CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD ZACATENCO
DEPARTAMENTO DE FÍSICA

"Búsqueda de nueva física en procesos que violan sabor leptónico en el experimento Belle II"

Tesis que presenta

Marcela García Hernández

para obtener el Grado de

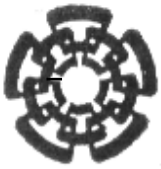
Doctora en Ciencias

en la Especialidad de

Física

Director de tesis: Dr. Iván Heredia de la Cruz

Ciudad de México Junio, 2023



CENTER FOR RESEARCH AND ADVANCED STUDIES OF THE NATIONAL POLYTECHNIC INSTITUTE

PHYSICS DEPARTMENT

"Search for new physics in charged lepton flavor violating processes at the Belle II experiment"

by

Marcela García Hernández

In order to obtain the

Doctor of Science

degree, speciality in

Physics

Advisor: Ph. D. Iván Heredia de la Cruz

Mexico City June, 2023

Search for new physics in charged lepton flavor violating processes at the Belle II experiment

by

Marcela García Hernández

Supervisor: Dr. Ivan Heredia de la Cruz

Submitted in partial fulfillment of the requirements
of the Degree of Doctor of Science in Physics
Department of Physics, CINVESTAV-IPN

Project Period: March, 2019 – June, 2023

Agradecimientos

Quiero expresar mi más sincero agradecimiento a todas las personas que me apoyaron y contribuyeron de diversas formas en la realización de esta tesis.

En primer lugar, deseo agradecer a mi asesor de tesis, Dr. Iván Heredia de la Cruz, por su guía y constante apoyo a lo largo de todo el proceso de investigación. Sus enseñanzas y orientación fueron fundamentales para mi desarrollo profesional y personal. Agradezco sinceramente por su confianza en mi trabajo. Además de ser mi asesor, también fue un colega de trabajo.

A mis sinodales, Dr. Pablo Roig Garcés, Dr. Eduard De La Cruz Burelo, Dr. Heriberto Castilla Valdez, Dr. Pedro Luis Manuel Podesta Lerma y Dr. Josué De Santiago Sanabria por su tiempo, dedicación y valiosos comentarios que contribuyeron a mejorar la calidad de esta tesis.

A los miembros del experimento Belle II, quienes formaron parte de esta emocionante experiencia de investigación. Gracias por contribuir al gran ambiente científico en el que tuve el privilegio de trabajar. En particular a los miembros del grupo de análisis de la búsqueda del decaimiento $\tau^{\pm} \to \ell^{\pm} + \alpha$, Dr. Ivan Heredia, Dr. Eduard de la Cruz, Dr. Armine Rostomyan, Dr. Petar Rados, Dr. Francesco Tenchini, Dr. Alejandro De Yta y Dr. Thomas Kraetzschmar, por su fundamental contribución a nuestro trabajo.

No puedo dejar de mencionar a mis profesores, compañeros y a toda la comunidad del departamento de física del Cinvestav.

A mis amigos y familia, en especial a Horacio Crotte, les estoy profundamente agradecida por su apoyo incondicional durante todo este proceso.

Finalmente, me gustaría expresar mi agradecimiento al Consejo Nacional de Humanidades, Ciencias y Tecnologías (Conahcyt) por el apoyo económico para llevar a cabo esta investigación.

Resumen

En física de partículas, los procesos de violación de sabor leptónico cargado no han sido observados. Si fueran observados, necesariamente conllevarían a escenarios de nueva física. El leptón τ es el leptón más pesado en el modelo estándar de la física de partículas. Es el único leptón suficientemente pesado para decaer en leptones y hadrones; además, sus mecanismos de producción son bien entendidos. Luego, la física del leptón τ es extremadamente conveniente para buscar violación de sabor.

En el experimento Belle II se espera alcanzar la muestra más grande en el mundo de leptones τ en los siguientes años. Esto nos permitirá buscar más de 50 procesos de violación de sabor leptónico en el sector del leptón τ con gran precisión.

Dentro del marco del experimento Belle II, trabajamos en la búsqueda del decaimiento $\tau^{\pm} \to \ell^{\pm} + \alpha$, donde ℓ es un electrón o un muón, y α es un bosón de espín 0 invisible. Además, revisamos el decaimiento $\tau^{\pm} \to \ell^{\pm} + \gamma$, el cual es un canal dorado para buscar violación de sabor leptónico, ya que tiene la tasa de decaimiento más grande en escenarios de nueva física.

Abstract

In particle physics, charged Lepton Flavor Violating (cLFV) processes have not been observed. If they were observed, they would necessarily lead to new physics scenarios. The τ lepton is the heaviest lepton in the Standard Model of particle physics. It is the only lepton heavy enough to decay into leptons or hadrons; moreover, its production mechanisms are well-understood. Thus, τ lepton physics is extremely convenient to search for flavor violation.

The Belle II experiment is expected to achieve the world's largest τ sample in the coming years. This will allow us to search for more than 50 charged lepton flavor violating processes in the τ lepton sector with high precision.

Within the Belle II experiment, we work on the search of the cLFV $\tau^{\pm} \to \ell^{\pm} + \alpha$ decay, where ℓ is either an electron or a muon, and α is an invisible spin-0 boson. In addition, we review the $\tau^{\pm} \to \ell^{\pm} + \gamma$ decay, which is a golden mode to search for lepton flavor violation since it has the largest branching ratio in new physics scenarios.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Standard Model (SM) of particle physics is the most successful theory that describes the behavior of the fundamental particles and forces that constitute our universe.

Regarding the lepton sector, the quantum number known as lepton flavor is found to be conserved. However, in the neutral sector, Lepton Flavor Violation (LFV) can occur due to neutrino oscillations. Neutrino oscillations describe the phenomenon where a neutrino of one flavor transforms into a neutrino of a different flavor, for example, $\nu_e \to \nu_\mu$ or ν_τ. The mixing of the neutrino flavor states is described by the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) unitary matrix. On the charged lepton sector, the LFV is predicted by extensions of the SM (considering ν masses) to be strongly suppressed, on the orders of $\sim 10^{-54} - 10^{-56}$ [14, 15]. So far, no evidence of LFV in the charged sector (cLFV) has been found.

In the field of particle physics, there are still many unsolved questions that the SM is unable to answer. For this reason, new physics models have arisen that introduce new mechanisms to address these questions. In some of these models, cLFV can take place at orders of magnitude close to the current experimental bounds. Therefore, in these orders, any observation of cLFV in the charged sector would be a clear indication of the presence of physics Beyond the Standard Model (BSM).

An important part of the Belle II physics program focuses on the search for cLFV [10, 16]. The Belle II experiment is instrumented at the SuperKEKB electron-positron asymmetric beams collider with a center-of-mass energy $\sqrt{s}$ centered at the $\Upsilon(4S)$ resonance, mainly producing B mesons. This is the reason why Belle II is known as a super B-factory. Moreover, the production of τ lepton pairs is approximately the same as B meson pairs; hence, Belle II can also be regarded as a τ-factory.

The Belle II detector is planned to collect a total integrated luminosity of 50 ab^{-1}, which will correspond to 46 billion τ pairs, the world's largest τ lepton sample [7]. This will allow us to search for more than 50 LFV processes (see Figure 3.9) in the τ sector with high precision [10].

In this thesis, we focus on the search of the cLFV $\tau^\pm \to \ell^\pm + \alpha$ decay, where ℓ is either an electron or a muon, and α is a light, BSM boson that is not directly detectable (invisible). This cLFV process is predicted in new physics models, e.g., involving axion-like particles [3, 17–19]. Moreover, in such models, the introduction of these new particles aims to explain the discrepancies of the anomalous magnetic moment of the muon, a_μ, and they are potential dark matter candidates. Here, we present the latest results on the branching-fraction ratio $\mathcal{B}(\tau^\pm \to \ell^\pm + \alpha)/\mathcal{B}(\tau^\pm \to \ell^\pm + \nu\bar{\nu})$ using 62.8 fb^{-1} of data collected with the Belle II detector [12].

In addition, we review the $\tau^\pm \to \ell^\pm + \gamma$ decay, where ℓ is either an electron or a muon. This process is a golden channel for studying cLFV, since it has the largest branching ratio predicted by new physics models where LFV is induced by one-loop diagrams including heavy

particles, such as the ones proposed by supersymmetric models [20,21]. However, the search for this decay represents a challenge since it is affected by irreducible backgrounds from Bhabha and $\tau^{\pm} \to \ell^{\pm} + \nu\bar{\nu}$ processes.

During the last decade, the Belle and BaBar collaborations led the search for the $\tau^{\pm} \to \ell^{\pm}+\gamma$ process, being the latest result presented by the Belle collaboration in 2021 with an integrated luminosity of 988 fb^{-1} [4, 5]. Here, we present the sensitivity analysis on the branching ratio $\mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \gamma)$ for an integrated luminosity of 400 fb^{-1}. This luminosity is almost two times lower than the Belle study; on the other hand, it is the luminosity that the Belle II experiment has collected so far. Currently, the operations of the Belle II detector are stopped due to the Long Shutdown 1 (LS1). The operations and data-taking will resume in October 2023. We hope to achieve a larger integrated luminosity to perform a more precise study in the coming years. In the meantime, we are working on improving and developing new techniques to reach better bounds even with less luminosity.

In Chapter 2 the physics motivation is presented. In Chapter 3 a brief introduction to the Belle II experiment is provided. The cLFV searches are presented in Chapters 4 and 5. Results are summarized in Chapter 6 and prospects are listed in Chapter 7.

It is worth mentioning that the search for the decay $\tau^{\pm} \to \ell^{\pm} + \alpha$ at Belle II, as presented in Chapter 4, was conducted in collaboration with professors Ivan Heredia, Eduard De La Cruz, Armine Rostomyan, postdoctoral researchers Petar Rados, Francesco Tenchini, and doctoral students Alejandro De Yta [22] and Thomas Kraetzschmar [23]. This research has been documented in an internal note of the Belle II experiment [11] and published in Physical Review Letters [12].

Chapter 2

Physics motivation

2.1 The Standard Model of particle physics

The Standard Model of particle physics (SM) is the theoretical model that describes the behavior and interactions of fundamental particles. The SM is a gauge theory based on the symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, where $SU(3)_C$ describes the strong nuclear force, $SU(2)_L$ the weak nuclear force, and $U(1)_Y$ the electromagnetic force. The gauge bosons associated with these symmetries are the gluons, the W and Z bosons, and the photon, respectively.

The SM consists of leptons and quarks, the force mediators gauge bosons, and the Higgs boson, which is responsible for giving mass to all the particles throughout the Higgs mechanism. The summary table of the SM is represented in Figure 2.1.

The SM is the most successful theory describing the behavior of fundamental particles (tested up to the TeV scale). However, it is not a complete theory, as many questions remain unanswered, such as the nature of dark matter and dark energy, the matter-antimatter asymmetry in the universe, the existence of three generations of fermions, and the hierarchy of fermion masses. The question of whether the Higgs boson is the only scalar particle is also of interest. Hence, there is motivation to search for undiscovered particles or mechanisms that help us to address these questions.

