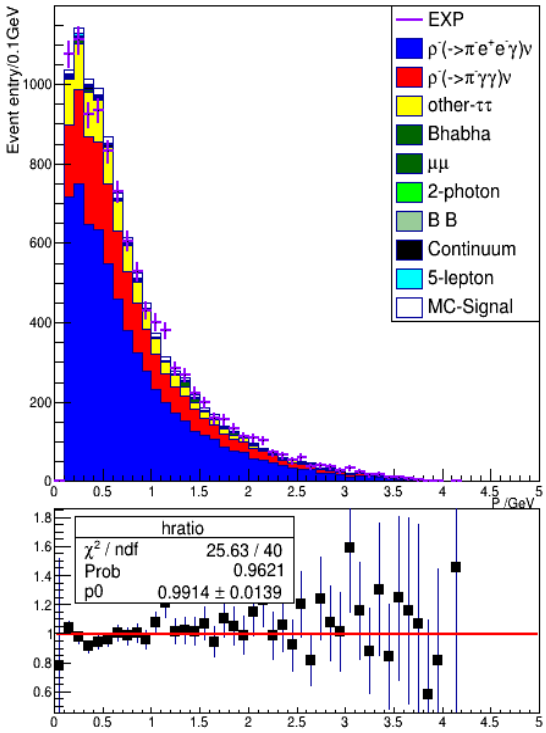


Figure 34: Zoom in plot of the signal box region (MC only).

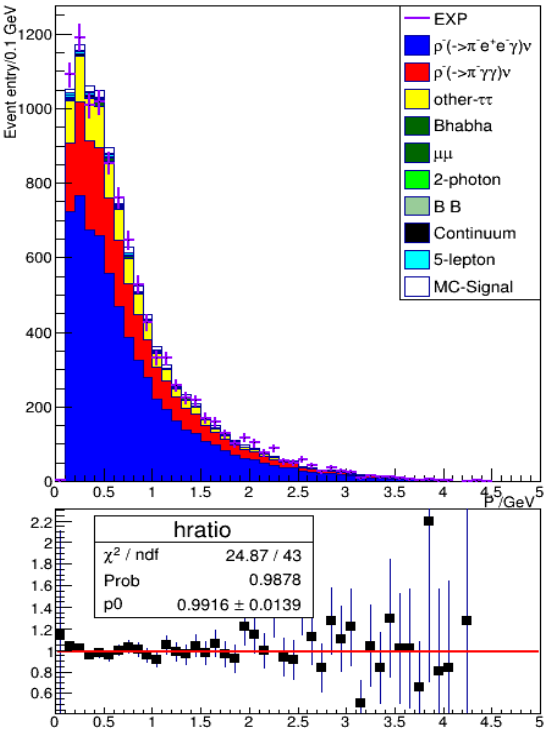
5.1.2 Sideband results

After applying corrections (particle identification, tracking and π^0) on MC samples, in the sideband, MC predicts 10051.6 BKG events and 202.5 signal events (assuming Br is 1.7×10^{-5}). The signal contamination in sideband is 1.97%. 10243 EXP events are observed. The momentum spectra of final state particles in CMS frame are also presented in Figure 35. The angle between the e^+e^- pair ($\cos(e^+e^-)$) in CMS frame is shown in Figure 36. Good consistency is observed between the EXP data and the corrected MC data in the sideband.

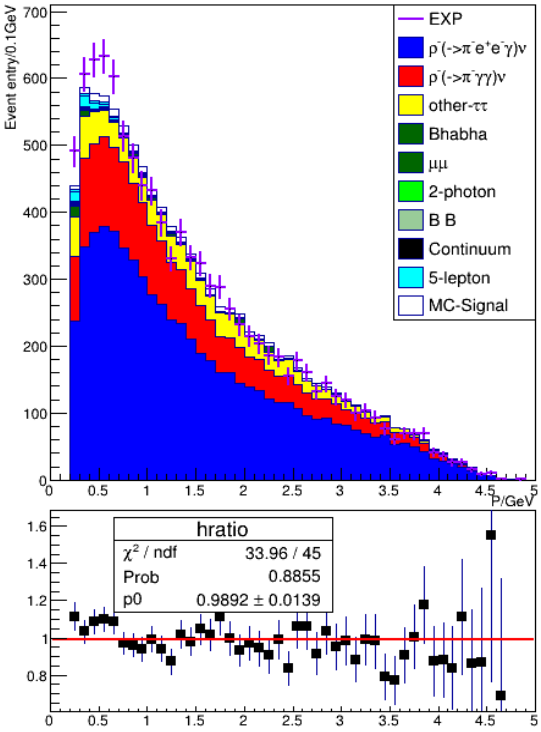
The distribution of the selection cuts are shown as well. The distribution of transverse position of the secondary e^+e^- vertex is shown in Figure 37. The distribution of longitudinal position of the secondary e^+e^- vertex is shown in Figure 38. The distribution of $\frac{\sum_i |\vec{p_i} \cdot \vec{n_T}|}{\sum_i |\vec{p_i}|}$ (thrust magnitude) is shown in Figure 39. The distribution of $\cos(\tau - 3\text{prongs})$ is shown in Figure 40. The distribution of number of photon in signal hemisphere is shown in Figure 41. The distribution of energy of photon in signal hemisphere is shown in Figure 42. These plots are drawn with the cut on that specific variable released while other cuts are still functioning. Finally, the distribution of the invariant mass of the π^0 candidates is shown in Figure 43, where a huge discrepancy that cannot be accounted for by the accuracy of the MC samples is observed, as shown in the left plot. Due to this discrepancy, the study of the reference mode $\tau^- \to \pi^- e^+ e^- \nu_\tau \gamma$ has been carried out and explained in the Section 4.2.1. After applying the π^0 correction, the normalisation of the MC events agrees with the EXP events as shown in the right plot of Figure 43.



(a) Momentum spectrum of electron in CMS frame.



(b) Momentum spectrum of positron in CMS frame.



(c) Momentum spectrum of pion in CMS frame.

Figure 35: Momenta of final state particles in CMS frame. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

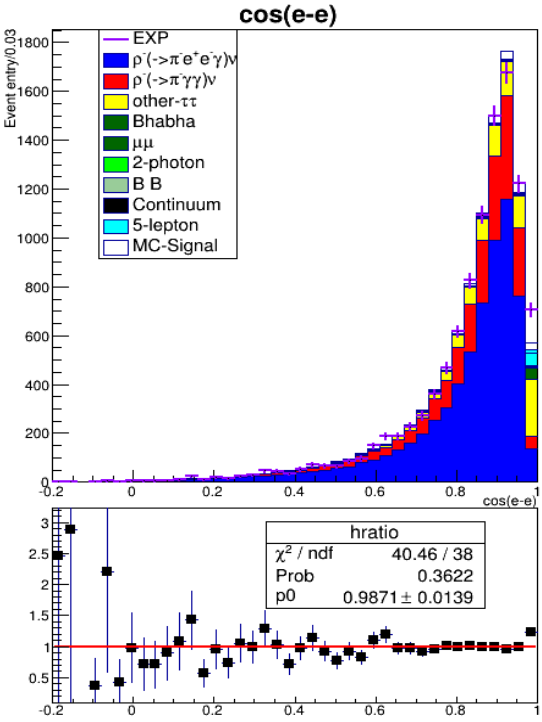


Figure 36: The distribution of angle between the e^+e^- pair ($\cos(e^+e^-)$). Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

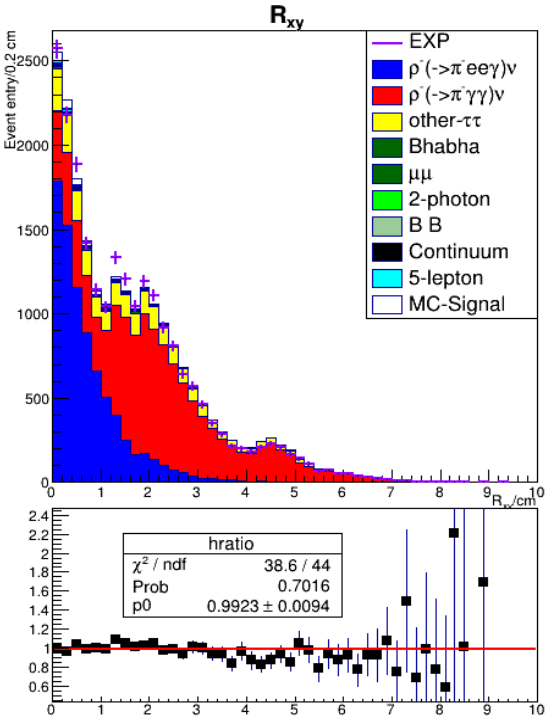


Figure 37: The distribution of transverse position (R_{xy}) of the secondary e^-e^+ vertex. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

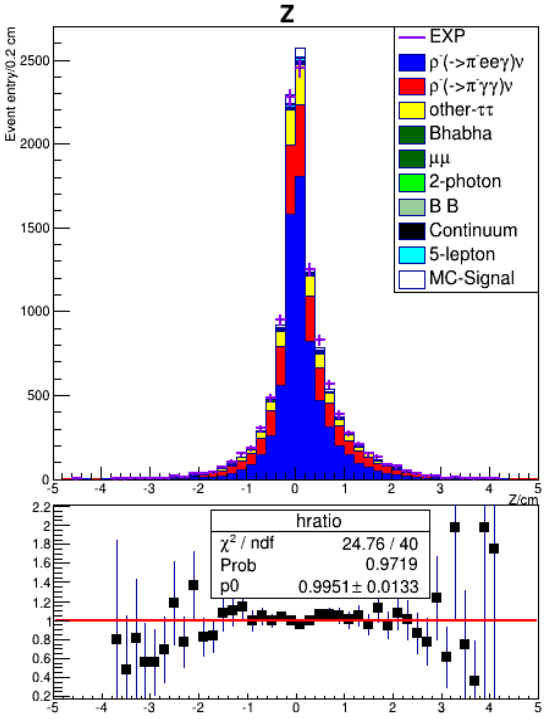


Figure 38: The distribution of longitudinal position (Z) of the secondary e^-e^+ vertex. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

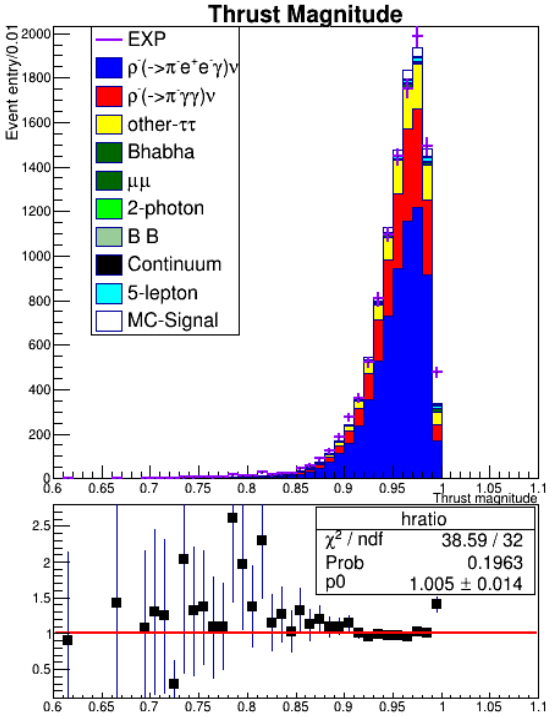


Figure 39: The distribution of thrust magnitude. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

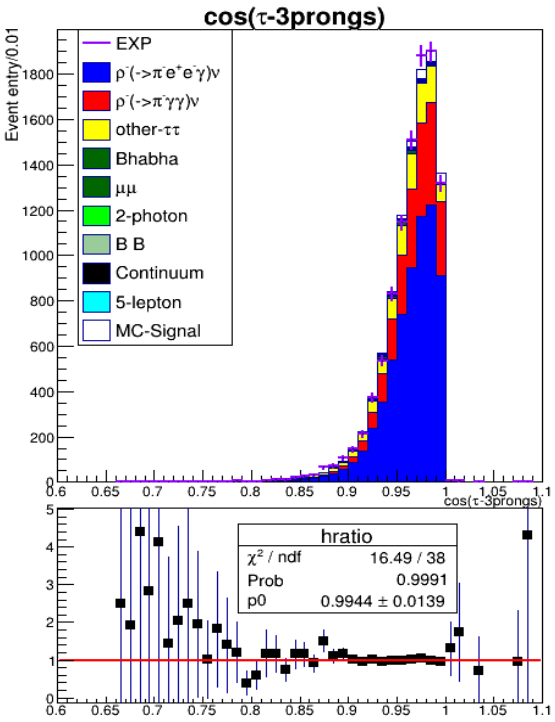


Figure 40: The distribution of $\cos(\tau - 3\text{prongs})$. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