2.2 Lepton flavor violation

In the SM there are three generations of fermions. In particular, for leptons, we have the electron (e), muon (μ), and tau (τ), with their respective neutrinos ν_e, ν_μ, and ν_τ. Each of these leptons have an associated lepton flavor, that is, the electron, the muon, and the tau flavor. This quantity is 1 for particles and -1 for antiparticles. Additionally, we can define the lepton number as the difference between the leptons (n_l) and the anti-leptons (n_{-l}), $L = n_l - n_{-l}$. In the SM, lepton and flavor numbers are conserved.

In (recent) extensions of the SM, where neutrino masses are considered, Lepton Flavor Violation (LFV) takes place due to neutrino oscillations. However, in the charged sector, these processes have not been observed. If they were observed, they would lead to new physics scenarios, since they are strongly suppressed considering only neutrino oscillations.

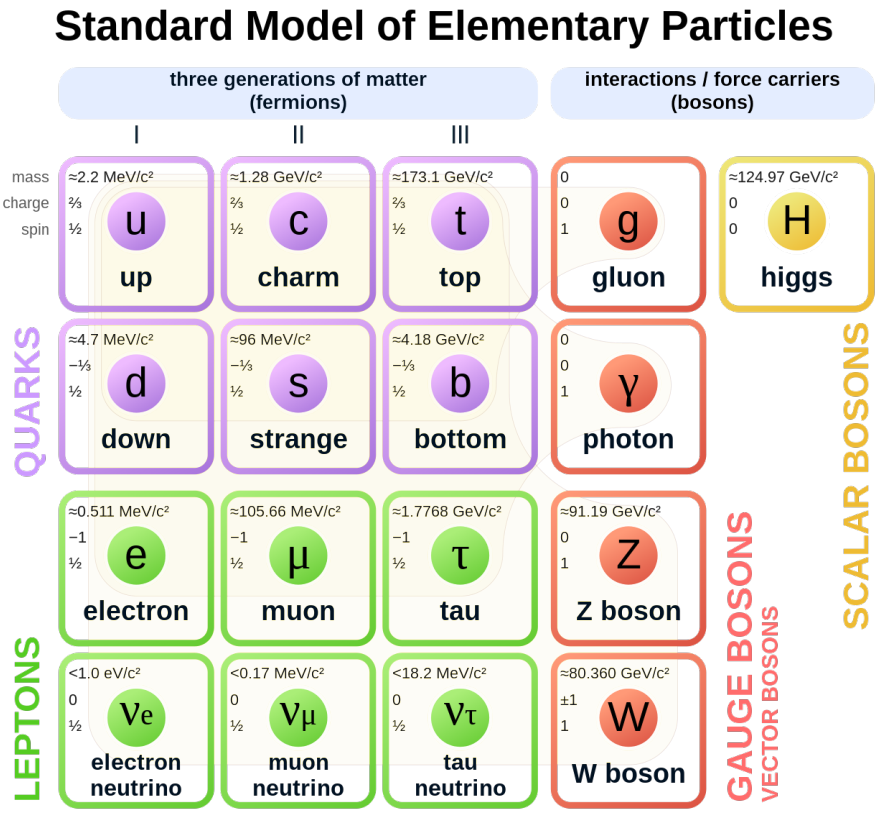


Figure 2.1: Overview of the Standard Model of particle physics. The Standard Model includes 12 fermions (leptons and quarks) grouped in three families and 5 gauge bosons which are the force mediators. The Higgs boson gives mass to all particles [1].

2.3 Search of the decay $\tau^{\pm} \to \ell^{\pm} + \alpha$

Several studies in the search for LFV processes have taken place in electron-positron colliders. In particular, the MARK III (1993) and ARGUS (1995) collaborations searched for $\tau^{\pm} \to \ell^{\pm} + \alpha$ [24, 25]. It was the ARGUS paper that established a novel methodology to search for $\tau^{\pm} \to \ell^{\pm} + \alpha$ with respect to the SM process $\tau^{\pm} \to \ell^{\pm} + \nu\bar{\nu}$. As no signal was found, ARGUS set an Upper Limit (UL) on the orders of $(3 - 36) \times 10^{-3}$ (depending on the α mass) at 90% confidence level (C.L.) for $\mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \alpha / \mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \nu\bar{\nu})$, where ℓ is a muon or an electron.

2.3.1 Searching technique

In this work, we use the ARGUS technique [25] to search for $\tau^{\pm} \to \ell^{\pm} + \alpha$ at Belle II, which considers a 3x1 prong topology, that is, one tau decaying as $\tau^{\pm} \to \ell^{\pm} + \alpha$ (signal decay) and the other as $\tau^{\mp} \to h^{\mp}h^{-}h^{+}\bar{\nu_{\tau}}$ (tag decay).

For the two-body decay $\tau^{\pm} \to \ell^{\pm} + \alpha$ in the τ rest frame, the lepton momentum spectrum is shaped as a monochromatic peak at a value depending on the α mass. Therefore, we expect a signal peak on top of an irreducible SM background, mainly from the decay $\tau^{\pm} \to \ell^{\pm} + \nu\bar{\nu}$. In order to boost the lepton into the τ rest frame, we must know the flight direction of the τ lepton. Unfortunately, due to the missing particles involved, the flight direction can not be accurately reconstructed. It is, however, possible to approximate the momentum of the τ signal

using information from the tag decay and boost the lepton into a τ *pseudo-rest frame*.

Following the ARGUS method, the direction of the signal τ lepton is approximated to the direction of the total momentum of the 3-prong system in the center-of-mass system (CMS),

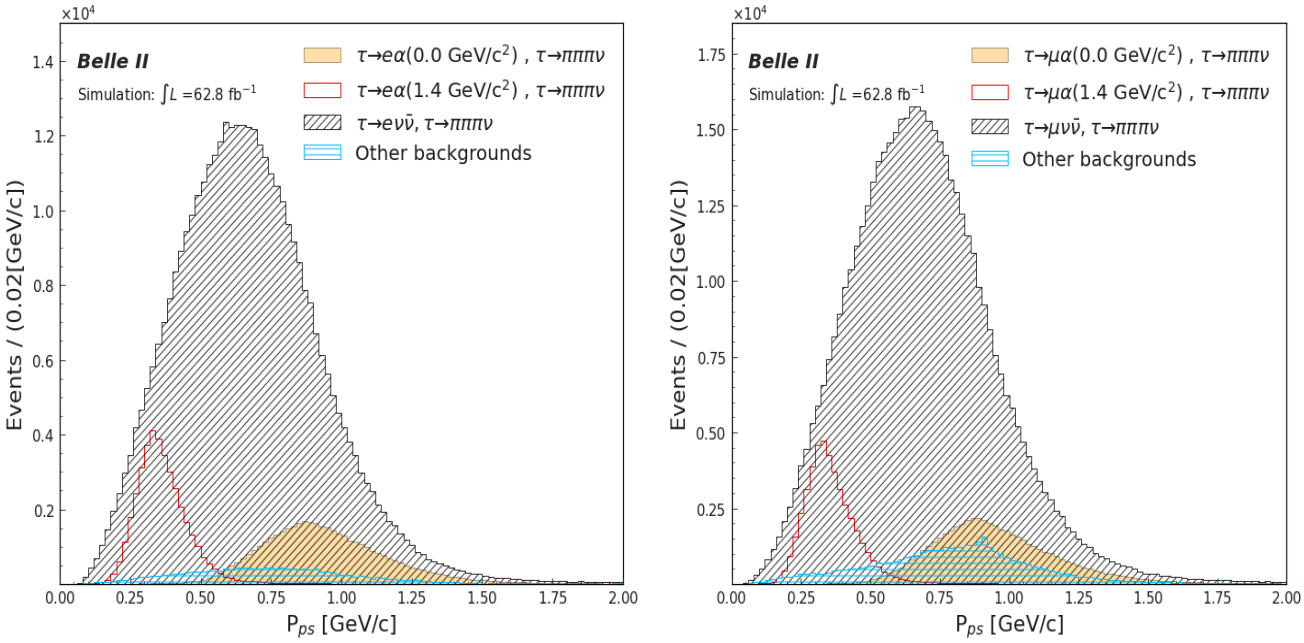
$$\hat{p}_{\tau} \approx -\frac{\vec{p}_{3h}}{|\vec{p}_{3h}|} \ , \tag{2.1}$$

while the τ energy (up to initial state radiative corrections) is fixed to the beam energy E_{beam},

$$E_{\tau} = E_{\text{beam}} = \sqrt{s}/2 \ . \tag{2.2}$$

With these approximations a transformation to a τ pseudo-rest frame becomes possible. Figure 2.2 shows the lepton momentum distributions for $\tau^{\pm} \to \ell^{\pm} + \nu\bar{\nu}$ and $\tau^{\pm} \to \ell^{\pm} + \alpha$ decays in the τ pseudo-rest frame. We set the branching ratio of $\tau^{\pm} \to \ell^{\pm} + \alpha$ to 0.1 in these plots.

To measure $\mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \alpha)$, we use different background suppression and confidence interval calculation approaches. The details are given in Chapter 4.



(a) Electron channel.

(b) Muon channel.

Figure 2.2: Lepton momentum distributions in the τ pseudo-rest frame calculated with the ARGUS method. Examples of the $\tau^{\pm} \to \ell^{\pm} + \alpha$ channel for masses equal to 0 and 1.4 GeV/c^2 are shown on the plot. The $\mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \alpha)$ is set to 0.1 for both α masses [2].

Spin

Resulting from the conservation of the angular momentum, the α particle in the $\tau^{\pm} \to \ell^{\pm} + \alpha$ decay should be a boson. For this search, we assume a spin-0 α boson.

Indirect but tighter constraints on the existence of a light spin-1 α boson that decays beyond the detector (long-lived) can be obtained from LFV decays to 3 leptons, $\mu^{\pm} \to e^{\pm}e^{+}e^{-}$ or $\tau^{\pm} \to \ell^{\pm}\ell^{+}\ell^{-}$, when the α is off-shell [26]. In any case, our search technique relies on the lepton momentum, that is spin-independent.

2.3.2 New physics models

The $\tau^{\pm} \to \ell^{\pm} + \alpha$ process is predicted in new physics models, e.g., involving axion-like particles [3, 17–19]. Axion-like particle (ALP) models introduce new, light, and weakly interacting particles (α) that share some of the properties of the axion. The axion is a hypothetical particle proposed in the 1970s to solve the Charge-Parity (CP) problem in Quantum Chromodynamics (QCD). In some models, these weakly interacting particles can cause LFV. Moreover, in such models, the introduction of these new particles aims to explain the discrepancies of the anomalous magnetic moment of the muon, a_μ, and they have the potential to be dark matter candidates. Some representative models of ALPs with LFV couplings are the LFV QCD axion, the LFV axiflavon, the leptonic familon, and the majoron.

The QCD axion model provides a mechanism (via the spontaneous breaking of a U(1) Peccei-Quinn (PQ) symmetry) for CP symmetry to be preserved in the strong force. It also introduces a new coupling between the QCD axion and the lepton Yukawa couplings, which allows the axion to interact with leptons in a way that violates lepton flavor. The LFV axiflavon is a realization of the QCD axion model, where the PQ symmetry is also responsible for explaining the smallness of the SM Yukawas and the QCD axion is the axiflavon with flavor-violating couplings [3, 27].

The leptonic familon is a (pseudo-) Nambu-Goldstone boson (PNGB) arising from the spontaneous breaking of a global horizontal symmetry, which can be the Froggatt-Nielsen (FN) flavor symmetry $\mathrm{U}(1)_{FN}$, where $\mathrm{U}(1)_{FN}$ only acts in the leptonic sector. In the LFV familon setup the strengths of the LFV couplings depend on the texture of the PMNS matrix [3, 28].

The majoron is the PNGB due to the spontaneous breaking of the lepton number. A natural context where this kind of ALP arises are the seesaw models, where the breaking scale of the lepton number is associated with the mass scale of heavy new fields. In a non-minimal class of seesaw models, the majoron has parametrically enhanced LFV couplings [3, 29, 30]. The status and future projections of the LFV ALP searches are summarized in Figure 2.3.

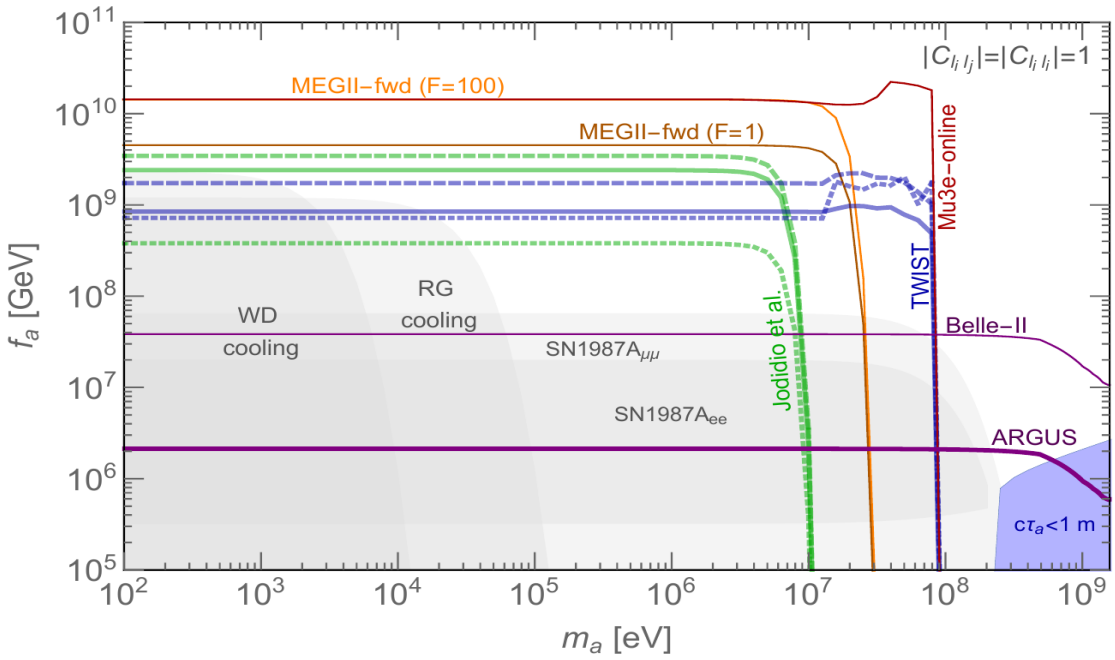


Figure 2.3: Summary of the present bounds and future projections on the decay constant f_a vs. m_a, for an ALP with generic couplings to leptons, where $\mathcal{B} \propto 1/f_\alpha$. The projection for the Belle II experiment is shown considering the total integrated luminosity of 50 ab^{-1} [3].

2.4 Search of the decay $\tau^{\pm} \to \ell^{\pm} + \gamma$

The $\tau^{\pm} \to \ell^{\pm} + \gamma$ process can occur in extensions of the SM, where neutrinos masses are considered. In this case, the LFV is allowed via neutrino oscillations, illustrated with a Feynman diagram in Figure 2.4.

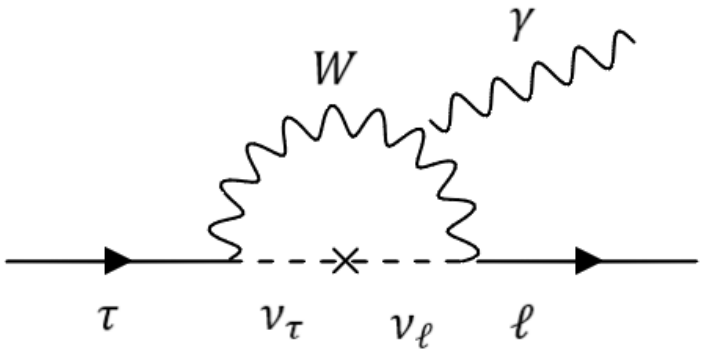


Figure 2.4: Feyman diagram of the $\tau \to \ell + \gamma$ process, where ℓ stands for muon or electron channels.

However, this process is strongly suppressed since its branching ratio is proportional to $(\frac{\Delta m_\nu}{m_W})^4$ due to the Glashow–Iliopolous–Maiani (GIM) mechanism [15, 31], where the mass difference of the neutrinos Δm^2_ν is $\sim 10^{-3}$ eV^2 and m_W is the W boson mass.

$$\mathcal{B}(\tau^{\pm} \to \ell^{\pm}\gamma) = \frac{3\alpha}{32\pi}|U^*_{\tau i}U_{\ell i}\frac{\Delta m^2_{3i}}{m^2_W}|^2 \sim 10^{-54} \tag{2.3}$$

If we were able to see any signature of LFV, this will need to be explained by new physics scenarios.

2.4.1 New physics models

There are several new physics (NP) models where the mass of the neutrinos is considered and LFV is enhanced to levels that are reachable for new experiments, such as Belle II. Some of these models are listed in Table 2.1.

Model	$\tau^{\pm} \to \mu^{\pm} + \gamma$	Reference
SUSY + Higgs	10^{-10}	[20]
Little Higgs	10^{-10}	[32]
SM + Seesaw	10^{-9}	[33]
Non-universal Z'	10^{-9}	[34]
SUSY SO(10)	10^{-8}	[21, 35]

Table 2.1: Predictions for different new physics models of the $\tau^{\pm} \to \mu^{\pm} + \gamma$ branching ratio.

In general, LFV is predicted by the following mechanisms:

- SUSY models introduce new particles, known as superpartners, that are heavier and more massive than their SM counterparts. These particles can interact with leptons and cause LFV.

- The Seesaw Mechanism can explain the small masses of neutrinos by introducing heavy right-handed neutrinos. These right-handed neutrinos can interact with leptons and cause LFV.

- In the Z$'$ model, the presence of the new gauge boson, denoted as Z$'$, can lead to various new physics phenomena, including the possibility of LFV. In this case, LFV arises when the Z$'$ boson couples differently to different lepton flavors.

- In the Little Higgs model with T-parity. The LFV arises from the mixing between the SM leptons and their T-parity leptons partners (with masses of order 1 TeV).

2.4.2 Searching techniques

During the last decade, the Belle and BaBar collaborations led the search for the golden decay $\tau^{\pm} \to \ell^{\pm} + \gamma$. The best UL for the electron channel was presented by Babar in 2009 [4], and the muon channel by Belle in 2021 [5]. To give an idea of these limits, the results for both collaborations are shown in Table 2.2.

Experiment	Integrated luminosity (fb^{-1})	Channel	UL @ 90 C.L.	Reference
Belle 2021	988	$\tau^{\pm} \to \mu^{\pm} + \gamma$	4.2×10^{-8}	[5]
		$\tau^{\pm} \to e^{\pm} + \gamma$	5.6×10^{-8}	
Babar 2009	516	$\tau^{\pm} \to \mu^{\pm} + \gamma$	4.4×10^{-8}	[4]
		$\tau^{\pm} \to e^{\pm} + \gamma$	3.3×10^{-8}	

Table 2.2: The most recent UL results at 90 % of C.L. presented by the Babar and Belle collaborations.

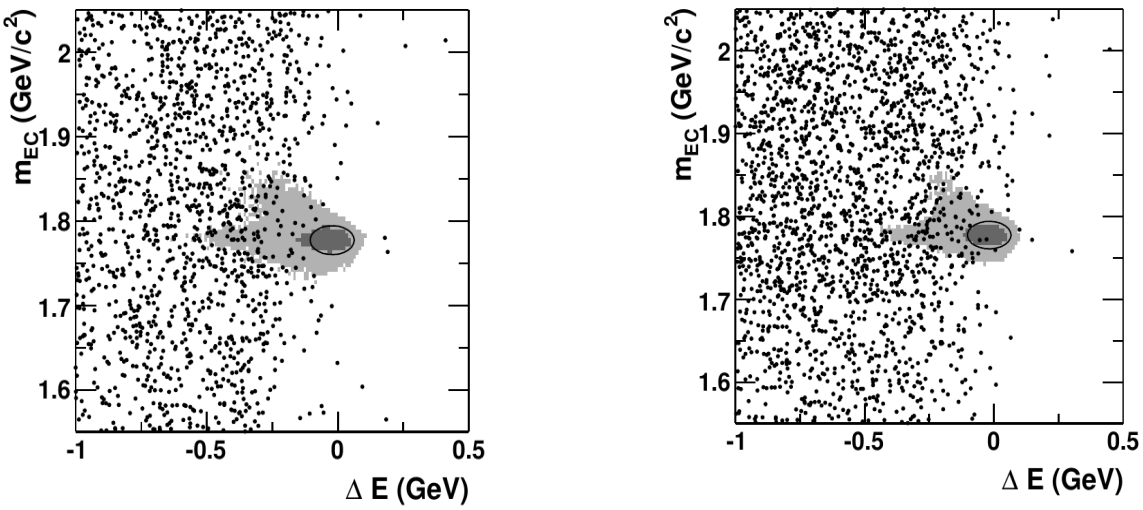
In these searches, the signal region is defined using the beam-constrained mass

$$M_{bc} = \sqrt{(E_{beam}^{CMS})^2 + (P_{\ell\gamma}^{CMS})^2}\,, \tag{2.4}$$

and the energy difference:

$$\Delta E = E_{\ell\gamma}^{CMS} - E_{beam}\,. \tag{2.5}$$

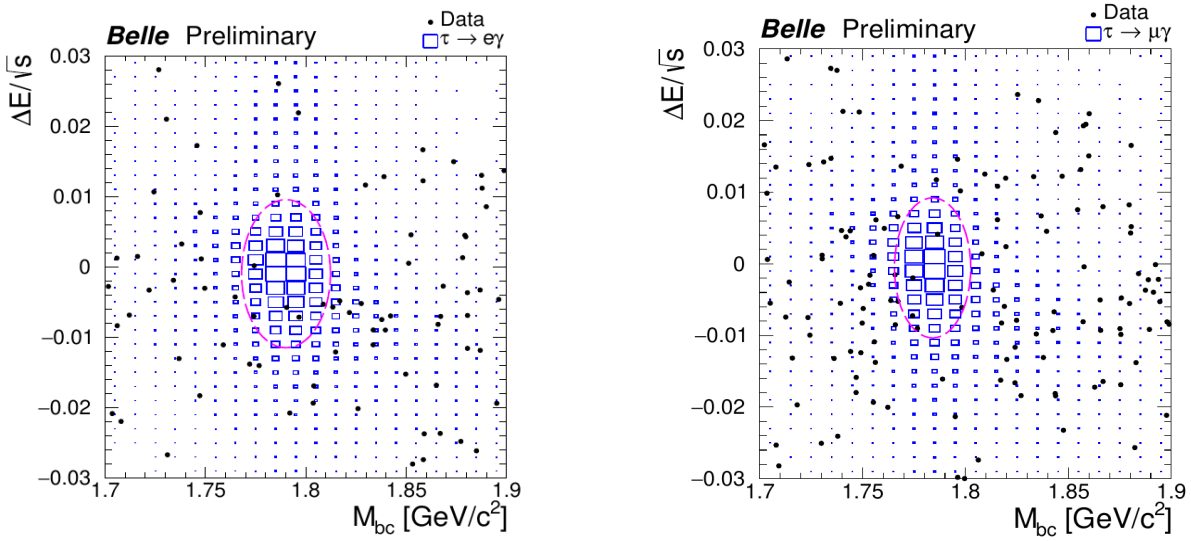
Babar in 2009 used the $M_{bc} - \Delta E$ plane and Belle in 2021 used the normalized energy difference instead, $M_{bc} - \Delta E/\sqrt{s}$ (in previous analyses from Belle, the invariant mass $M_{inv} = \sqrt{(E_{\ell\gamma})^2 + (P_{\ell\gamma})^2}$ was used instead of M_{bc}). The 2-D distributions for the electron and muon channels can be seen in Figure 2.5 and Figure 2.6, respectively.