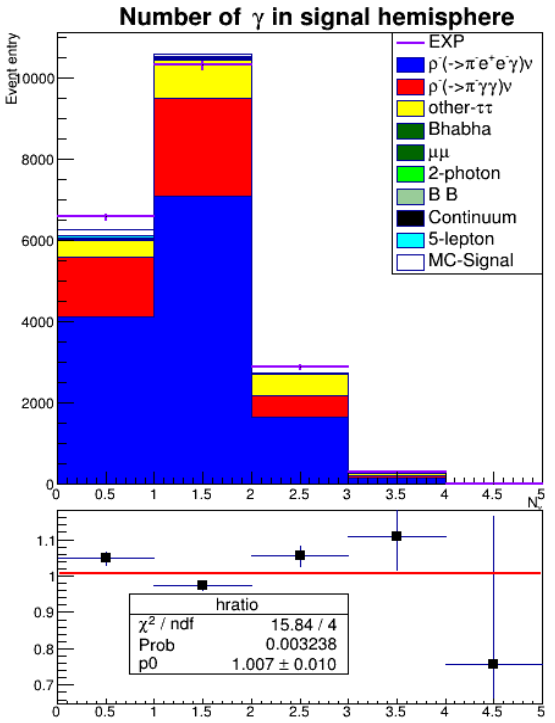


Figure 41: The distribution of number of photon in signal hemisphere. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

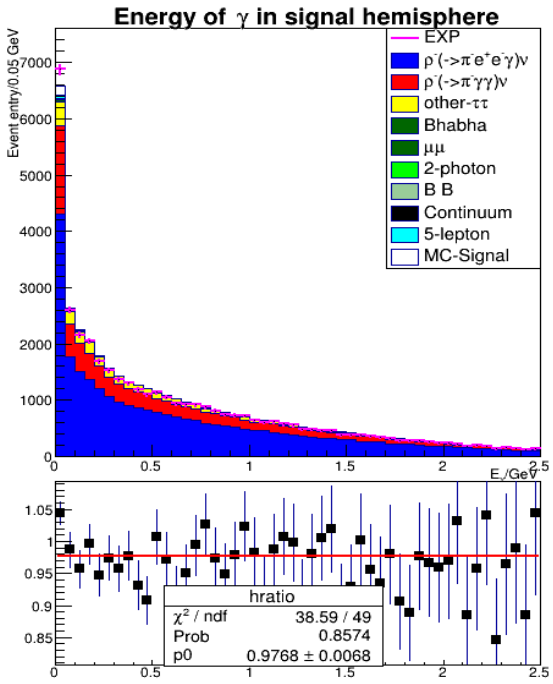
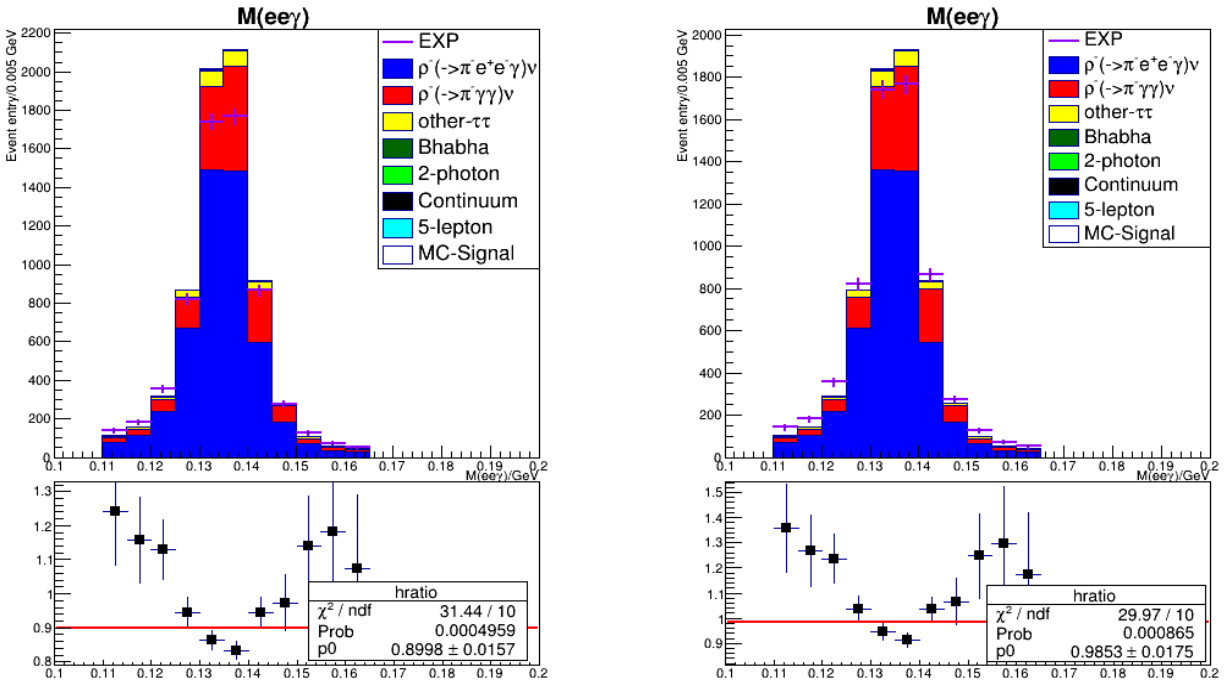


Figure 42: The distribution of energy of photon in signal hemisphere. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.



(a) M(eeγ) of the π^0 candidates without correction.

(b) M(eeγ) of the π^0 candidates with correction.

Figure 43: The distribution of M(eeγ) of the π^0 candidates without π^0 correction is shown in (a) and with π^0 correction is shown in (b). Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

5.2 Sideband study for $\tau^- \to \pi^-\mu^+\mu^-\nu_\tau$

5.2.1 Definition of the sideband

In this decay mode, the invariant mass of 3 prongs does not offer strong discrimination, as shown in Figure 44, where it is clearly demonstrated that the distribution of BKG MC is quite similar to the distribution of signal MC. Thus, we seek for other variables. Among all the selection cuts on $\tau^- \to \pi^-\mu^+\mu^-\nu_\tau$, the transverse radius of $\mu^+\mu^-$ vertex serves as an excellent cut to reject background events as shown in Figure 45. Background components located on both sides of this cut own similar scale ratio. Therefore, this variable is used to define the signal box. The range of transverse radius > 0.2 cm is taken as the sideband, see Figure 46 (again, all selections in the last chapter are applied, except that now $R_{xy} > 0.2$ cm is required). In order to validate the MC, distributions of crucial cuts on sideband events are demonstrated below, together with a few other kinematic parameters of the particles. Both EXP data and MC data are presented in these histograms, and MC data is normalised with respect to the luminosity of experimental data. Daggers indicate the EXP data and colourful bars indicate the corrected MC data. The ratio between EXP events and MC events including backgrounds and signal are shown in the lower plot and the MC signal events is assumed of a branching fraction of 1.0×10^{-5}.

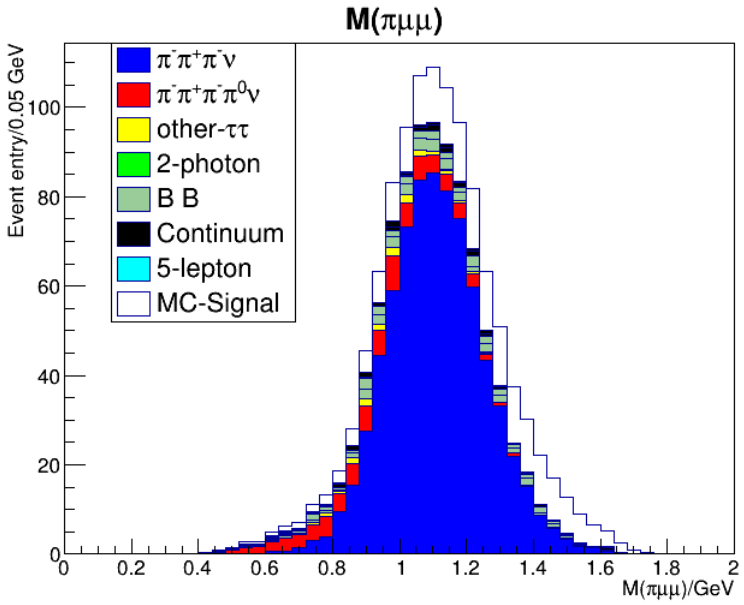


Figure 44: The distribution of the invariant mass of 3 prongs. Colourful bars indicate BKG MC, while the white bars indicate signal MC. Unlike the case of $\tau^- \to \pi^-e^+e^-\nu_\tau$, this variable does not offer strong discrimination against BKG events.

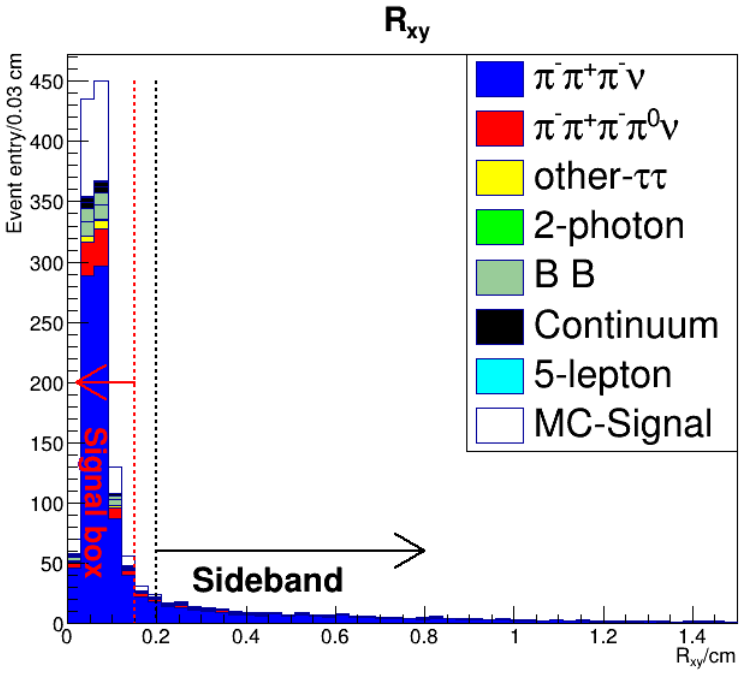


Figure 45: The distribution of transverse position of $\mu^+\mu^-$ vertex (R_{xy}) of MC from 0 to 1.5 cm. The region ($R_{xy} < 0.15$ cm) on the left side of the first red vertical dotted line is the signal box. The region ($R_{xy} > 0.2$ cm) on the right side of the second red vertical dotted line is the sideband. A narrow gap is kept between them for the purpose of blinding. It is shown clearly that the signal MC mainly located in the signal box region.

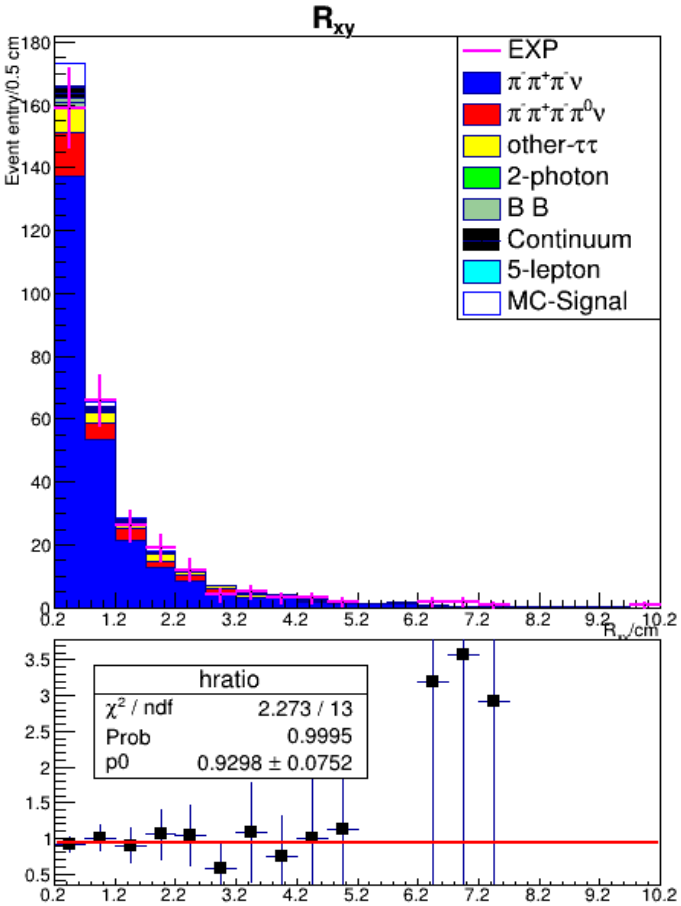
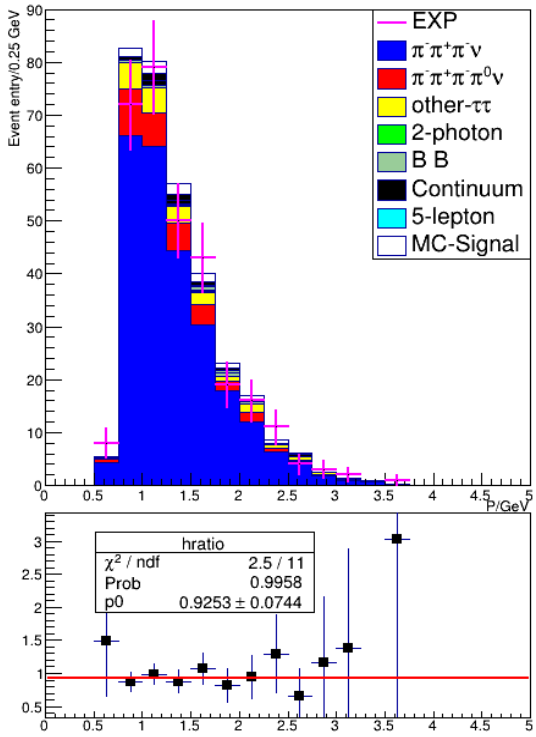