(a) Electron channel.

(b) Muon channel.

Figure 2.5: BaBar $\tau^{\pm} \to \ell^{\pm} + \gamma$ signal regions for the electron (left) and muon (right) channels. The $x-$axis represents the energy difference of the τ and the beam energy in the center of mass system (CMS). The beam-constrained mass is displayed on the $y-$axis [4].



(a) Electron channel.

(b) Muon channel.

Figure 2.6: Belle $\tau^{\pm} \to \ell^{\pm} + \gamma$ signal regions for the electron (left) and muon (right) channels. The $x-$axis represents the normalized energy difference of the τ and the beam energy in the center of mass system (CMS). The beam-constrained mass is displayed on the $y-$axis [5].

Since the Belle 2021 analysis [5] is the most recent update, we summarize the distinctive points and improvements below:

- A total integrated luminosity of 988 fb^{-1} (912×10^6 τ pairs) was used.
- A calibration of the photon energy was performed for the first time using radiative $e^{\pm}e^{\mp} \to \mu^{\pm}\mu^{\mp}(\gamma)$ events.
- To define the signal region, M_{bc} is used instead of M_{inv}, as M_{bc} is less sensitive to the photon energy calibration.
- The τ tag side was reconstructed explicitly for the following 1-prong possibilities: $\tau^{\pm} \to e^{\pm} + \nu\bar{\nu}$, $\tau^{\pm} \to \mu^{\pm} + \nu\bar{\nu}$, $\tau^{\pm} \to \pi^{\pm} + \nu$, and $\tau^{\pm} \to \rho^{\pm} + \nu$.
- The energy asymmetry between the lepton and the photon in the signal side was introduced to improve the analysis: $|E_{\ell}^{CMS} - E_{\gamma}^{CMS}|/(E_{\ell}^{CMS} + E_{\gamma}^{CMS})$.
- The cosine of the angle between the τ and the track on the tag side was also used for the first time. Since the signal side is fully reconstructed, the momentum of the tag track can be written as $\vec{p}_{\tau(tag)} = -\vec{p}_{\tau(sig)} = -\vec{p}_{\ell} - \vec{p}_{\gamma}$. Then, the cosine of the τ and the track in the tag side is calculated as

$$\cos\theta_{\tau(tag),track(tag)} = \frac{\vec{p}_{\tau(tag)} \cdot \vec{p}_{track(tag)}}{|\vec{p}_{\tau(tag)}||\vec{p}_{track(tag)}|} \ . \tag{2.6}$$

  This variable was used to remove $\tau^{\pm}\tau^{\mp}$ events as can be seen in Figure 2.7.
- The final signal region was selected using an elliptical region given by

$$\frac{(M_{bc} - \mu M_{bc})^2}{2\sigma_{M_{bc}}} + \frac{(\Delta E/\sqrt{s} - \mu\Delta E/\sqrt{s})^2}{2\sigma_{\Delta E/\sqrt{s}}} < 1.0 \tag{2.7}$$

  where $\sigma_{M_{bc}}$ and $\sigma_{\Delta E/\sqrt{s}}$ were obtained from a fit of an asymmetric Gaussian to a specific region on the M_{bc} and $\Delta E/\sqrt{s}$ variables, as can be seen in Figure 2.8.
- The signal efficiency in the signal region is 3.7% for $\tau^{\pm} \to \mu^{\pm} + \gamma$ events and 2.9 % for $\tau^{\pm} \to e^{\pm} + \gamma$.
- Regarding the statistical model, for the background estimation, MC simulations were used (except for $e^{\pm}e^{\mp}(\gamma)$). Its probability density function (PDF) was modeled as

$$F(M_{bc}, \Delta E/\sqrt{s}) = f(M_{bc}) \times g(\Delta E/\sqrt{s}) \ , \tag{2.8}$$

  where $f(M_{bc})$ is a constant function, and $g(\Delta E/\sqrt{s})$ a sum of a Landau and an exponential function for $\tau^{\pm}\tau^{\mp}$ events, and a sum of a Gaussian and a Landau function for $\mu^{\pm}\mu^{\mp}(\gamma)$ and $e^{\pm}e^{\mp}(\gamma)$ events. For signal, the PDF was obtained by smoothening the corresponding MC distribution.
- The following systematic uncertainties were considered: track reconstruction efficiency, photon reconstruction efficiency, luminosity, PID efficiency, trigger efficiency, photon energy calibration, background PDF model, and MC statistics.

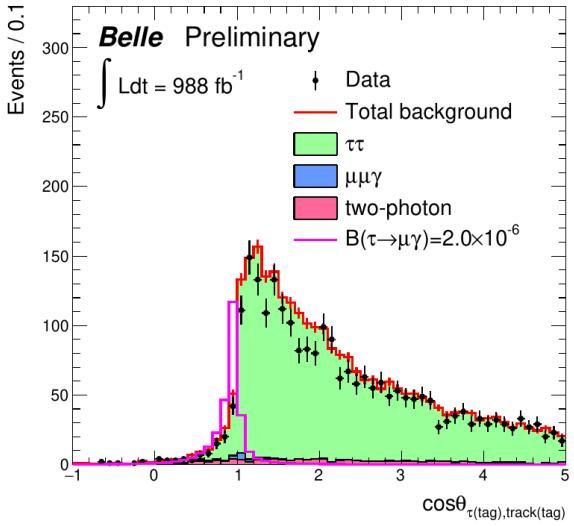


Figure 2.7: Distribution of signal and background components on the cosine of the angle between the τ and the track in the tag side, described in Equation 2.6. The green-filled histogram shows that the $\tau^{\pm}\tau^{\pm}$ background is located at larger values than the $\tau^{\pm} \to \ell^{\pm} + \gamma$ channel [5].

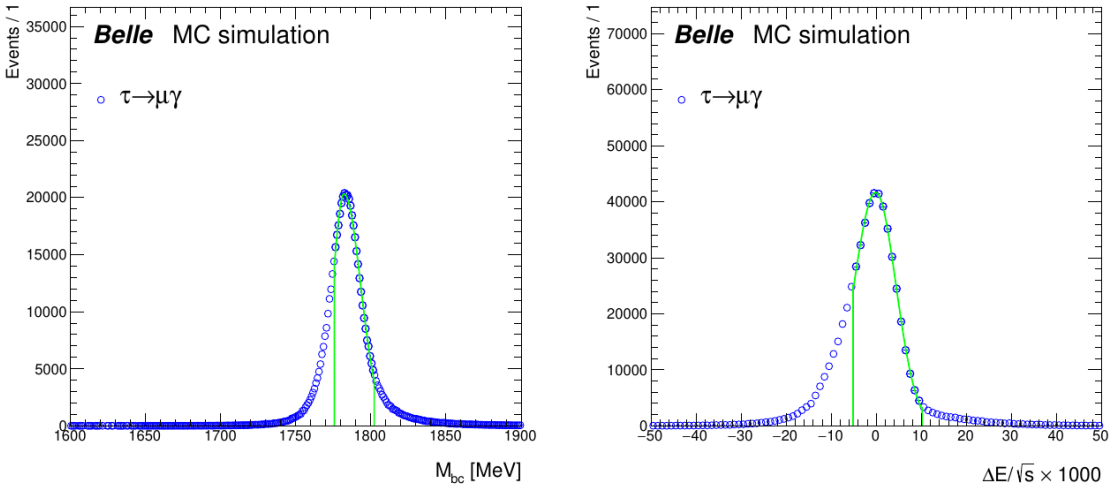


Figure 2.8: Fit to simulation on the beam-constrained mass and energy difference variables. The distributions are modeled using an asymmetric Gaussian, and the green lines define the region used for the fit [5].

Chapter 3

Belle II experiment

The Belle II experiment aims to solve unanswered fundamental questions, e.g. regarding the matter composition of the universe, and explore new phenomena, beyond the Standard Model of particle physics, by using the intensity frontier approach[1].

The Belle II experiment, instrumented at the SuperKEKB is illustrated in Figure 3.1. It is located at the High Energy Accelerator Research Organization (KEK) in Tsukuba, Japan. In Belle II, electron-position asymmetric beams collide at a center of mass energy of $\sqrt{s} = m_{\Upsilon(4S)} \approx 10.58$ GeV, where the production of B meson pairs is enhanced, which is the reason why the facility is commonly known as a super B-factory.

The Belle II experiment is a major upgrade to the Belle experiment. Belle II is designed to reach a target luminosity of 6.5×10^{35} cm^{-2} s^{-1}, that is, 30 times more than Belle, and a total integrated luminosity of 50 ab^{-1}, which is 50 times more data than the Belle experiment [7]. One of the features of the high luminosity implementation of Belle II is the narrow beams at the interaction point, which are achieved by nanobeam technology [36].

The Belle II collaboration currently consists of more than 1160 physicists and engineers from 1130 organizations in more than 27 countries, including Mexico.

3.1 SuperKEKB accelerator

The SuperKEKB accelerator is a circular collider of 3 km circumference, built 10 meters underground. The main components are the High Energy Ring (HER), Low Energy Ring (LER), and an injector linear accelerator. It collides electrons and positions at an energy of 7 and 4 GeV, respectively, at the interaction point, where the Belle II detector is located.

Presently, the operation of SuperKEKB is stopped (since summer 2022). A Long Shutdown 1 (LS1) was scheduled to replace the Vertex Detector (VXD). The instantaneous luminosity reached before LS1 was 4.7×10^{34} cm^{-2} s^{-1}, which is the world record for instantaneous luminosity. The collected integrated luminosity before LS1 is 424 ± 3 fb^{-1}. SuperKEKB will resume operations in October 2023.

[1]The intensity frontier approach in particle physics focuses on creating high-intensity particle beams and studying the rare processes that occur in loops when these beams interact. On the other hand, the energy frontier approach focuses on accelerating particles to higher and higher energies, creating collisions with higher center-of-mass energies.

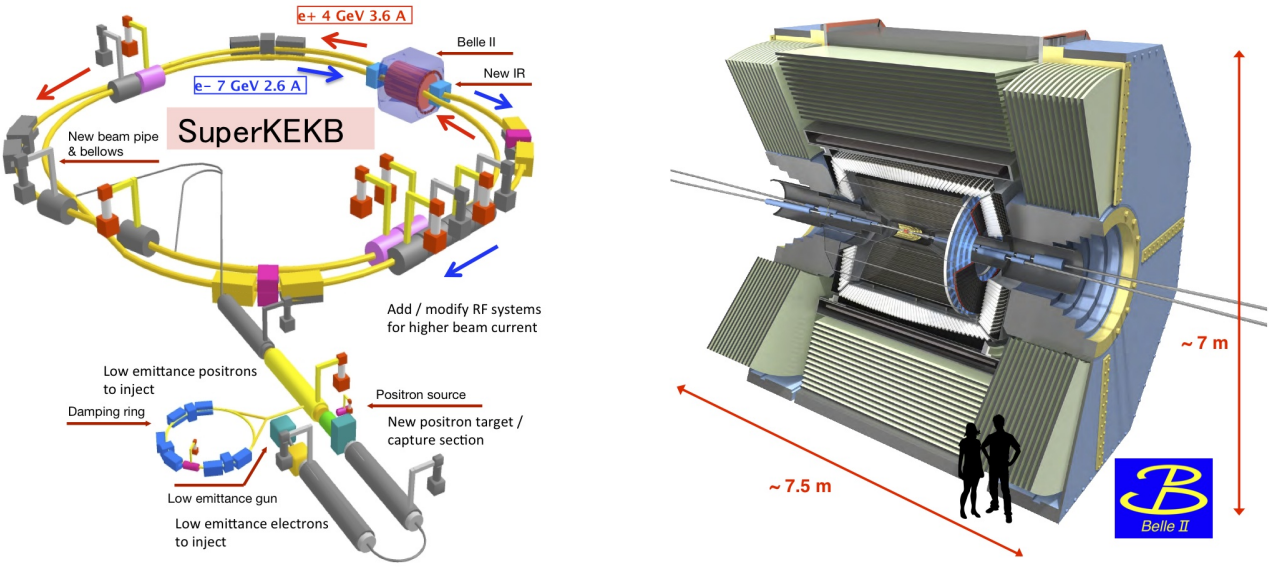


Figure 3.1: Schematic visualization of the SuperKEKB collider and Belle II detector [6].

3.2 Belle II detector

The Belle II detector is 7×7.5 m^2 size and weighs 1,400 tons. It consists of several subdetectors arranged in cylindrical layers around the interaction point. The main components of the Belle II detector are illustrated in Figure 3.2 and explained briefly in the following sections.

3.2.1 Tracking system

The tracking system consists of the Vertex Detector (VXD) and the Central Drift Chamber (CDC). The tracking system is responsible to reconstruct tracks and vertices.

Vertex detector

The vertex detector is comprised of the Pixel Detector (PXD) and the Strip Vertex Detector (SVD), altogether with six layers located around a 10 mm radius beam pipe, as can be seen in Figure 3.3.

The first two layers located at $r = 14$ mm and $r = 22$ mm constitute the PXD. The PXD detector uses pixelated DEPFET sensors [7].

The following four layers at radii of 38 mm, 80 mm, 115 mm, and 140 mm constitute the SVD detector. The SVD is equipped with double-sided silicon strip sensors.

The VXD acceptance is $17° < \theta < 150°$ (polar angle measured from the electron beam direction) and its resolution is $\simeq 50$ μm. The VXD reconstructs the vertex of B-meson decays, with lifetimes on the order of a picosecond, and of other short-lived particles [9].

Central Drift Chamber

The Central Drift Chamber (CDC) is a large-volume drift chamber with small drift cells.

The CDC contains, in a radius of 1,130 mm, 14,336 sense wires arranged in 56 layers, either in “axial” orientation (aligned with the solenoidal magnetic field) or “stereo” (skewed with

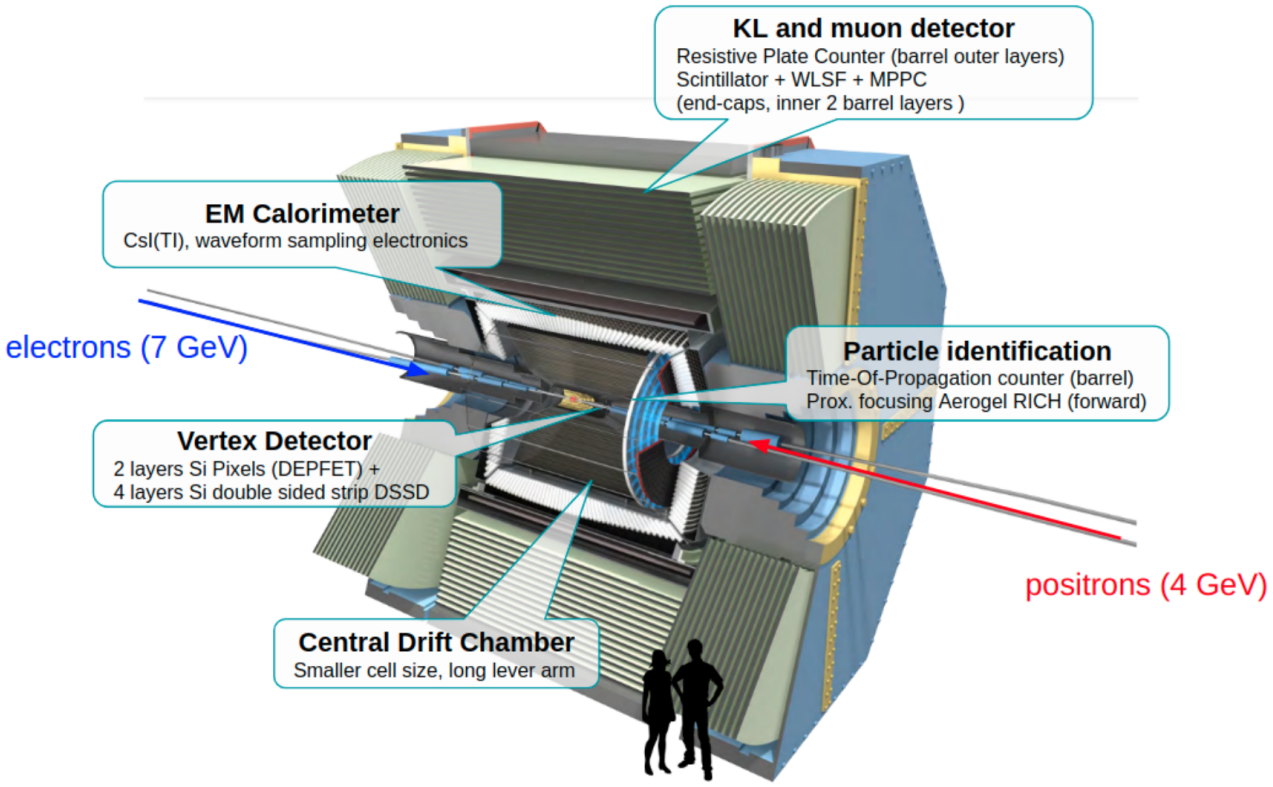


Figure 3.2: Schematic overview of the Belle II principal components.

respect to the axial wires). The drift cells are 17 mm in size and are filled with a He-C_2H_6 50:50 mixture, which results in an average drift velocity of 3.3 cm/μs and a maximum drift time of about 350 ns. By combining information from axial and stereo layers it is possible to reconstruct a full 3D helix track [7]. A schematic view of the CDC design can be seen in Figure 3.4.

The CDC performs precise measurements of the momentum of charged particles, which is very important for the trigger system. The CDC acceptance is $17° < \theta < 150°$ and its spacial resolution is $\simeq$100 μm. Additionally, it provides particle identification information using measurements of energy loss within its gas volume (dE/dx). Moreover, tracks with low momentum that does not reach the particle identification system can be identified using the CDC alone [9].

3.2.2 Particle identification system

The particle identification system is conformed by the Time of Propagation (TOP) counter and the proximity-focusing Aerogel Ring-Imaging Cherenkov (ARICH) detector. The main objective of the particle identification systems is to separate kaons from pions.

TOP

The TOP counter is located in the barrel region. It is a special type of Cherenkov detector where the two-dimensional information of a Cherenkov ring image is given by the time of arrival and impact position of Cherenkov photons at the photo-detector at one end of a long quartz bar. The detector consists of 16 quartz bars of 2.6 m long x 45 cm wide x 2 cm thick size, with

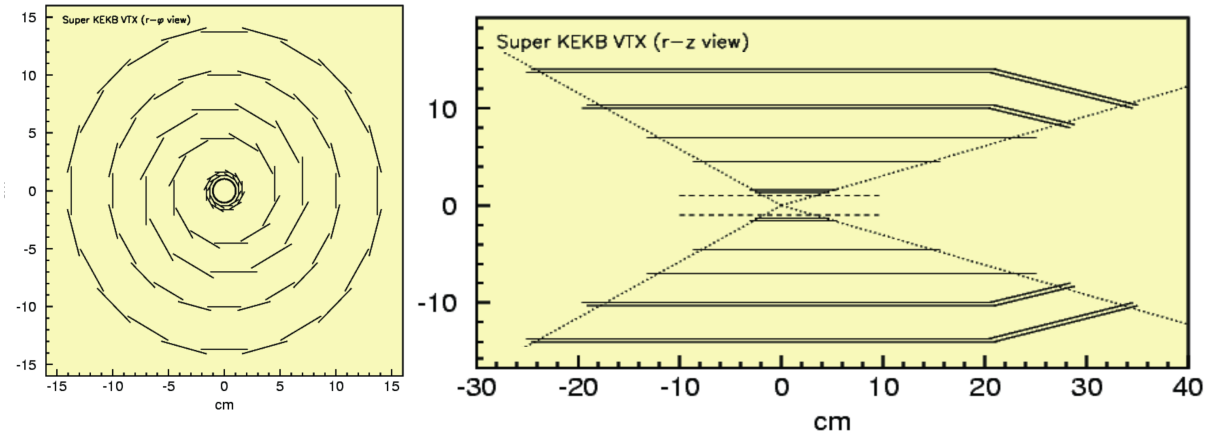


Figure 3.3: Schematic transverse and longitudinal view of the vertex detector [7].