Figure 46: The distribution of transverse position of $\mu^+\mu^-$ vertex (R_{xy}) in the sideband region. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

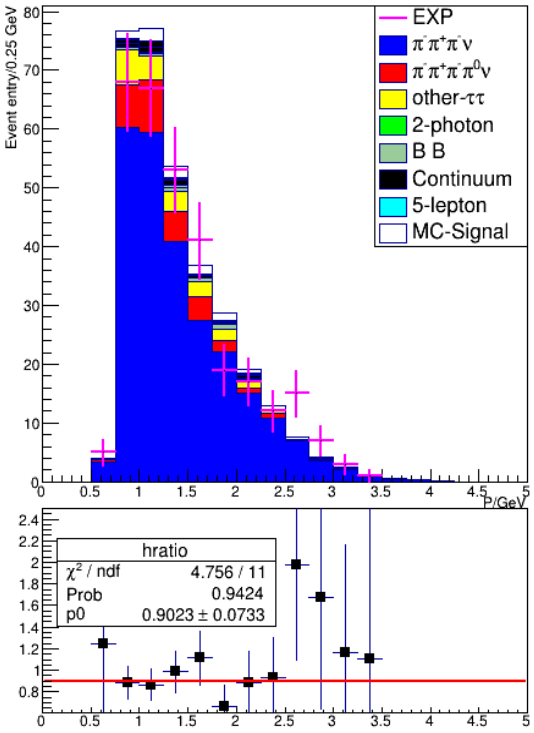
5.2.2 Sideband results

After applying corrections (particle identification, tracking and $\pi \to \mu$ mis-identification) on MC samples, in the sideband, MC predicts 315 BKG events and 10 signal events (assuming Br is 1.0×10^{-5}). The signal contamination in sideband is 3.15%. 308 EXP events are observed. The momentum spectra of final state particles in CMS frame are presented in Figure 47. The angle between the $\mu^+\mu^-$ pair ($\cos(\mu^+\mu^-)$) in CMS frame is shown in Figure 48. Good consistency is observed between the EXP data and the corrected MC data in the sideband.

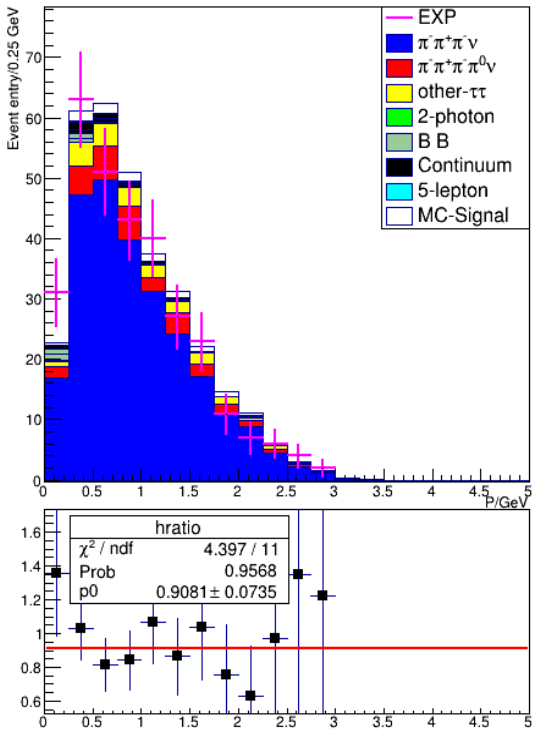
The distribution of the selection cuts are shown as well. The distribution of the thrust magnitude is shown in Figure 49. The distribution of pseudo mass of 3 prongs is shown in Figure 50. The distribution of the invariant mass of $\mu^+\mu^-$ is shown in Figure 51. The distribution of number of photon is shown in Figure 52. The distribution of energy of photon in signal hemisphere is shown in Figure 53. These plots are drawn with the cut on that specific variable released while other cuts are still functioning.



(a) Momentum spectrum of muon in CMS frame.



(b) Momentum spectrum of antimuon in CMS frame.



(c) Momentum spectrum of pion in CMS frame.

Figure 47: Momenta of final state particles in CMS frame. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

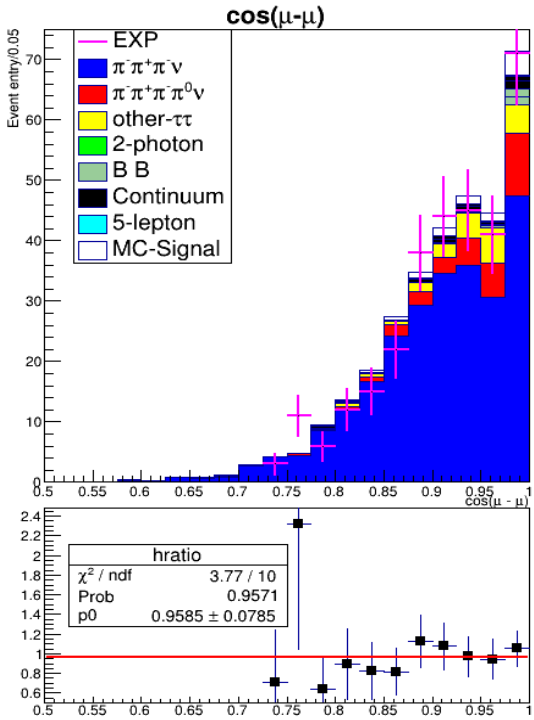


Figure 48: The distribution of angle between the $\mu^+\mu^-$ pair ($\cos(\mu^+\mu^-)$). Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

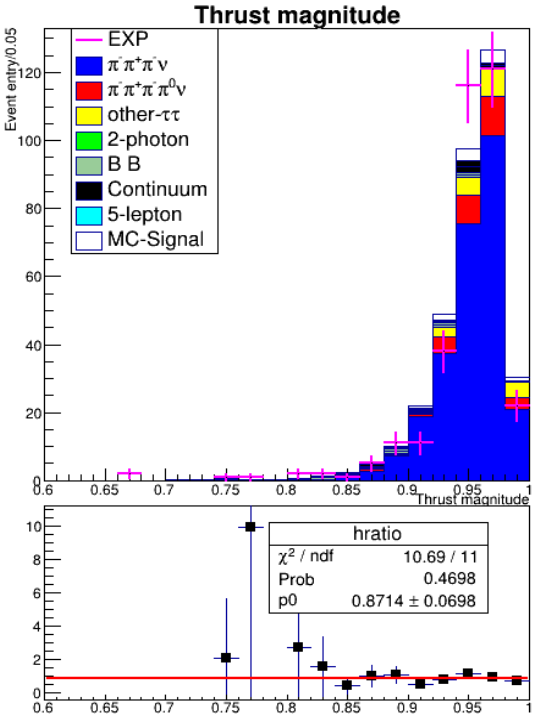


Figure 49: The distribution of thrust magnitude. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

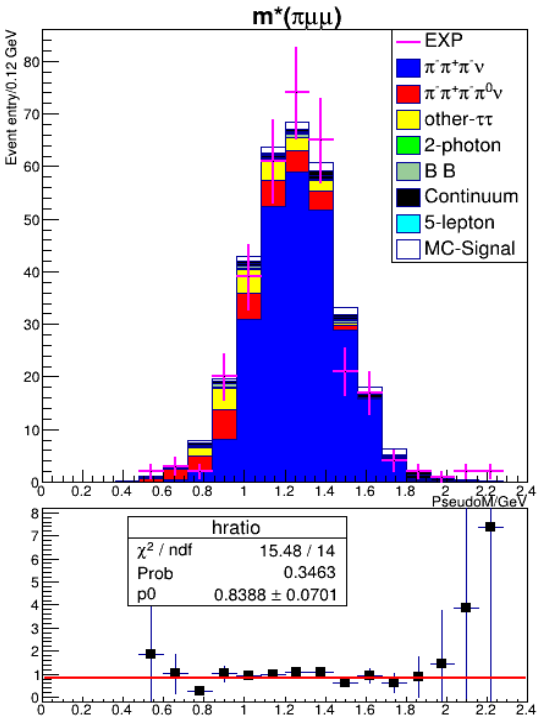


Figure 50: The distribution of pseudo mass of 3 prongs. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

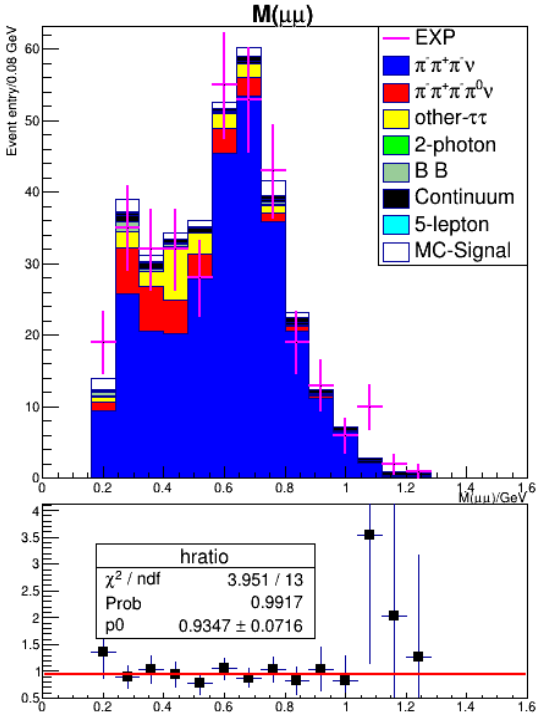


Figure 51: The distribution of invariant mass of $\mu^+\mu^-$. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

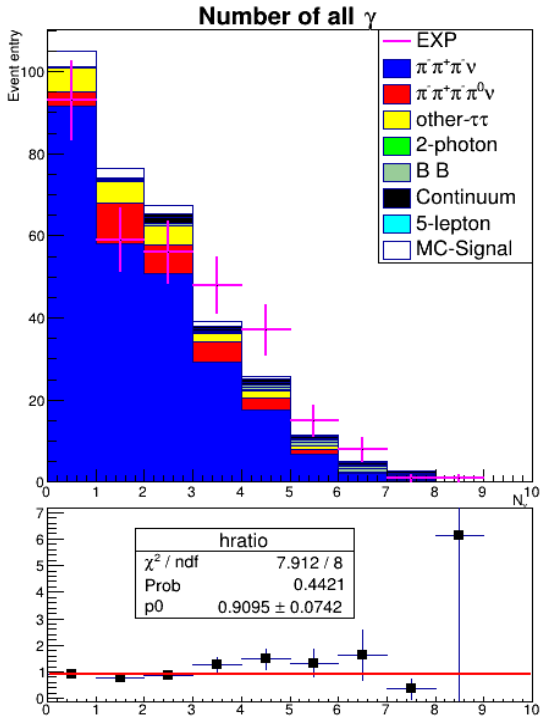


Figure 52: The distribution of number of photon. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

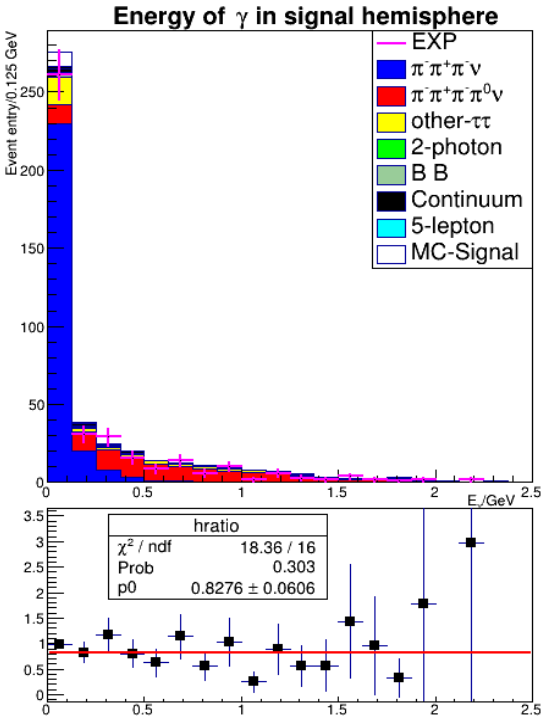


Figure 53: The distribution of energy of photon in the signal hemisphere. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data.

5.3 Expected number of events in the signal box

As shown in the previous two sections of this chapter, the corrected MC data in the sideband, where background processes are the dominant contribution, agrees with experimental data very well, which is a convincing proof that the corrected MC data in signal box would predict accurate background contamination, owing to the similar constituents of background processes in the sideband and the signal box. As a result, the MC data with corrections is used to estimate the number of background events in the signal box.

The number of BKG events is calculated as: $N_{\rm BKG} = \mathcal{L} \cdot (\sum_i \sigma_i \cdot \epsilon_i)$, where $\mathcal{L} = 562$ fb^{-1} indicates luminosity, σ indicates cross section, ϵ indicates detection efficiency and index i runs over all BKG processes at $\Upsilon(4S)$.

By analogy, the number of signal events is calculated as: $N_{\rm sig} = \mathcal{L} \cdot \sigma_{\tau\tau} \cdot \mathcal{B} \cdot \epsilon_{\rm sig}$, where $\mathcal{L} = 562$ fb^{-1} indicates luminosity, $\sigma_{\tau\tau} = 0.919$ nb indicates cross section of $e^+e^- \to \tau^+\tau^-$ at the energy of $\Upsilon(4S)$, $\mathcal{B}$ indicates the branching fraction of the decay mode of concern and $\epsilon_{\rm sig}$ indicates detection efficiency . In the calculation of expected number below, the branching fraction of $\tau^\pm \to \pi^\pm e^+ e^- \nu_\tau$ is taken as 1.7×10^{-5} and the branching fraction of $\tau^\pm \to \pi^\pm \mu^+ \mu^- \nu_\tau$ is taken as 1×10^{-5}.

The method to estimate errors from every factor is the same for the terms in the uncertainty of branching fraction. Therefore, details of the estimation can be referred to in Section 6.2.

5.3.1 $\tau^- \to \pi^- e^+ e^- \nu_\tau$

Using the MC samples that have been corrected on the particle identification, tracking and π^0 veto, contents of events surviving the selection criteria have beed studied.As shown in Figure 34, in the signal box, in total, there are 623 $\pm$ 23 events that pass the selection, including 458 $\pm$ 22 BKG events and 165.4 $\pm$ 6 signal events, as shown in Table 12.

The relative error on the number of BKG events $\triangle N_{\rm BKG}/N_{\rm BKG} = 4.8\%$, $\triangle N_{\rm BKG} = 22$, is obtained from the quadrature sum of luminosity (1.4%), detection efficiency (1.6%), branching fractions of BKG modes (1.9%), PID (3.5%), tracking (1.4%), trigger (0.4%) and π^0 veto corrections (0.8%), as summarised in Table 13. The relative error on the number of signal events $\triangle N_{\rm sig}/N_{\rm sig} = 3.8\%$, $\triangle N_{\rm sig} = 6.3$, arises from luminosity (1.4%), detection efficiency (0.4%), $\tau\tau$ cross section (0.3%), PID (3.2%), tracking (1.4%), trigger (0.5%) and π^0 veto corrections (0.02%), as summarised in Table 14.

Background events remaining in the signal box can be categorised as follows, based on MC studies: generic $\tau\tau$ processes account for 93.63% of the background MC events, B decays 0.04%, continuum process 2.70%, 2-photon 2.66%, Bhabha 0.88% and five lepton tau decay 0.09%, as shown in Table 15. In particular, the dominant contributions of generic $\tau\tau$ processes are $\tau^- \to \pi^-\pi^0(\to e^+e^-\gamma)\nu_\tau$ (55.20% of total BKG) and $\tau^- \to \pi^-\pi^0(\to \gamma\gamma)\nu_\tau$ (23.99% of total BKG).

Table 12: Summary of MC results of $\tau^- \to \pi^- e^+ e^- \nu_\tau$.

τ^- decay mode	$\pi^- e^+ e^- \nu_\tau$
Detection efficiency ($\epsilon_{\rm sig}$)	$(1.884 \pm 0.072)\%$
Main background	$\pi^-\pi^0(e^+e^-\gamma)\nu_\tau$, $\pi^-\pi^0(\gamma\gamma)\nu_\tau$
Expected number of BKG events ($N_{\rm BKG}$)	458 ± 22
Expected number of signal events ($N_{\rm sig}$)	165.4 ± 6.3
Figure of Merit	6.6

Table 13: The breakdown of $\triangle N_{\rm BKG}$ from MC studies.

Luminosity	1.4%
Detection efficiency	1.6%
$\mathcal{B}$s of BKG modes	1.9%
PID	3.5%
Tracking	1.4%
Trigger	0.4%
π^0 veto	0.8%
In total	4.8%

Table 14: The breakdown of $\triangle N_{\rm sig}$ from MC studies.

Luminosity	1.4%
Detection efficiency	0.4%
$\tau\tau$ cross section	0.3%
PID	3.2%
Tracking	1.4%
Trigger	0.5%
π^0 veto	0.02%
In total	3.8%

Table 15: The contents of BKG from MC studies.

$\tau^- \to \pi^-\pi^0(e^+e^-\gamma)\nu_\tau$	55.20%
$\tau^- \to \pi^-\pi^0(\gamma\gamma)\nu_\tau$	23.99%
Other τ decays	14.44%
Continuum processes	2.70%
2-photon processes	2.66%
B decays	0.04%
Five lepton	0.09%
Bhabha	0.88%

5.3.2 $\tau^+ \to \pi^+ e^+ e^- \bar{\nu}_\tau$

Using the MC samples that have been corrected on the particle identification, tracking and π^0 veto, contents of events surviving the selection criteria have beed studied. In total, 621 $\pm$ 22 events would pass the selection, including 455 $\pm$ 21 BKG events and 165.6 $\pm$ 6 signal events, as shown in Table 16.

The relative error on the number of BKG events $\triangle N_{\rm BKG}/N_{\rm BKG} = 4.7\%$, $\triangle N_{\rm BKG} = 21$, arises from luminosity (1.4%), detection efficiency (1.6%), branching fractions of BKG modes (1.9%), PID (3.3%), tracking (1.4%), trigger (0.4%) and π^0 veto corrections (0.8%), as summarised in Table 17. The relative error on the number of signal events $\triangle N_{\rm sig}/N_{\rm sig} = 3.7\%$, $\triangle N_{\rm sig} = 6.1$, arises from luminosity (1.4%), detection efficiency (0.4%), $\tau\tau$ cross section (0.3%), PID (3.1%), tracking (1.4%), trigger (0.5%) and π^0 veto corrections (0.02%), as summarised in Table 18.

Background events remaining in the signal box can be categorised as follows, based on MC studies: generic $\tau\tau$ processes account for 94.66% of the background MC events, continuum process 2.30%, 2-photon 2.69%, B decays 0.10%, di-muon 0.16% and five lepton tau decay 0.09%, as shown in Table 19. In particular, the dominant contributions of generic $\tau\tau$ processes are $\tau^+ \to \pi^+\pi^0(\to e^+e^-\gamma)\nu_\tau$ (56.59% of total BKG) and $\tau^+ \to \pi^+\pi^0(\to \gamma\gamma)\nu_\tau$ (22.97% of total BKG).

Table 16: Summary of MC results of $\tau^+ \to \pi^+ e^+ e^- \bar{\nu}_\tau$.

τ^- decay mode	$\pi^- e^+ e^- \nu_\tau$
Detection efficiency ($\epsilon_{\rm sig}$)	(1.886 $\pm$ 0.070)%
Main background	$\pi^+\pi^0(e^+e^-\gamma)\nu_\tau$, $\pi^+\pi^0(\gamma\gamma)\nu_\tau$
Expected number of BKG events ($N_{\rm BKG}$)	455 $\pm$ 21
Expected number of signal events ($N_{\rm sig}$)	165.6 $\pm$ 6.1
Figure of Merit	6.6

Table 17: The breakdown of $\triangle N_{\rm BKG}$ from MC studies.

Luminosity	1.4%
Detection efficiency	1.6%
$\mathcal{B}$s of BKG modes	1.9%
PID	3.3%
Tracking	1.4%
Trigger	0.4%
π^0 veto	0.8%
In total	4.7%

Table 18: The breakdown of $\triangle N_{\rm sig}$ from MC studies.

Luminosity	1.4%
Detection efficiency	0.4%
$\tau\tau$ cross section	0.3%
PID	3.1%
Tracking	1.4%
Trigger	0.5%
π^0 veto	0.02%
In total	3.7%

Table 19: The contents of BKG from MC studies.

$\tau^+ \to \pi^+\pi^0(e^+e^-\gamma)\nu_\tau$	56.59%
$\tau^+ \to \pi^+\pi^0(\gamma\gamma)\nu_\tau$	22.97%
Other τ decays	15.10%
Continuum processes	2.30%
2-photon processes	2.69%
B decays	0.10%
Five lepton	0.09%
Di-muon	0.16%

5.3.3 $\tau^- \to \pi^-\mu^+\mu^-\nu_\tau$

Using the MC samples that have been corrected on the particle identification, tracking and $\pi \to \mu$ mis-identification (only on generic $\tau\tau$ MC), contents of events surviving the selection criteria have beed studied. As shown in Figure 45, in the signal box, in total, there are 1142 $\pm$ 45 events that pass the selection, including 937 $\pm$ 44 BKG events and 205 $\pm$ 8 signal events, as shown in Table 20.

The relative error on the number of BKG events $\triangle N_{\rm BKG}/N_{\rm BKG} = 4.8\%$, $\triangle N_{\rm BKG} = 44$, arises from luminosity (1.4%), detection efficiency (1.7%), branching fractions of BKG modes (1.0%), PID (3.6%), tracking (1.4%), trigger (0.4%) and $\pi \to \mu$ mis-identification correction factor (1.5%), as summarised in Table 21. The relative error on the number of signal events $\triangle N_{\rm sig}/N_{\rm sig} = 3.8\%$, $\triangle N_{\rm sig} = 8$, arises from luminosity (1.4%), detection efficiency (0.3%), $\tau\tau$ cross section (0.3%), PID (3.2%), tracking (1.4%) and trigger (0.7%), as summarised in Table 22.

Background events remaining in the signal box can be categorised as follows, based on MC studies: generic $\tau\tau$ processes account for 90.80% of the background MC events, B decays 6.37%, continuum processes 2.40%, 2-photon 0.13% and five lepton tau decay 0.30%, as shown in Table 23. In particular, the dominant contributions of generic $\tau\tau$ processes is $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ (81.06% of total BKG) and $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$ (8.14% of total BKG).

Table 20: Summary of MC results of $\tau^- \to \pi^-\mu^+\mu^-\nu_\tau$.

τ^- decay mode	$\pi^-\mu^+\mu^-\nu_\tau$
Detection efficiency ($\epsilon_{\rm sig}$)	$(3.969 \pm 0.151)\%$
Main background	$\pi^-\pi^+\pi^-\nu_\tau$
Expected number of BKG events ($N_{\rm BKG}$)	937 ± 44
Expected number of signal events ($N_{\rm sig}$)	205 ± 8
Figure of Merit	6.1

Table 21: The breakdown of $\triangle N_{\rm BKG}$ from MC studies.

Luminosity	1.4%
Detection efficiency	1.7%
$\mathcal{B}$s of BKG modes	1.0%
PID	3.6%
Tracking	1.4%
Trigger	0.4%
$\pi \to \mu$ mis-id	1.5%
In total	4.8%

Table 22: The breakdown of $\triangle N_{\rm sig}$ from MC studies.

Luminosity	1.4%
Detection efficiency	0.3%
$\tau\tau$ cross section	0.3%
PID	3.2%
Tracking	1.4%
Trigger	0.7%
In total	3.8%

Table 23: The contents of BKG from MC studies.

$\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$	81.06%
$\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$	8.14%
Other τ decays	1.60%
B decays	6.37%
Continuum processes	2.40%
2-photon processes	0.13%
Five lepton	0.30%

5.3.4 $\tau^+ \to \pi^+\mu^+\mu^-\bar{\nu}_\tau$

Using the MC samples that have been corrected on the particle identification, tracking and $\pi \to \mu$ mis-identification (only on generic $\tau\tau$ MC), contents of events surviving the selection criteria have beed studied. In total, 1142 ± 45 events would pass the selection, including 936 ± 44 BKG events and 206 ± 8 signal events, as shown in Table 24.

The relative error on the number of BKG events $\triangle N_{\rm BKG}/N_{\rm BKG} = 4.7\%$, $\triangle N_{\rm BKG} = 44$, arises from luminosity (1.4%), detection efficiency (1.7%), branching fractions of BKG modes (1.0%), PID (3.5%), tracking (1.4%), trigger (0.4%) and $\pi \to \mu$ mis-identification correction factor (1.5%), as summarised in Table 25. The relative error on the number of signal events $\triangle N_{\rm sig}/N_{\rm sig} = 3.8\%$, $\triangle N_{\rm sig} = 8$, arises from luminosity (1.4%), detection efficiency (0.3%), $\tau\tau$ cross section (0.3%), PID (3.1%), tracking (1.4%) and trigger (0.6%), as summarised in Table 26.