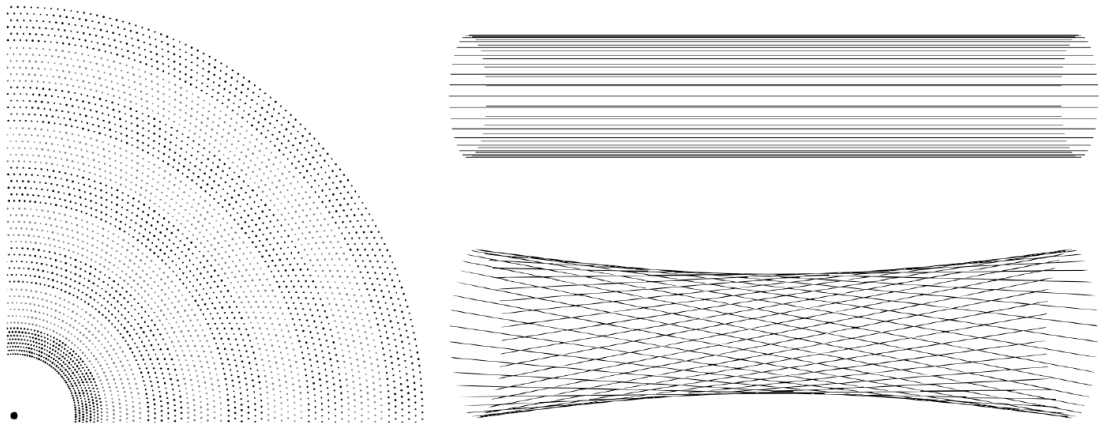


Figure 3.4: Schematic view of a quadrant of a slide of the $r - \phi$ projection of the CDC and a visualization of the stereo (top) and axial (bottom) wires design. In this figure, the skew is exaggerated [8].

an additional volume of 10 cm long where the photo-detector is located. An illustration of the TOP counter can be seen in Figure 3.5.

ARICH

The ARICH detector is located outside the CDC and on the forward end-cap region and consists of two layers of aerogel with different refractive indices, which are $n = 1.045$ upstream and $n = 1.055$ downstream. The different refractive indices are used to increase the yield without degrading the Cherenkov angle resolution. For the detection, a hybrid avalanche photon detector (HAPD) is used. The HAPD is a 73×73 mm^2 size sensor with 144 channels, where photo-electrons are accelerated over a potential difference of 8 kV and are detected in avalanche photodiodes (APD). Figure 3.6 shows the ARCIH detector.

The ARICH detector has been designed to separate kaons from pions over most of their momentum spectrum and to provide discrimination between pions, muons, and electrons below 1 GeV/c [9].

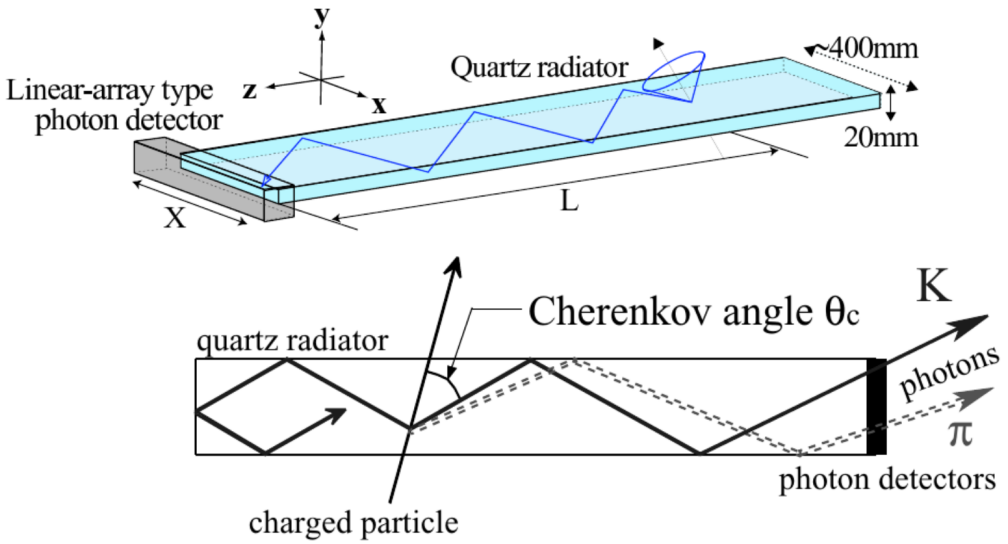


Figure 3.5: Schematic visualization of the TOP counter and its working principle using internal reflecting Cherenkov photons [9].

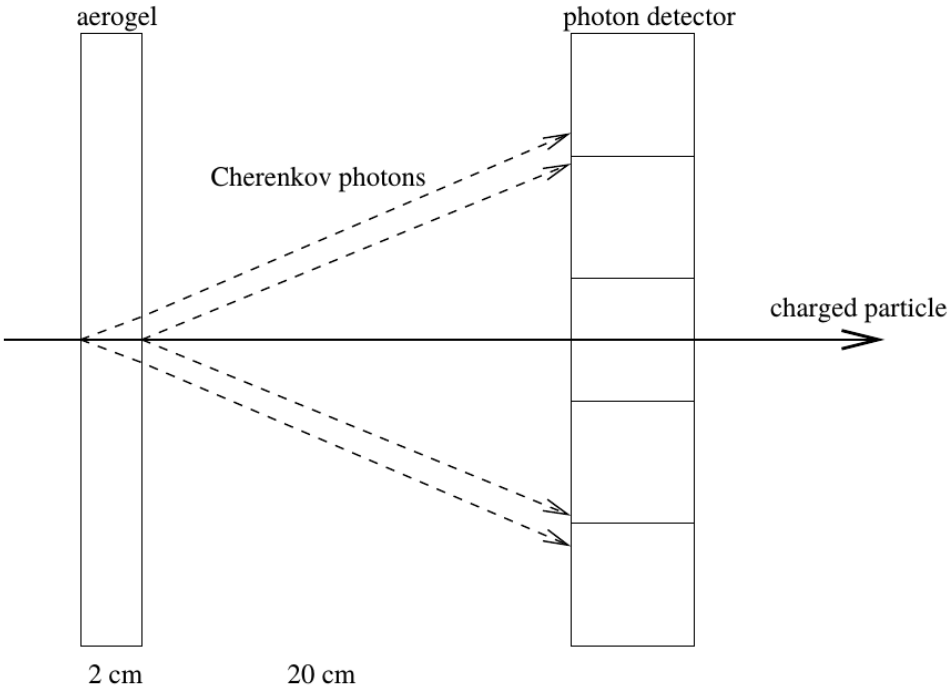


Figure 3.6: Schematic visualization of the working principle of the ARICH detector [9].

3.2.3 Electromagnetic calorimeter

The Electromagnetic Calorimeter (ECL) is located at three detector regions: barrel, forward, and backward end-caps. It consists of a highly-segmented array of 8,736 crystals made of a mixture of thallium-doped caesium iodide CsI(Tl).

The calorimeter covers a polar angle region of $12.4° < \theta < 155.1°$, except for two gaps of about 1° wide between the barrel and endcaps.

The ECL detects photons and electrons with high efficiency. It has a precise determination of the photon energy and angular coordinates. In addition, the ECL is responsible for producing proper signals for the trigger system, providing luminosity measurements, and helping with K_L^0 detection together with the KLM [9]. Both the ECL and the KLM can be visualized in Figure 3.7.

3.2.4 K_L^0 and muon detector

The K_L^0 and muon detector, or KLM, is located outside the superconducting solenoid of 1.5 T. It consists of an alternating design of 4.7 cm thick iron plates and active detector elements (RPCs and scintillators strips). There are 15 detector layers and 14 iron plates in the barrel, and 14 detector layers and 14 iron plates in each endcap.

The octagonal barrel covers the polar angle region of $45° < \theta < 125°$, while the endcaps $20° < \theta < 155°$.

Both the K_L^0 and the muon particle have a large lifetime. Thus, the outer detector KLM can only be reached and penetrated by muons and K_L^0 particles [9].

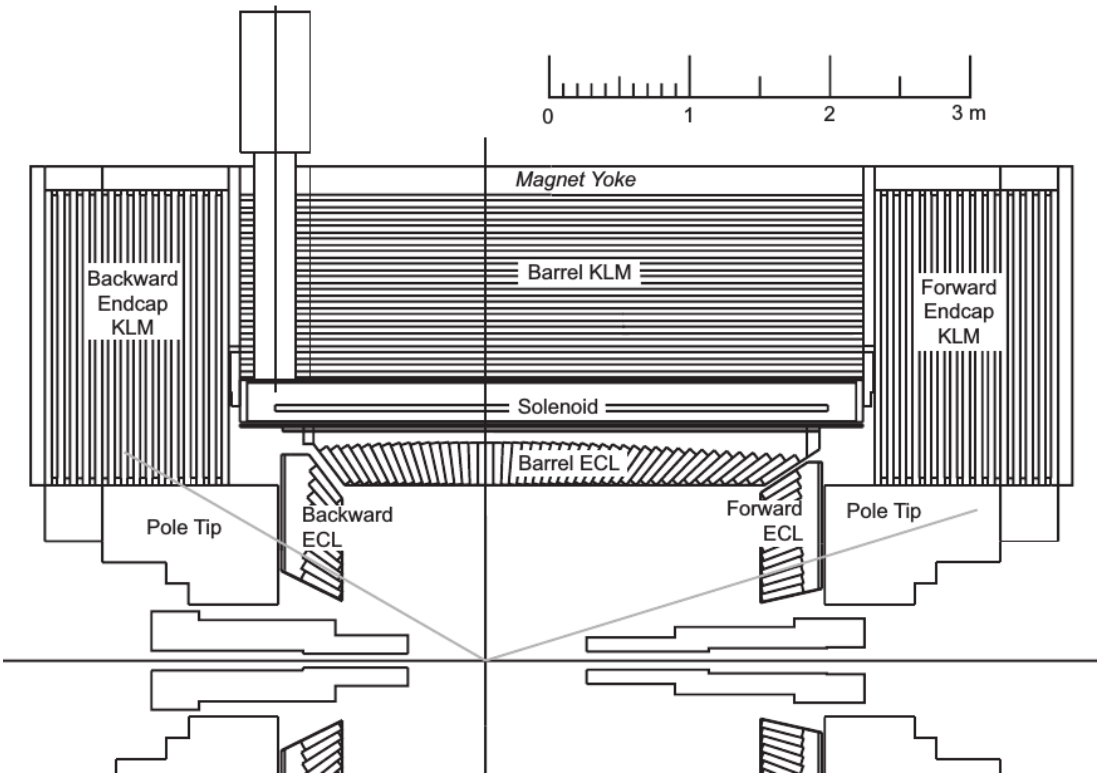


Figure 3.7: Longitudinal design of the Belle II detector, where the ECL and the KLM detector can be visualized [9].

3.3 Trigger and Data Acquisition

The trigger system performs the selection of events of interest during the data taking. It is a two-level system: a hardware based low-level trigger (L1), and a software based high-level trigger (HLT). The nominal L1 trigger has a latency of 5 μs, and a maximum trigger output rate of 30 kHz, limited by the read-in rate of the Data Acquisition (DAQ) system.

Belle II has a wide trigger menu for all the different physics analysis targets. These triggers must work efficiently in the presence of the much higher background rates expected from SuperKEKB (compared to its predecessor, KEKB). The expected beam backgrounds are the Touschek effect, Beam-gas Scattering, Synchrotron radiation, Radiative Bhabha process, two-photon process, and beam-beam effects [7].

The DAQ system is in charge of reading out detector L1 trigger signals. The system transfers the data from the front-end electronics through several steps of data processing, and finally to the storage system.

Moreover, a GRID system is used for data processing, Monte Carlo production, and computing power for physics analysis.