Background events remaining in the signal box can be categorised as follows, based on MC studies: generic $\tau\tau$ processes account for 91.13% of the background MC events, B decays 6.18%, continuum processes 2.19%, 2-photon 0.20% and five lepton tau decay 0.30%, as shown in Table 27. In particular, the dominant contributions of generic $\tau\tau$ processes is $\tau^+ \to \pi^+\pi^+\pi^-\nu_\tau$ (80.77% of total BKG) and $\tau^+ \to \pi^+\pi^+\pi^-\pi^0\nu_\tau$ (8.35% of total BKG).

Table 24: Summary of MC results of $\tau^+ \to \pi^+\mu^+\mu^-\bar{\nu}_\tau$.

τ^+ decay mode	$\pi^+\mu^+\mu^-\nu_\tau$
Detection efficiency ($\epsilon_{\rm sig}$)	(3.990 ± 0.152)%
Main background	$\pi^+\pi^+\pi^-\nu_\tau$
Expected number of BKG events ($N_{\rm BKG}$)	936 ± 44
Expected number of signal events ($N_{\rm sig}$)	206 ± 8
Figure of Merit	6.1

Table 25: The breakdown of on $\triangle N_{\rm BKG}$ from MC studies.

Luminosity	1.4%
Detection efficiency	1.7%
$\mathcal{B}$s of BKG modes	1.0%
PID	3.5%
Tracking	1.4%
Trigger	0.4%
$\pi \to \mu$ mis-id	1.5%
In total	4.7%

Table 26: The breakdown of $\triangle N_{\rm sig}$ from MC studies.

Luminosity	1.4%
Detection efficiency	0.3%
$\tau\tau$ cross section	0.3%
PID	3.1%
Tracking	1.4%
Trigger	0.6%
In total	3.8%

Table 27: The contents of BKG from MC studies.

$\tau^+ \to \pi^+\pi^+\pi^-\nu_\tau$	80.77%
$\tau^+ \to \pi^+\pi^+\pi^-\pi^0\nu_\tau$	8.35%
Other τ decays	2.01%
B decays	6.18%
Continuum processes	2.19%
2-photon processes	0.20%
Five lepton	0.30%

6 Extraction of branching fractions

6.1 Formula of branching fractions

The branching fraction of $\tau^{\pm} \to \pi^{\pm} l^{+} l^{-} \nu_{\tau}$ is measured by counting events, where signal yield is assumed to be the difference between the number of observed events and the number of BKG events predicted by the MC study, in adhering to the following formula:

$$\mathcal{B}(\tau^{\pm} \to \pi^{\pm} l^{+} l^{-} \nu_{\tau}) = \frac{N_{\rm obs} - N_{\rm BKG}}{\sigma_{\tau\tau} \cdot \mathcal{L} \cdot \epsilon_{\rm sig}}, \tag{1}$$

$$N_{\rm BKG} = \mathcal{L} \cdot (\sum_{i} \sigma_i \cdot \epsilon_i), \tag{2}$$

where $\sigma_{\tau\tau} = 0.919 \pm 0.003$ nb is the cross section of $\tau\tau$ production at $\Upsilon(4S)$, $\mathcal{L}$ the luminosity of exploited EXP data, $\epsilon_{\rm sig}$ the detection efficiency of signal events, $N_{\rm obs}$ the number of observed events from EXP data and $N_{\rm BKG}$ the number of BKG events predicted by the MC study with index i running over all BKG processes that can occur at $\Upsilon(4S)$. It is worthwhile to emphasize that the value of $N_{\rm BKG}$ and $\epsilon_{\rm sig}$ are taken from the MC study in the last chapter, section 5.3.

6.2 Methodology of systematic uncertainty's estimation

As signal yield is evaluated by the number of observed events subtracting the number of background events, the uncertainty of branching fraction can be classified in three groups, the uncertainty associated with efficiency estimate (terms in the denominator in Formula (1)), the uncertainty associated with background estimate (the right hand side of Formula (2)) and the correlated terms. The correlations are considered for luminosity, tracking and particle identification.

$$\begin{aligned}\left(\frac{\triangle\mathcal{B}}{\mathcal{B}}\right)^2 = & \left(\frac{\triangle\sigma_{\tau\tau}}{\sigma_{\tau\tau}}\right)^2 + \left(\frac{\frac{\triangle\mathcal{L}}{\mathcal{L}} \cdot N_{\rm obs}}{N_{\rm obs} - N_{\rm BKG}}\right)^2 + \left(\frac{\frac{\triangle R_{\rm trk}}{R_{\rm trk}} \cdot N_{\rm obs}}{N_{\rm obs} - N_{\rm BKG}}\right)^2 + \left(\frac{\frac{\triangle R_{\rm PID}}{R_{\rm PID}} \cdot N_{\rm obs}}{N_{\rm obs} - N_{\rm BKG}}\right)^2 \\ & + \left(\frac{\triangle\epsilon_{\rm sig}}{\epsilon_{\rm sig}}\right)^2 + \left(\frac{\triangle N_{\rm BKG}}{N_{\rm obs} - N_{\rm BKG}}\right)^2 + \left(\frac{\triangle N_{\rm obs}}{N_{\rm obs} - N_{\rm BKG}}\right)^2, \end{aligned} \tag{3}$$

The uncertainty of the branching fraction is expressed by formula (3). The last term, $\frac{\triangle N_{\rm obs}}{N_{\rm obs} - N_{\rm BKG}}$, is regarded as the statistical uncertainty, and the rest terms are used to estimate the systematic uncertainty of the measurement on branching fractions. The physical origins of these terms are expressed as followed.

Tracking efficiency

The efficiency of charged tracks' reconstruction has been investigated thoroughly in Belle [45]. Here we apply the latest result [46] obtained by using the partially reconstructed technique with the decay mode, $D^{*} \to \pi D^{0}$, subsequently $D^{0} \to \pi\pi K^{0}_{S}$ and

$K^0_S \to \pi^+\pi^-$. First, one daughter pion of K^0_S is not reconstructed (this is referred to "partially"), using the kinematic information (mass of D^*, D^0, K^0_S), the momentum of the unreconstructed pion can be restored. Then we search for this unreconstructed daughter pion (this is referred to "fully"), see whether it is detected or not. A track finding efficiency is defined as the number of events that the initially unreconstructed pion is found divided by the number of all events. The associated systematic uncertainty is estimated by comparing the tracking finding efficiencies in EXP data and MC data. As a result, for charged tracks of transverse momenta larger than 200 MeV/c, a systematic uncertainty of 0.35% per track is recommended. Therefore, in this study, an uncertainty of $\frac{\triangle R_{\rm trk}}{R_{\rm trk}}$ =1.4% associated with tracking efficiency is taken into account, as four charged tracks are involved in the final state of the decay mode of concern. Meanwhile, in order to cover the low momentum range ($P_t \leq 200$ MeV/c), an additional investigation of the reconstruction efficiency of low momentum $\pi^\pm$ [47] has be carried out, where another decay chain $B^0 \to D^{*-}\pi^+$ is exploited. The probe is the pion of slow momentum which is emitted in the D^*. The track finding efficiency at low momentum is obtained by comparing the slow momentum pion's distributions in EXP and MC. Normalisation is done by the ratio of the highest momentum bin. Following this approach, the systematic uncertainties of the charged tracks of low momenta in the concerned study have been checked and found to be consistent within 0.35% per track. Hence, an uncertainty of 1.4% is taken for this effect. Its impact on both the number of BKG, the signal yield and the correlation between BKG and signal are taken into account so as to give the systematic uncertainty associated with tracking efficiency of the measurement of branching fraction ($\frac{\frac{\triangle R_{\rm trk}}{R_{\rm trk}} \cdot N_{\rm obs}}{N_{\rm obs} - N_{\rm BKG}}$).

Particle identification correction

As the final state contains a pion and two leptons, uncertainties associated with π identification [43] and lepton identification [44] are considered and estimated based on the following approaches [36].

π identification: The systematic study of the pion identification is implemented via an inclusive decay of D^*, $D^{*+} \to D^0\pi^+_{\rm slow} \to K^-\pi^+\pi^+_{\rm slow}$, where $\pi^+_{\rm slow}$ serves as the tag while K^- and π^+ serve as the probe and get identified by their charge with respect to $\pi^+_{\rm slow}$. The data sample is divided into 384 bins according to the momentum and polar angle of the track (32 momentum divisions $\otimes$ 12 polar angle divisions). The correction factor $R_l = \epsilon_{\rm data}/\epsilon_{\rm MC}$ and its statistical and systematic errors are tabulated with respect to the 384 bins. This result is applied to the concerned study, and $R_{\pi\rm ID} = \frac{1}{N}\sum_l n_l R_l$, where the index l runs over the 384 bins, n_l indicates the number of tracks fall in the corresponding bin l and N indicates the summation of all n_l. The uncertainty associated with π-ID is calculated as: $\triangle R_{\pi\rm ID} = \frac{1}{N}\left(\sqrt{\sum_l (n_l \cdot \triangle R^{\rm stat}_l)^2} + \sum_l n_l \cdot \triangle R^{\rm syst}_l\right) + 0.003$, where it is assumed that the statistical errors in different bins are independent while the systematic errors in different bins are completely correlated and 0.003 is an additional constant error for the possible experiment dependence.

Lepton identification: The systematic study of the lepton identification is implemented via two processes: $\gamma\gamma \to l^+l^-$ and $J/\psi \to l^+l^-$. The $\gamma\gamma \to l^+l^-$ process is a very clean channel with low multiplicity tracks while the influence from hadronic environ-

ment is supplemented by the $J/\psi \to l^+l^-$. The correction factor $R_l = \epsilon_{\rm data}/\epsilon_{\rm MC}$ and its statistical and systematic errors are tabulated with respect to 70 bins (7 polar angle divisions $\otimes$ 10 momentum divisions). Similar to the case of π-ID, the lepton-ID correction factor $R_{\rm leptonID}$ and its associated uncertainties are applied to the concerned study by the same binned method while the systematic error of $R_{\rm leptonID}$ is obtained by a quadratic summation from the results of the aforementioned two-photon and J/ψ decays,

$$\triangle R_{\rm leptonID} = \frac{1}{N} \cdot \sqrt{\sum_l (n_l \cdot \triangle R_l^{\rm stat})^2 + (\sum_l n_l \cdot \triangle R_l^{\rm syst\_\gamma\gamma})^2 + (\sum_l n_l \cdot \triangle R_l^{\rm syst\_J/\psi})^2}.$$

The total particle identification uncertainty is obtained as a quadratic summation of the particle identification uncertainties of three daughter tracks, $R_{\rm PID} = R_{\pi{\rm ID}} \oplus R_{\rm lepton} \oplus R_{\rm anti-lepton}$. In the mode $\tau^- \to \pi^-e^+e^-\nu_\tau$ ($\tau^+ \to \pi^+e^+e^-\nu_\tau$), the PID uncertainty on BKG is 3.5% (3.3%), and 3.2% (3.1%) on signal. In the mode $\tau^- \to \pi^-\mu^+\mu^-\nu_\tau$ ($\tau^+ \to \pi^+\mu^+\mu^-\nu_\tau$), the PID uncertainty on BKG is 3.6% (3.5%), and 3.2% (3.1%) on signal. When calculating its effect on branching fraction, the PID uncertainty from the BKG, from the signal yield and the correlation all are taken into account ($\frac{\frac{\triangle R_{\rm PID}}{R_{\rm PID}} \cdot N_{\rm obs}}{N_{\rm obs} - N_{\rm BKG}}$).

Luminosity

The uncertainty associated with luminosity is obtained to be $\frac{\triangle \mathcal{L}}{\mathcal{L}} = 1.4\%$ [48] by using Bhabha events, where the final state electrons of these Bhabha events are required to be detected in the barrel part of the detector. The uncertainty is dominated by the accuracy of the generator for Bhabha process. Not only its effect in the number of BKG and the signal yield, but also the correlation between them are taken into account ($\frac{\frac{\triangle \mathcal{L}}{\mathcal{L}} \cdot N_{\rm obs}}{N_{\rm obs} - N_{\rm BKG}}$).

Trigger

At Belle, the main function of the trigger is to suppress Bhabha events so as to record other physics events as much as possible due to the limited readout power. In the case of events with more than four tracks and energy sum larger than 5 GeV, this uncertainty does not need to be considered [50], as their signatures are exceedingly different towards Bhabha events. In the concerned decay modes, the number of tracks is equal to four. As a result, the uncertainty associated with trigger is conservatively estimated to be half of the deviation of trigger efficiency to unity. In the mode $\tau^- \to \pi^-e^+e^-\nu_\tau$, the trigger efficiency on BKG is 99.2%, and 99.0% on signal. As a result, the trigger uncertainty for BKG is taken as 0.4% and for signal is taken as 0.5%. In the mode $\tau^- \to \pi^-\mu^+\mu^-\nu_\tau$, the trigger efficiency on BKG is 99.2%, and 98.7% on signal. As a result, the trigger uncertainty for BKG is taken as 0.4% and for signal is taken as 0.7%. When calculating the uncertainty on branching fraction, both its effect arising from BKG (in $\frac{\triangle N_{\rm BKG}}{N_{\rm obs} - N_{\rm BKG}}$) and signal (in $\frac{\triangle \epsilon_{\rm sig}}{\epsilon_{\rm sig}}$) are considered. The trigger efficiency is included in the detection efficiency $\epsilon_{\rm sig}$ and ϵ_i.