3.4 Belle 2 Simulation

The Belle II simulations consist of two steps: first, the generation of the primary physics process; and second, the detailed Belle II detector simulation (by Geant 4), explained in Section 3.2. The simulations include the beam backgrounds mentioned above and, depending on the analysis, the conditions of the data taking are taken or not into account to a greater or lesser extent.

Cross sections

The cross sections of e^+e^- collisions at $\sqrt{s} = 10.58$ GeV for the simulated backgrounds used in this work are listed in Table 3.1.

Physics process	Cross section (nb)	Generator
$\Upsilon(4S)$	1.110 ± 0.008	KKMC
$u\bar{u}(\gamma)$	1.61	KKMC
$d\bar{d}(\gamma)$	0.40	KKMC
$s\bar{s}(\gamma)$	0.38	KKMC
$c\bar{c}(\gamma)$	1.30	KKMC
$e^+e^-(\gamma)$	$300 \pm 3^*$	BABAYAGA.NLO
$\mu^+\mu^-(\gamma)$	1.148	KKMC
$\tau^+\tau^-(\gamma)$	0.919	KKMC
$e^+e^-e^+e^-$	39.7 ± 0.1	AAFH
$\mu^+\mu^-e^+e^-$	18.9 ± 0.1	AAFH
$e^+e^-p^+p^-$	0.0117	TREPS
$e^+e^-K^+K^-$	0.0798	TREPS
$e^+e^-\pi^+\pi^-$	1.895	TREPS

Table 3.1: Total production cross sections and event generators for various physics processes. *Considering preselection cuts for $e^+e^-(\gamma)$ on the energy and polar angle. [7]

3.5 Belle 2 Analysis Software Framework

The Belle 2 Analysis Software Framework (basf2) is based on C++ and Python modules. It is designed to allow independent processing blocks, called modules, to perform relatively small tasks, which are executed linearly within a defined path. The information between modules is shared using common data storage, as can be seen in Figure 3.8.

Basf2 is an open-source framework that consists of simulation, reconstruction, and analysis functions [37]. We extensively use basf2 to obtain our data and Monte-Carlo (MC) samples

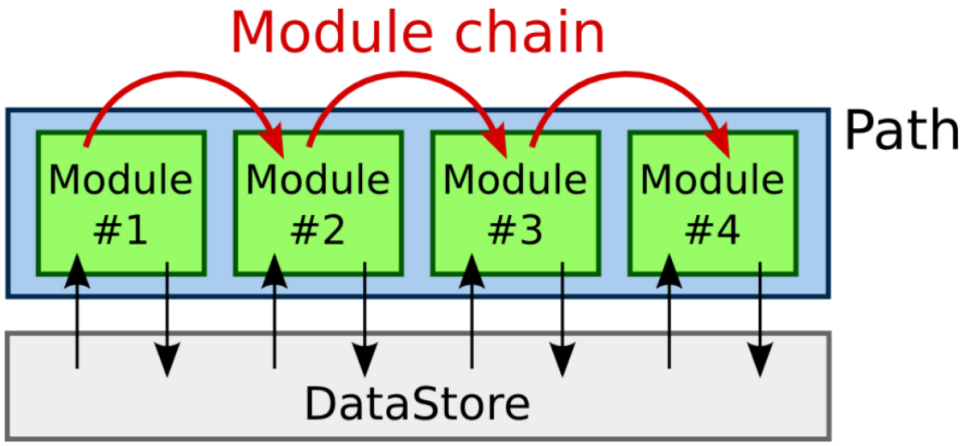


Figure 3.8: Schematic visualization of the Belle 2 Analysis Software Framework workflow [7].

3.6 τ lepton physics at Belle II

In Belle II, the production of B mesons (particles containing a b-quark) is enhanced; however, the production cross section of τ lepton pairs [38], $\sigma(e^+e^- \to \tau^+\tau^-) = 0.919$ nb, is approximately the same as B meson pairs, $\sigma(e^+e^- \to B\bar{B}) = 1.05$ nb. Then, Belle II can also be regarded as a $\tau-$factory. The τ lepton is the heaviest lepton in the SM. It is the only lepton heavy enough to decay into leptons or hadrons [39]. Since the mechanisms that govern its production and weak decay are well-understood, τ lepton physics is extremely convenient to:

- Perform SM precision measurements.
  - Determination of fundamental SM parameters: τ mass, τ lifetime, Michel parameters, etc.
  - Tests of the SM: Lepton Flavor Universality (LFU).
- Search for new physics.
  - Lepton Flavor Violation (LFV).
  - Lepton Number Violation (LNV).
  - Baryon Number Violation (BNV).
  - Charge-Parity Violation (CPV).

Moreover, τ decays at Belle II are produced in low-background conditions and are kinematically constrained.

In Belle II, the target integrated luminosity of 50 ab^{-1} will correspond to 46 billion τ-pairs. This will bring unprecedented precision in τ physics.

In the past, B factories provided the most interesting results in τ lepton physics. Right now, in Belle II, we are working to lead the progress in τ lepton physics.

3.7 Searches of cLFV at Belle II

It is expected that within the next 2-3 years, Belle II will collect the world-largest τ lepton sample. This will allow us to search for new physics with high precision. In particular, in this work, we are focused on cLFV searches.

Since the τ is the most massive lepton, it is possible to search for more than 50 different cLFV processes at Belle II.

In Figure 3.9 the projections for 52 decay channels are presented, for an integrated luminosity of 5 ab^{-1} and 50 ab^{-1}. As a result, we expect to improve the current upper limits by at least one order and up to two orders of magnitude. These projections are calculated based on the limits of Belle analyses, considering the absence of background (assuming precision proportional to $\mathcal{L}^{-1}$) for all the decays, except for $\tau^{\pm} \to \ell^{\pm} + \gamma$ (in this case, a precision proportional to $\mathcal{L}^{-\frac{1}{2}}$ was assumed). For the $\tau^{\pm} \to \ell^{\pm} + \gamma$ decay, it is very difficult to achieve a free background environment without significantly affecting the signal efficiency.

These improvements in the current experimental bounds will be key to constrain new physics models such as those mentioned in Section 2.4.

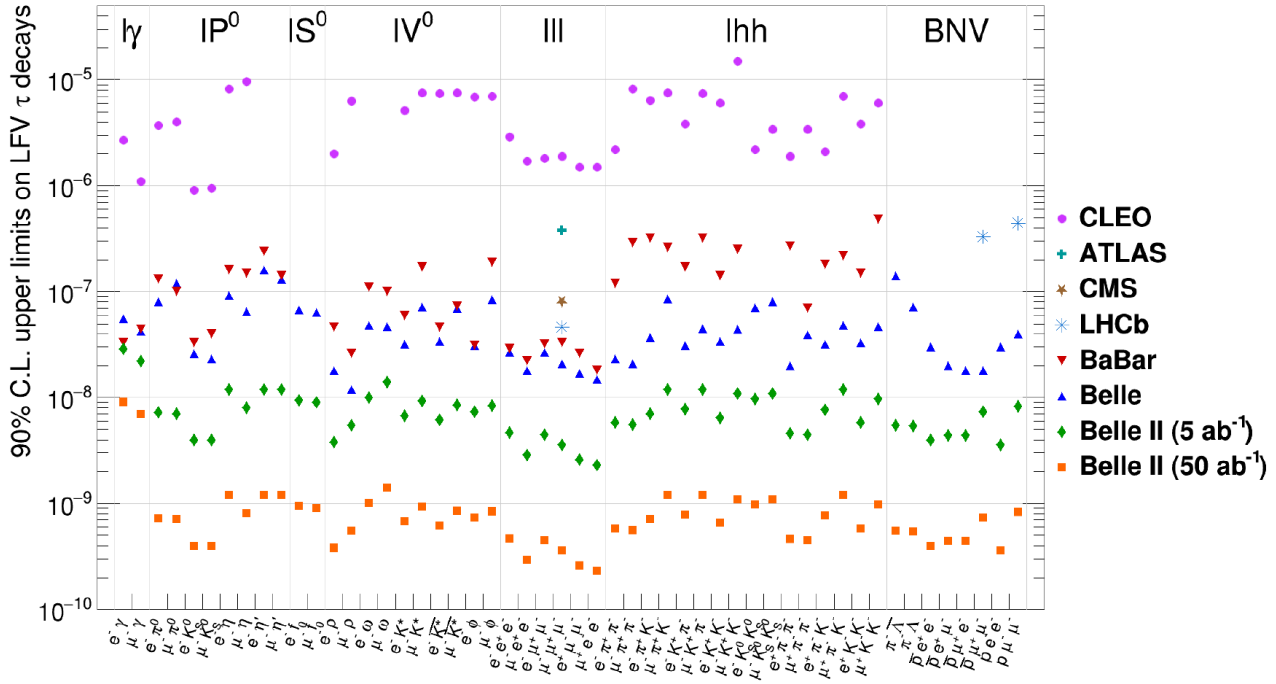


Figure 3.9: Expected Belle II upper limits for τ-LFV decays at 90 % of C.L. Belle II results are estimated considering 5 and 50 ab^{-1} of integrated luminosity, extrapolating Belle limits. A comparison between different experimental results is shown [10].

Chapter 4

Search of the decay $\tau^{\pm} \to \ell^{\pm} + \alpha$ at Belle II

In this chapter, we present the analysis for the search of the LFV $\tau^{\pm} \to \ell^{\pm} + \alpha$, where α is an invisible spin-0 boson.

The $\tau^{\pm} \to \ell^{\pm} + \alpha$ decay is a Lepton Flavor Violating (LFV) process that involves a light beyond the standard model boson that is not directly detectable, i.e., invisible. This decay is predicted in different new physics models as mentioned in Chapter 2.

The last search was performed by the ARGUS collaboration in 1995 [25] using an integrated luminosity of 476 pb^{-1}. The analysis presented in this chapter corresponds to an integrated luminosity of 62.8 fb^{-1} collected with the Belle II experiment. This represents an increase of more than 2 orders of magnitude in integrated luminosity.

4.1 Data and Monte Carlo simulation samples

4.1.1 Monte Carlo simulation

To generate signal Monte Carlo (MC) samples of $\tau^{\pm} \to \ell^{\pm} + \alpha$ decays for different α masses, we use the official (centralized) MC production of the Belle II experiment. We generate 10^7 events for seven different masses M_{α} of the invisible particle, $\text{M}_{\alpha} = 0, 0.5, 0.7, 1.0, 1.2, 1.4$ and 1.6 GeV/c^2. The samples correspond to the "Early Phase 3 MC13a" production with nominal run-independent beam background overlays (`'BGx1'`).

To quantify the background composition after the event reconstruction and selections, we generate generic and low multiplicity samples with the same beam background and data taking conditions as the signal samples. The samples are listed in Table 4.1.

The total integrated luminosity of these samples is 200 fb^{-1} (some samples require an appropriate rescaling). The samples are simulated using the production cross sections listed in Table 3.1.

4.1.2 Data

The data samples were collected by the Belle II experiment from e^+e^- collisions at a center-of-mass energy of 10.58 GeV, during 2019 and 2020.