$e^+e^- \to \tau^+\tau^-$ cross section

A dedicated study on the precision of the cross section of $e^+e^- \to \tau^+\tau^-$ at the energy scale of $\Upsilon(4S)$ resonance for KKMC Monte Carlo simulation is conducted by developers of KKMC [49], in which the $\sigma_{\tau\tau}$ is obtained to be 0.919 ± 0.003 nb. Therefore, the uncertainty arising from $\frac{\triangle \sigma_{\tau\tau}}{\sigma_{\tau\tau}}$ is estimated to be 0.3%.

$\mathcal{B}$ of Background modes

The accuracies of branching fractions ($\mathcal{B}$) of Background modes are considered, which give

rise to the fluctuation of $N_{\rm BKG}$ and introduce uncertainty in $\frac{\triangle N_{\rm BKG}}{N_{\rm obs}-N_{\rm BKG}}$. The accuracies of the main background decay modes are taken from PDG [38] and listed in Table 28.

Table 28: Summary of accuracies of BRs of main BKG modes. The accuracy of $\tau^{\pm} \to \pi^{\pm}\pi^0(e^+e^-\gamma)\nu_\tau$ is calculated by summing the accuracy of $\tau^{\pm} \to \pi^{\pm}\pi^0\nu_\tau$ (0.35%) and $\pi^0 \to e^+e^-\gamma$ (2.98%) in quadrature. Similarly, the accuracy of $\tau^{\pm} \to \pi^{\pm}\pi^0(\gamma\gamma)\nu_\tau$ is calculated by summing the accuracy of $\tau^{\pm} \to \pi^{\pm}\pi^0\nu_\tau$ (0.35%) and $\pi^0 \to \gamma\gamma$ (0.03%) in quadrature.

BKG decay modes in $\tau^{\pm} \to \pi^{\pm}e^+e^-\nu_\tau$	
$\tau^{\pm} \to \pi^{\pm}\pi^0(e^+e^-\gamma)\nu_\tau$	3.00%
$\tau^{\pm} \to \pi^{\pm}\pi^0(\gamma\gamma)\nu_\tau$	0.35%
BKG decay modes in $\tau^{\pm} \to \pi^{\pm}\mu^+\mu^-\nu_\tau$	
$\tau^{\pm} \to \pi^{\pm}\pi^+\pi^-\nu_\tau$	0.57%
$\tau^{\pm} \to \pi^{\pm}\pi^+\pi^-\pi^0\nu_\tau$	2.59%

Detection efficiency of BKG (Finite size of BKG MC)

This effect mainly represents the uncertainty stemming from the statistical fluctuation of MC background events due to the finite size of BKG MC samples. The statistical error is calculated as a binomial variation of MC samples. In the mode $\tau^{\pm} \to \pi^{\pm}e^+e^-\nu_\tau$, $\frac{\triangle N_{\rm BKG-eff}}{N_{\rm BKG}} = 1.6\%$. In the mode $\tau^{\pm} \to \pi^{\pm}\nu_\tau\mu^+\mu^-$, $\frac{\triangle N_{\rm BKG-eff}}{N_{\rm BKG}} = 1.7\%$. The impact on branching fraction is included in $\frac{\triangle N_{\rm BKG}}{N_{\rm obs}-N_{\rm BKG}}$.

Detection efficiency of signal (Finite size of signal MC)

Similarly, owing to the finite size of signal MC samples, a term of uncertainty on signal detection efficiency calculated as a binomial variation of the signal MC sample, has been taken into account (in $\frac{\triangle\epsilon_{\rm sig}}{\epsilon_{\rm sig}}$). In the mode $\tau^{\pm} \to \pi^{\pm}e^+e^-\nu_\tau$, $\frac{\triangle\epsilon_{\rm sig}}{\epsilon_{\rm sig}} = 0.4\%$. In the mode $\tau^{\pm} \to \pi^{\pm}\nu_\tau\mu^+\mu^-$, $\frac{\triangle\epsilon_{\rm sig}}{\epsilon_{\rm sig}} = 0.3\%$.

π^0 veto correction

From sideband study of $\tau^- \to \pi^-e^+e^-\nu_\tau$, a discrepancy of π^0 veto between EXP data and MC data is observed. A correction factor $R = 0.913 \pm 0.011$ is introduced in Section 4.2.1. Thus, its uncertainty is considered as $\triangle N_{\pi^0} = N_{\pi^0} \times 1.2\%$, where N_{π^0} is the number of events that trigger the π^0 veto. In the case of BKG, the uncertainty ($\frac{\triangle N_{\pi^0}}{N_{\rm BKG}}$) is 0.8%, while in the case of signal, the uncertainty ($\frac{\triangle N_{\pi^0}}{N_{\rm sig}}$) is 0.02%. They are included in $\frac{\triangle N_{\rm BKG}}{N_{\rm obs}-N_{\rm BKG}}$ and $\frac{\triangle\epsilon_{\rm sig}}{\epsilon_{\rm sig}}$, respectively.

$\pi \to \mu$ mis-identification

In the analysis of $\tau^{\pm} \to \pi^{\pm}\nu_\tau\mu^+\mu^-$, the mis-identification of $\pi \to \mu$ is investigated by a reference decay mode $\tau^{\pm} \to \pi^{\pm}\pi^+\pi^-\nu_\tau$. As shown in Section 4.3.1, tabulated correction factors, $R^{\pi^-\to\mu^-}$ ($= \frac{\epsilon_{\rm EXP}(\pi^-\to\mu^-)}{\epsilon_{\rm MC}(\pi^-\to\mu^-)}$) and $R^{\pi^+\to\mu^+}$ ($= \frac{\epsilon_{\rm EXP}(\pi^+\to\mu^+)}{\epsilon_{\rm MC}(\pi^+\to\mu^+)}$), on $\pi \to \mu$ mis-identification are introduced. The uncertainty associated with this effect is considered as $\triangle N_{\pi\to\mu} = \sqrt{\sum_l (n_l \cdot \triangle R_l^{\pi^-\to\mu^-})^2 + \sum_l (n_l \cdot \triangle R_l^{\pi^+\to\mu^+})^2}$, where l runs over all bins. The uncertainty $\frac{\triangle N_{\pi\to\mu}}{N_{\rm BKG}} = 1.5\%$, and its impact on branching fraction is included in $\frac{\triangle N_{\rm BKG}}{N_{\rm obs}-N_{\rm BKG}}$.

6.3 Results of branching fractions and total systematic uncertainties

6.3.1 Opening the signal box of $\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}$

The signal box of $\tau^{-} \to \pi^{-}e^{+}e^{-}\nu_{\tau}$ is opened and shown in the Figure 54, where daggers represent the experimental events and histogram bars represent MC predictions including backgrounds and signal events, the "hratio" in the lower plot is defined as the signal yield divided by the MC signal events assuming a branching fraction of 1.7×10^{-5}. In the signal box, 676 EXP events are observed. Subtracting the expected 458 background events, a signal yield of 218 events is obtained, which give rise to a corresponding branching fraction, $\mathcal{B}(\tau^{-} \to \pi^{-}e^{+}e^{-}\nu_{\tau}) = (2.24 \pm 0.27 \pm 0.30) \times 10^{-5}$, where the first error is statistical and the second error is systematic. The contributions of systematic error is shown in Table 29.

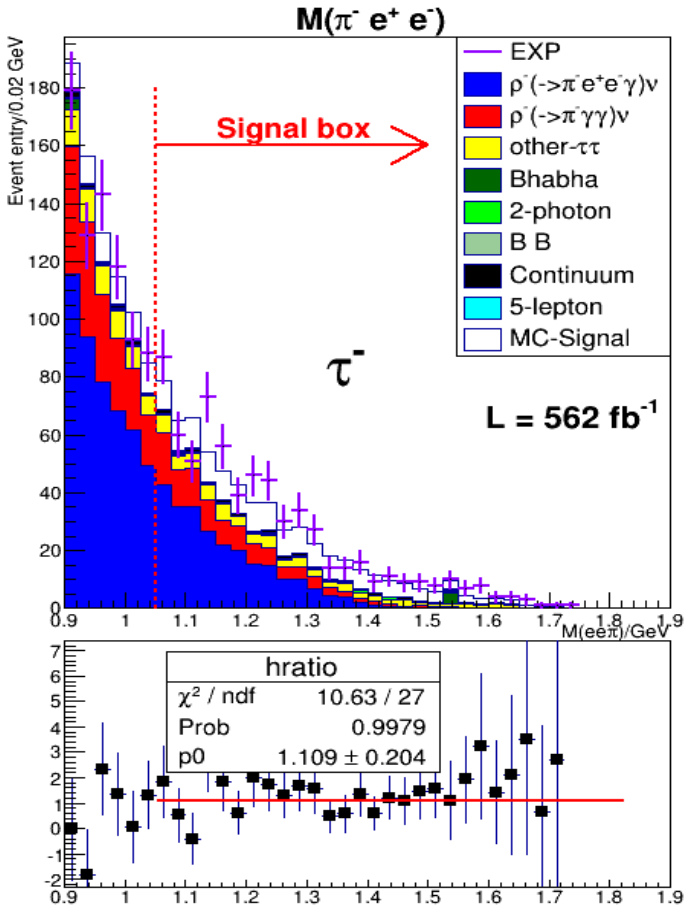


Figure 54: The distribution of invariant mass of 3 prongs. The range of invariant mass of $\pi^{-}e^{+}e^{-}$ from 1.05 to 1.8 GeV/c^2 (the right side of red vertical dotted line) is chosen to define the signal box. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data. Unlike last chapter, in this chapter, in the lower plot, hratio=(EXP − BKG MC)/(Signal MC).

Table 29: The breakdown of systematic uncertainties of the $\tau^- \to \pi^- e^+ e^- \nu_\tau$ branching fraction.

	contents	syst. error
Signal only	Detection efficiency	0.4%
	$\tau\tau$ cross section	0.3%
	Trigger	0.5%
	π^0 veto correction	0.02%
BKG only	Detection efficiency	3.4%
	$\mathcal{B}$s of BKG	4.0%
	Trigger	0.8%
	π^0 veto correction	1.7%
Correlated terms	Luminosity	4.3%
	Tracking	4.3%
	PID	10.5%
Total		13.4%

The signal box of the charge-conjugated mode, $\tau^+ \to \pi^+ e^+ e^- \bar{\nu}_\tau$, is opened and shown in the Figure 55, where daggers represent the experimental events and histogram bars represent MC predictions including backgrounds and signal events, the "hratio" in the lower plot is defined as the signal yield divided by the MC signal events assuming a branching fraction of 1.7×10^{-5}. In the signal box, 689 EXP events are observed. Subtracting the expected 455 background events, a signal yield of 234 events is obtained, which give rise to a corresponding branching fraction, $\mathcal{B}(\tau^+ \to \pi^+ e^+ e^- \bar{\nu}_\tau) = (2.40 \pm 0.27 \pm 0.29) \times 10^{-5}$, where the first error is statistical and the second error is systematic. The contributions of systematic error is shown in Table 30.

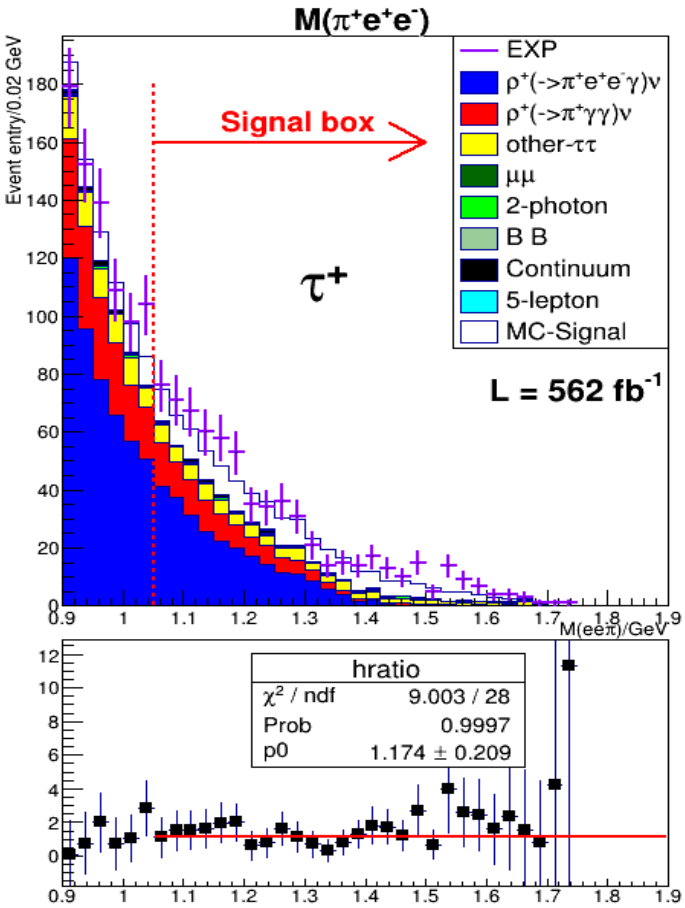


Figure 55: The distribution of invariant mass of 3 prongs. The range of invariant mass of $\pi^+e^+e^-$ from 1.05 to 1.8 GeV/c^2 (the right side of red vertical dotted line) is chosen to define the signal box. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data. In the lower plot, hratio=(EXP − BKG MC)/(Signal MC).

Table 30: The breakdown of systematic uncertainties of the $\tau^+ \to \pi^+e^+e^-\bar{\nu}_\tau$ branching fraction.

	contents	syst. error
Signal only	Detection efficiency	0.4%
	$\tau\tau$ cross section	0.3%
	Trigger	0.5%
	π^0 veto correction	0.02%
BKG only	Detection efficiency	3.1%
	$\mathcal{B}$s of BKG	3.7%
	Trigger	0.8%
	π^0 veto correction	1.6%
Correlated terms	Luminosity	4.1%
	Tracking	4.1%
	PID	9.4%
Total		12.2%

Combining two results of $\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}$ together, the branching fraction of $\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}$ is measured as $\mathcal{B}(\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}) = (2.33 \pm 0.19 \pm 0.30) \times 10^{-5}$, where the first error is statistical and the second error is systematic. The result is consistent with the theoretical prediction, $[1.4, 2.8] \times 10^{-5}$.

6.3.2 Opening the signal box of $\tau^{\pm} \to \pi^{\pm}\mu^{+}\mu^{-}\nu_{\tau}$

The signal box of $\tau^{-} \to \pi^{-}\mu^{+}\mu^{-}\nu_{\tau}$ is opened and shown in the Figure 56, where daggers represent the experimental events and histogram bars represent MC predictions including backgrounds and signal events, the "hratio" in the lower plot is defined as the signal yield divided by the MC signal events assuming a branching fraction of 1.0×10^{-5}. In the signal box, 1001 EXP events are observed. The difference between EXP events and BKG MC is 64 events (=1001-937), which is less than 2 sigma of BKG uncertainty (see Table 20). This signal could be due to the BKG uncertainty so that an upper limit is set to this mode. Here, a frequentist approach taking into account the uncertainty of background [51, 52], a ROOT package TRolke [53], which is an extension of the classic approach done by Feldman and Cousins [54], is applied to calculate the upper limit.

TRolke enables users to compute desired confidence intervals for a Poisson process (in this case, number counting) in the presence of uncertain background and efficiency. 7 different model assumptions are prepared in the TRolke class. For the analysis of concern, the model dealing with a Gaussian uncertainty in the background estimate and a Gaussian uncertainty in the efficiency estimate is adopted. With the results from MC studies (see Table 20) and EXP data, an upper limit on branching fraction at 90% confidence level is calculated to be 0.75×10^{-5}.

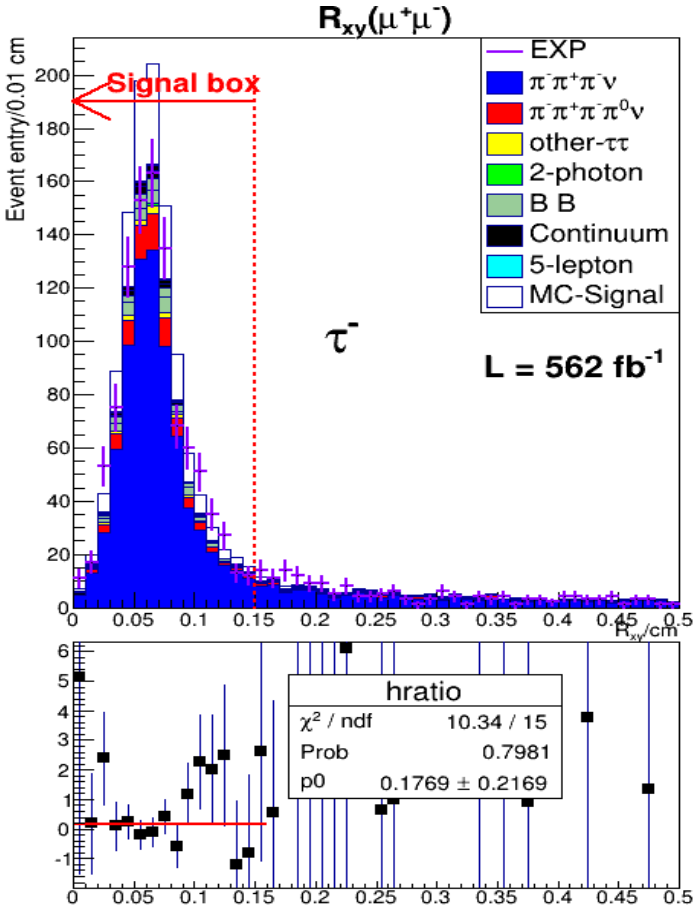


Figure 56: The distribution of transverse position of $\mu^+\mu^-$ vertex. The range of transverse position of $\mu^+\mu^-$ vertex from 0 to 0.15 cm (left side of the red vertical dotted line) is chosen to define the signal box. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data. In the lower plot, hratio=(EXP − BKG MC)/(Signal MC).

The signal box of the charge-conjugated mode, $\tau^+ \to \pi^+\mu^+\mu^-\bar{\nu}_\tau$, is opened and shown in the Figure 57, where daggers represent the experimental events and histogram bars represent MC predictions including backgrounds and signal events, the "hratio" in the lower plot is defined as the signal yield divided by the MC signal events assuming a branching fraction of 1.0×10^{-5}. In the signal box, 967 EXP events are observed. Similarly, an upper limit is set. With the results from MC studies (see Table 24) and EXP data, the upper limit of branching fraction of the charge-conjugated mode is calculated as 0.60×10^{-5}.

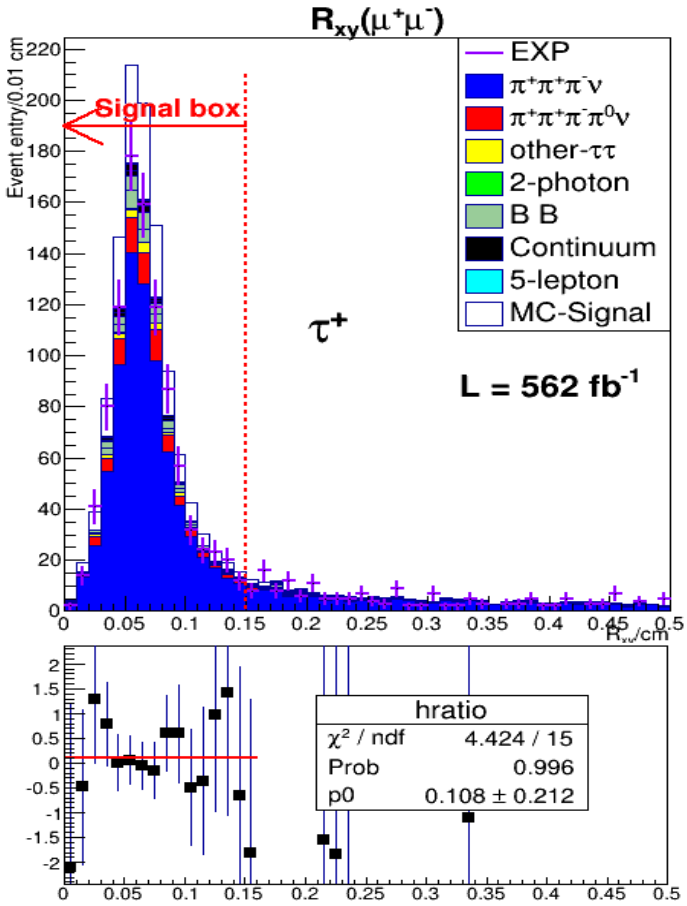


Figure 57: The distribution of transverse position of $\mu^+\mu^-$ vertex. The range of transverse position of $\mu^+\mu^-$ vertex from 0 to 0.15 cm (left side of the red vertical dotted line) is chosen to define the signal box. Daggers stand for EXP data with statistical error and bars for MC. MC data is normalised with respect to the luminosity of experimental data. In the lower plot, hratio=(EXP − BKG MC)/(Signal MC).

Combining two results of $\tau \to \pi\mu^+\mu^-\nu_\tau$ together, the upper limit of branching fraction of $\tau^\pm \to \pi^\pm\mu^+\mu^-\nu_\tau$ is calculated to be 0.55×10^{-5} at 90% confidence level, which is located inside the prediction range from theory $[0.03, 1] \times 10^{-5}$.

7 Conclusion

7.1 Summary of $\mathcal{B}(\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau})$

Using a 562 fb^{-1} dataset collected at the $\Upsilon(4S)$ resonance with Belle detector (SVD2) at the KEKB asymmetric-energy $e^{+}e^{-}$ collider, the branching fraction of $\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}$, $\mathcal{B}(\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}) = (2.33 \pm 0.19 \pm 0.30) \times 10^{-5}$, where the first error is statistical and the second error is systematic, is obtained for the first time with an approximate 7.9σ significance to reject null hypothesis (no signal). The measured value agrees with the theoretical prediction range, $(1.7^{+1.1}_{-0.3}) \times 10^{-5}$, as displayed in the Table 1. Dominant contribution arises from internal bremsstrahlung (corresponding Feynman Diagrams are shown in Figure 1, (a), (b) and (c)) whose calculated error stems from the uncertainties of τ lepton lifetime and pion decay constant. In comparison to internal bremsstrahlung contributions, vector and axial-vector contributions play a minor role in the decay mode $\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}$, while their calculated errors are relatively large.

The decay $\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}$ is the cross channel of $\pi^{\pm} \to l^{\pm}e^{+}e^{-}\nu_{l}$, where the $\gamma^{*}W^{*\mp}\pi^{\pm}$ vertex with two gauge boson off mass shell is also involved, and has been searched for in the past. The branching fraction of $\pi^{\pm} \to e^{\pm}e^{+}e^{-}\nu_{l}$ has been measured with 15% accuracy, $(3.2 \pm 0.5) \times 10^{-9}$, while the branching fraction of $\pi^{\pm} \to \mu^{\pm}e^{+}e^{-}\nu_{\mu}$ decay mode still remains unknown. The measured value of $\mathcal{B}(\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau})$ could help to infer the $\mathcal{B}(\pi^{\pm} \to \mu^{\pm}e^{+}e^{-}\nu_{\mu})$ and could be used to check lepton universality between $\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}$ and $\pi^{\pm} \to e^{\pm}e^{+}e^{-}\nu_{e}$. We have encouraged theorists to do so.

No significant CP Asymmetry is observed between the branching fractions of $\tau^{+} \to \pi^{+}e^{+}e^{-}\nu_{\tau}$, $(2.40 \pm 0.27 \pm 0.29) \times 10^{-5}$, and $\tau^{-} \to \pi^{-}e^{+}e^{-}\nu_{\tau}$, $(2.24 \pm 0.27 \pm 0.30) \times 10^{-5}$. In spite of no evident sign of New Physics (e.g., enhancement on branching fraction due to sterile neutrino), this measurement still stands out as a precise background prediction for lepton flavour violation searches.

7.2 Summary of $\mathcal{B}(\tau^{\pm} \to \pi^{\pm}\mu^{+}\mu^{-}\nu_{\tau})$

Concerning the rare τ decay mode $\tau^{\pm} \to \pi^{\pm}\mu^{+}\mu^{-}\nu_{\tau}$, the number of observed events is smaller than the expected value corresponding to the maximum of theoretical branching fraction. An upper limit $\mathcal{B}(\tau^{\pm} \to \pi^{\pm}\mu^{+}\mu^{-}\nu_{\tau}) < 0.55 \times 10^{-5}$ at 90% confidence level is obtained on this decay mode, which is the first upper limit on this mode. As shown in Table 1, the branching fraction due to internal bremsstrahlung on this mode is smaller than that of $\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}$ by two orders of magnitude. In contrast to $\tau^{\pm} \to \pi^{\pm}e^{+}e^{-}\nu_{\tau}$, the theoretical branching fraction of $\tau^{\pm} \to \pi^{\pm}\mu^{+}\mu^{-}\nu_{\tau}$ is dominated by the vector and axial-vector contributions that have dramatically large uncertainties.

Although the decay mode $\tau^{\pm} \to \pi^{\pm}\mu^{+}\mu^{-}\nu_{\tau}$ is not observed, the upper limit obtained could be used to constrain some New Physics models. In the previous calculation [8] of branching fraction with a sterile neutrino entering the diagram, the upper limit is predicted to be $< 1.3 \times 10^{-5}$, which is higher than the measured value. Therefore, the measured upper limit

could be used to constrain the sterile neutrino's mass, mixing strength and lifetime.

7.3 Prospects

The measurement of branching fractions of $\tau^{\pm} \to \pi^{\pm} l^{+} l^{-} \nu_{\tau}$ is still statistically limited, especially in the case of $l = \mu$. In the case of $l = e$, the uncertainties from statistics and systematics are comparable and at the order of 10%. Nevertheless, a high spatial resolution is the key in this analysis, as the source of severe background events mainly arises from gamma conversion, which can be rejected by the cut on secondary vertex position. Therefore, in this analysis, only SVD2 data is exploited here.

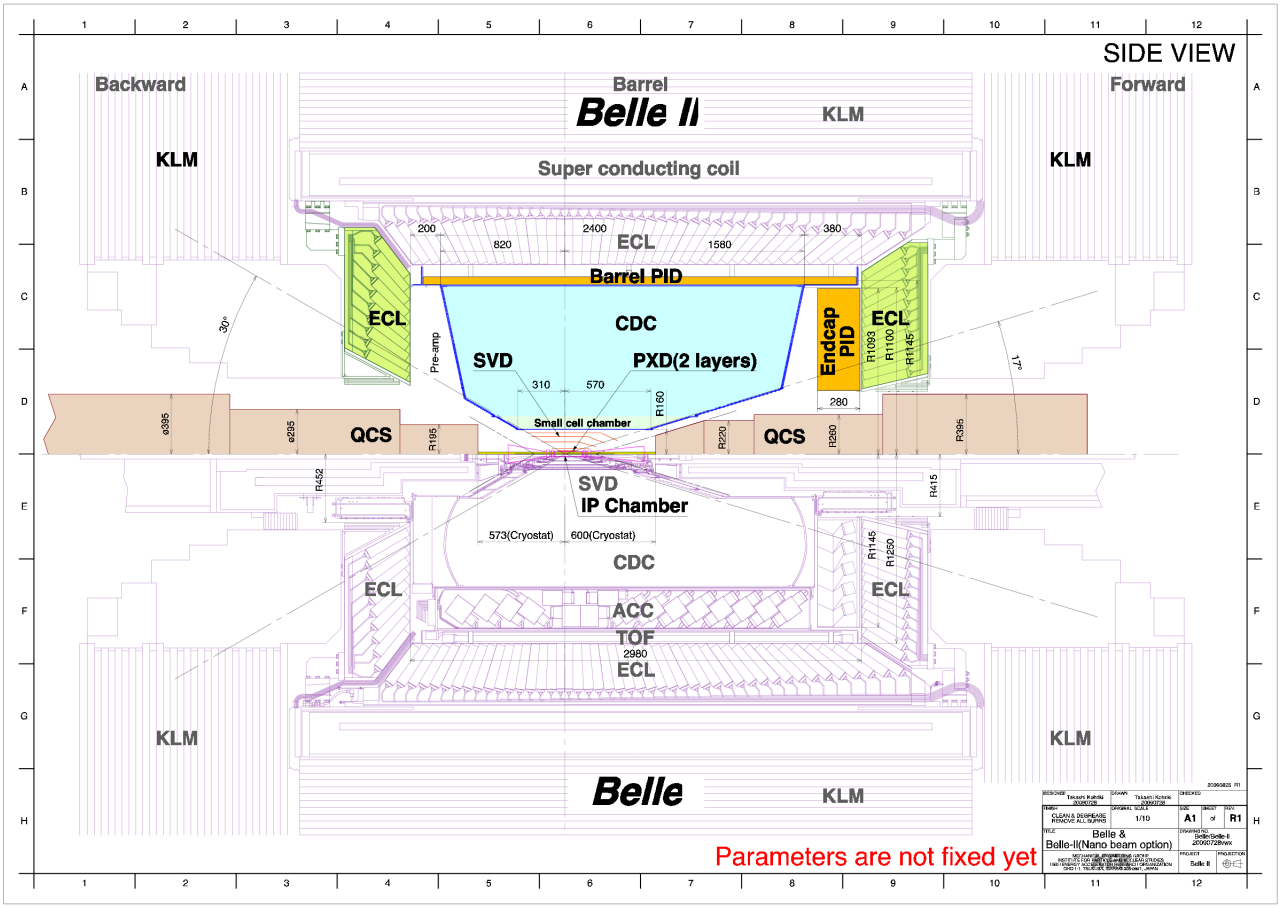


Figure 58: A side-view plot of the Belle II detector in comparison with the Belle detector [55]. The top half is Belle II while the bottom half is Belle detector.

Belle detector is upgraded into Belle II, aiming at searches for physics beyond the Standard Model. With the upcoming Belle II data set, not only a larger statistics will be obtained, also several aspects of the systematic uncertainties can be reduced. A side-view plot of the Belle II detector is shown in Figure 58. First, a new vertex detector (VXD) [56] will be installed in Belle II, which would contribute to an increasingly precise tracking efficiency and substantially improve the spatial resolution on vertex position and impact parameters of the track. As a result, it can result to a better discrimination against backgrounds from gamma conversion.

Second, the upgraded PID system, including a new sub-detector, Time-Of-Propagation detector [56], would contribute to a better separation among μ, π and K, as in this analysis, PID is the leading part of the systematic uncertainty. Third, with a 50 ab^{-1} ($\sim$ 90 times larger than the exploited one) data set, the branching fractions of backgrounds could also be measured more accurately, which will further reduce the systematic uncertainty. As a result, a more precise measurement of $\mathcal{B}(\tau^{\pm} \to \pi^{\pm} l^{+} l^{-} \nu_{\tau})$ will be available. It is very promising to observe the decay $\tau^{\pm} \to \pi^{\pm} \mu^{+} \mu^{-} \nu_{\tau}$ for the first time, whose central value of theoretical branching fraction is 1.938×10^{-6}. Assuming 90 times larger statics, Belle II experiment can reach to a limit 5.6×10^{-7} (with FOM > 5).

References

[1] S. Scherer, Introduction to Chiral Perturbation Theory, arXiv:hep-ph/0210398.

[2] G. Ecker et al., The Role of Resonances in Chiral Perturbation Theory, Nuclear Physics B, 321 (1989): 311-342.

[3] Richard. F. Lebed, Phenomenology of large N_c QCD, World Scientific.

[4] P. Roig, et al., Weak radiative pion vertex in $\tau^- \to \pi^- \nu_\tau l^+ l^-$ decays, Physical Review D, 2013, 88(3): 033007.

[5] A. A. Aguilar-Arevalo et al., Unexplained Excess of Electronlike Events from a 1-GeV Neutrino Beam, Phys. Rev. Lett. 102.10 (2009): 101802.

[6] A. A. Aguilar-Arevalo et al., Observation of a Significant Excess of Electron-Like Events in the MiniBooNE Short-Baseline Neutrino Experiment, arXiv:1805.12028.

[7] S. N. Gninenko, MiniBooNE Anomaly and Heavy Neutrino Decay, PHYSICAL REVIEW LETTERS, 103, 241802 (2009).

[8] C. Dib, et al., Heavy sterile neutrinos in tau decays and the MiniBooNE anomaly, Physical Review D, 85 (2012) 011301.

[9] `https://hflav.web.cern.ch/content/tau`

[10] `http://belle.kek.jp/belle/transparency/`

[11] A. Bevan et al., The Physics of the B Factories, Springer, 2016.

[12] A. Gaz, Indirect constraints on New Physics from the B-factories, arXiv:1411.1882.

[13] A. Abashian et al., The Belle detector, Nuclear Instruments and Methods in Physics Research A 479 (2002) 117–232.

[14] S. Kurokawa et al., Overview of the KEKB accelerators, Nuclear Instruments and Methods in Physics Research A 499 (2003) 1–7.

[15] K. Akai et al., RF systems for the KEK B-Factory, Nuclear Instruments and Methods in Physics Research A 499 (2003) 45–65.

[16] K. Hosoyama, et al., Development of the kek-b superconducting crab cavity, 11th European Particle Accelerator Conference. 2008.

[17] R. Abe et al., The new beampipe for the Belle experiment, Nuclear Instruments and Methods in Physics Research A 535 (2004) 558–561.

[18] M. Yokoyama et al., Radiation hardness of VA1 with submicron process technology. IEEE Transactions on Nuclear Science 48, 440 (2001).

[19] Z. Natkaniec et al., Status of the Belle silicon vertex detector, Nuclear Instruments and Methods in Physics Research A 560 (2006) 1–4.

[20] R. Abe et al., BELLE/SVD2 status and performance, Nuclear Instruments and Methods in Physics Research A 535 (2004) 379–383.

[21] H. Hirano et al., A high-resolution cylindrical drift chamber for the KEK B-factory, Nuclear Instruments and Methods in Physics Research A 455 (2000) 294–304.

[22] T. Sumiyoshi et al., Silica aerogel Cherenkov counter for the KEK B-factory experiment, Nuclear Instruments and Methods in Physics Research A 433 (1999) 385–391.

[23] H. Kichimi et al., The Belle TOF system, Nuclear Instruments and Methods in Physics Research A 453 (2000) 315–320.

[24] K. Miyabayashi et al., Belle electromagnetic calorimeter, Nuclear Instruments and Methods in Physics Research A 494 (2002) 298–302.

[25] A. Abashing et al., The K_L/μ detector subsystem for the BELLE experiment at the KEK B-factory, Nuclear Instruments and Methods in Physics Research A 449 (2000) 112–124.

[26] H. J. Kim et al., A fast programmable trigger for isolated cluster counting in the BELLE experiment, Nuclear Instruments and Methods in Physics Research A 457 (2001) 634–639.

[27] S. Jadach, B. Ward, Z. Was, The precision Monte Carlo event generator KK for two-fermion final states in e^+e^- collisions, Computer Physics Communications, 2000, 130(3): 260-325.

[28] F. A. Berends, P. H. Daverveldt, R. Kleiss, Monte Carlo simulation of two-photon processes: II: Complete lowest order calculations for four-lepton production processes in electron-positron collisions, Computer Physics Communications, 1986, 40(2-3): 285-307.

[29] Written for Belle based on HemiCosm code.

[30] D. J. Lange, The EvtGen particle decay simulation package, Nuclear Instruments and Methods in Physics Research Section A 462(2001) 152.

[31] N. Davidson, et al., Universal interface of TAUOLA: technical and physics documentation, Computer Physics Communications, 2012, 183(3): 821-843.

[32] E. Barberio, B. van Eijk, and Z. Was, Photos — a universal Monte Carlo for QED radiative corrections in decays, Computer Physics Communications, 1991, 66: 115.

[33] R. Brun et al., GEANT 3.21, CERN Report No.DD/EE/84-1, 1987.

[34] `http://belle.kek.jp/group/tautp/tauphys/tsim/tsim.html`

[35] `http://belle.kek.jp/secured/wiki/doku.php?id=physics:taup:start`

[36] `http://belle.kek.jp/group/pid_joint/`

[37] `http://belle.kek.jp/~hitoshi/private/track/`

[38] `http://pdg.lbl.gov/2016/reviews/rpp2016-rev-tau-branching-fractions.pdf`

[39] Kenji Inami, Bellenote # 629, Note on the tau-pair skim.

[40] Denis Epifanov, Bellenote # 855, Study of $\tau^- \to K_S\pi^-\nu_\tau$ decay at Belle.

[41] A. Stahl, Physics with Tau Leptons, Chapter 2, Springer.

[42] L. Piilonen, et al., Bellenote # 338, BELLE Muon Identification.

[43] S. Nishida, Bellenote # 779, Study of Kaon and Pion Identification Using Inclusive D^* Sample.

[44] L. Hinz, Bellenote # 954, Lepton ID efficiency correction and systematic error.

[45] P. Koppenburg, Bellenote # 621, A Measurement of the Track finding efficiency Using partially Reconstructed D^* Decays.

[46] Bipul Bhuyan, Bellenote # 1165, High P_t Tracking Efficiency Using Partially Reconstructed D^* Decays.

[47] Wolfgang Dungel, Bellenote # 1176, Systematic investigation of the reconstruction efficiency of low momentum $\pi^\pm$ and π^0.

[48] `http://belle.kek.jp/group/ecl/private/lum/lum6new.html`

[49] S. Banerjee, et al., Tau and muon pair production cross sections in electron-positron annihilations at $\sqrt{s} = 10.58$ GeV, Physical Review D, 2008, 77(5):054012.

[50] `http://belle.kek.jp/secured/wiki/doku.php?id=software:tsim&s[]=trigger`

[51] Wolfgang. A. Rolke, Angel. M. Lopez, Jan. Conrad, Limits and Confidence Intervals in the Presence of Nuisance Parameters, arXiv:physics/0403059.

[52] J. Lundberg, J. Conrad, W. Rolke, A. Lopez, Limits, discovery and cut optimization for a Poisson process with uncertainty in background and signal efficiency: TRolke 2.0, arXiv:0907.3450.

[53] `https://root.cern.ch/doc/v608/classTRolke.html`

[54] Gary. J. Feldman, Robert. D. Cousins, A Unified Approach to the Classical Statistical Analysis of Small Signals, arXiv:physics/9711021.

[55] C. Z. Yuan, The Belle II Experiment at the SuperKEKB, arXiv:1208.3813.

[56] T. Abe, et al., Belle II technical design report, arXiv:1011.0352, 2010.