Different data sets were collected in these years under different circumstances, as the SuperKEKB instantaneous luminosity gradually increased. A summary of the data sets used can be found in Table 4.2. The total integrated luminosity is $\text{L} = 62.8 \pm 0.005$ fb^{-1}.

The MC and data samples were processed without any pre-selection of events (a process named "skim").

Process	Name of production	Scale factor	$\int \mathcal{L}$ dt (fb^{-1})
Generic			
$e^+e^- \to B\bar{B}$	mixed /charged	1	200
$e^+e^- \to \tau^+\tau^-$	taupair	1	200
$e^+e^- \to u\bar{u}$	uubar	1	200
$e^+e^- \to d\bar{d}$	ddbar	1	200
$e^+e^- \to s\bar{s}$	ssbar	1	200
$e^+e^- \to c\bar{c}$	ccbar	1	200
Low multiplicity			
$e^+e^- \to e^+e^-(\gamma)$	ee	10	200
$e^+e^- \to \mu^+\mu^-(\gamma)$	mumu	1	200
$e^+e^- \to e^+e^-e^+e^-$	eeee	4	200
$e^+e^- \to e^+e^-\mu^+\mu^-$	eemumu	4	200
$e^+e^- \to e^+e^-K^-K^+$	eeKK	1	1000
$e^+e^- \to e^+e^-p^-p^+$	eepp	1	1000
$e^+e^- \to e^+e^-\pi^-\pi^+$	eepipi	1	1000

Table 4.1: Process description and integrated luminosity of the Early Phase 3 MC13a samples used in the analysis [11].

Data sets	Processing Label	Experiment	$\int$ L dt (pb^{-1})
Data set 1	Proc11	7	425.5 ± 0.3
Data set 2	Proc11	8	4597.4 ± 0.9
Data set 3	Proc11	10	3741.3 ± 1.1
Data set 4	Buckets 9-11, 13-15	12	54030.5 ± 4.8

Table 4.2: Belle II experiment data sets used in this work. The processing and experiment number labels are shown, as well as, the corresponding integrated luminosity [11].

4.2 Event Selection

4.2.1 Event reconstruction

We reconstruct $e^+e^- \to \tau^+\tau^-$ events where one tau lepton decays to a single charged particle ($\tau^\pm \to \ell^\pm + \alpha$ or $\tau^\pm \to \ell^\pm + \nu_\tau\nu_\ell$) and the other decays to three charged particles ($\tau^\mp \to h^\mp h^- h^+ \nu_\tau$). When reconstructing these 3×1 prong[1] topology events (see next Section 4.2.2), we require exactly four tracks with zero net charge, originating close to the interaction region:

- any displacement from the interaction point should be less than 3 cm in the z-axis,
- and less than 1 cm in the transverse plane.

The "1-prong" signal track is required to be within the momentum (p_{lab}) and polar angle (θ_{lab}) coverage of the lepton ID efficiency correction factors (see Section 4.4.2). To each track in data (not MC), we apply a correction associated to the global momentum scale. This accounts for imperfections in the magnetic field map used during data reconstruction. These corrections were provided by the Belle II Tracking Group [40].

Neutral pion candidates are constructed by combining two photons satisfying the following requirements:

- ECL clusters with energy deposits of at least 0.1 GeV ($E(\gamma)$ >0.1 GeV);
- should be within the CDC acceptance to ensure they are not matched to any charged particle;
- the sum of the number of crystals in an ECL cluster should be greater than 1.5 ;

with an additional requirement on the invariant mass of the two photon system, $115 < M_{\gamma\gamma} < 152$ MeV/c^2. To each π^0 in MC, we apply a correction associated to different π^0 reconstruction efficiencies between data and MC. [41].

Any photon passing the requirements listed above with a higher energy threshold,

- $E(\gamma)$ >0.2 GeV ,

and not coming from a π^0 candidate, is also saved separately.

4.2.2 3x1 prong topology

We use the thrust vector - computed from the four tracks, π^0 candidates, and photons - to separate the events into signal and tag hemispheres. The thrust axis $\widehat{n}_{trust}$ is defined such that the value V_{thrust},

$$V_{thrust} = \sum \frac{\mid \overrightarrow{p}_i^{CMS} \cdot \widehat{n}_{trust} \mid}{\mid \sum \overrightarrow{p}_i^{CMS} \mid} , \tag{4.1}$$

is maximized. Here, $\overrightarrow{p}_i^{CMS}$ is the momentum of each particle in the event in the center of mass (CMS) reference frame.

[1] A charged particle that is reconstructed from detector tracks is commonly known as a "prong".

As can be seen in Figure 4.1, the hemisphere corresponding to the signal side, $\tau^{\pm} \to \ell^{\pm} + \alpha$, should contain a single track ("1-prong"), while the three tracks ("3-prong") compose the tag-side hemisphere. These are mainly $\tau^{\mp} \to 3h^{\mp}(n\pi^0)\nu$ decays, most of them $\tau^{\mp} \to 3\pi^{\mp}\nu$. The 3x1 prong topology is imposed by requiring

$$(\overrightarrow{p}^{CMS}_{signal} \cdot \hat{n}_{trust}) \cdot (\overrightarrow{p}^{CMS}_{tag,i} \cdot \hat{n}_{trust}) < 0 \ , \forall \ i \in \text{tag}. \tag{4.2}$$

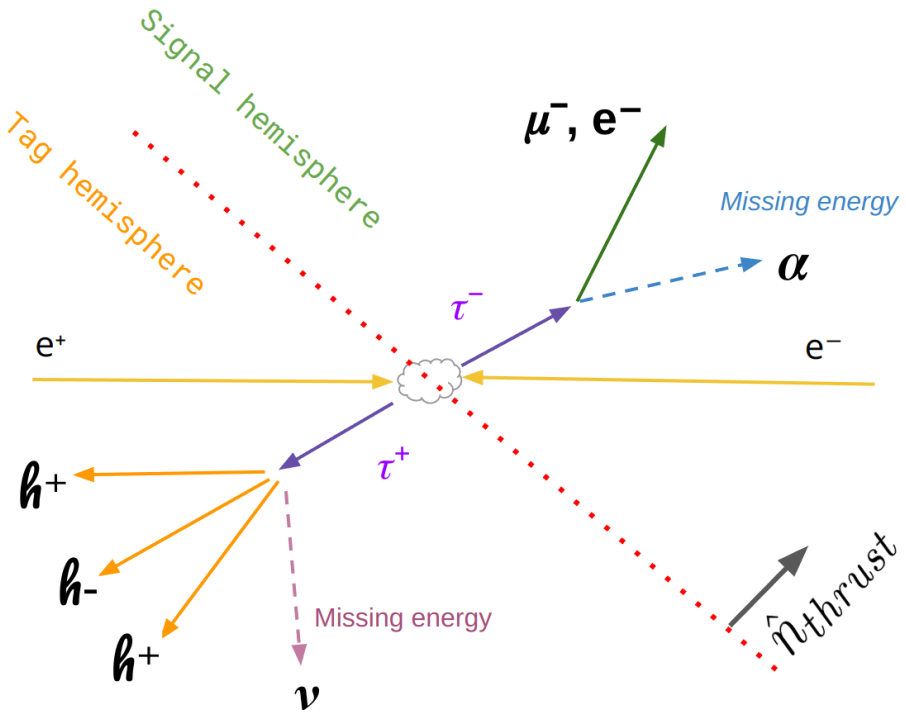


Figure 4.1: Ilustration of the 3×1 topology. The signal $\tau^{\pm} \to \ell^{\pm} + \alpha$ ("1-prong") and tag side ("3-prong") are in opposite hemispheres.

4.2.3 Particle identification

On the signal side, we select the charged track to be either an electron or a muon, by requiring the following particle identification (PID) conditions:

- Global electron ID likelihood > 0.9,

or

- Global muon ID likelihood > 0.9.

Since the lepton ID likelihoods have different performances in data and simulation, we apply correction factors to MC associated to both the lepton efficiency and the probability that a charged pion ($\pi^{\pm}$) is misidentified as a charged lepton ($\ell^{\pm}$) (or $\pi^{\pm} \to \ell^{\pm}$ fake rate). These correction factors are provided by the Lepton ID Group within the Belle II Collaboration [42] and their statistical uncertainties will be considered as a systematic uncertainty source in Section 4.5.

In the "3-prong" tag side, each track satisfy

- cluster $E/p \leq 0.8$,

where "clusterE" is the energy measured in the ECL and p its momentum in the laboratory system. Since electrons deposit most of their energy in the ECL and the ratio clusterE/p for them is close to 1, this requirement basically rejects electrons and keeps hadrons and muons. All "3-prong" branching ratios are negligible compared to $\tau^{\mp} \to 3h^{\mp}(n\pi^0)\nu$. Later, we will also require a neutrals particle veto ($n = 0$).

4.2.4 Trigger

Events in data are required to fire the logical OR of several un-prescaled low-multiplicity ECL (lml) triggers:

- **lml0** or **lml1** or **lml2** or **lml4** or **lml6** or **lml7** or **lml8** or **lml9** or **lml10** or **lml12**

A description of these trigger bits is provided in Appendix B. Hereafter, this combination of triggers will be referred to as *lmlX*.

4.2.5 Skim

For τ physics at Belle II, there are 3 skims available: "thrust", "generic" and "LFV". However, unskimmed samples are used for this analysis.

4.3 Background Suppression

Except for the lepton momentum, in general, the $\tau^{\pm} \to l^{\pm} + \alpha$ and the $\tau^{\pm} \to \ell^{\pm} + \nu\bar{\nu}$ decay are indistinguishable. On the other hand, we have no prior knowledge of the α mass. Thus, the aim of the analysis is to suppress all other SM processes except for $e^+e^- \to \tau^{\pm}(\pi^{\pm}\pi^-\pi^+\nu)\tau^{\mp}(\ell^{\mp} + \nu\bar{\nu})$. The selection is optimized to favor the SM decay $\tau^{\mp} \to \ell^{\mp} + \nu\bar{\nu}$ using the following figure of merit (FOM):

$$FOM = \frac{S}{\sqrt{S+B}} , \tag{4.3}$$

where S is the number of events with $\tau^{\mp} \to \ell^{\mp} + \nu\bar{\nu}$ decays, while B is the number of remaining SM events.

4.3.1 Discriminating variables used for the selection

To maximize the FOM, we use variables weakly correlated to the signal momentum:

- $\#\gamma_{\text{1prong}}$: Number of photons in the event on the signal side ("1-prong").
- $\#\gamma_{\text{3prong}}$: Number of photons in the event on the tag side ("3-prong").
- $\#\pi^0_{\text{1prong}}$: Number of neutral pions in the event on the signal side ("1-prong").
- $\#\pi^0_{\text{3prong}}$: Number of neutral pions in the event on the tag side ("3-prong").

- χ^2 probability: χ^2 probability of the "3-prong" vertex fit.
- p_T leading: Transverse momentum (p_T) of the leading track on the tag side.
- p_T sub-leading: Transverse momentum (p_T) of the sub-leading track on the tag side.
- p_T third: Transverse momentum (p_T) of the third track on the tag side.
- *Thrust*: Thrust magnitude, as calculated in Equation 4.1.
- $M_{\text{Inv}}^{\text{3prong}}$: Invariant mass of the "3-prong" system.

$$M_{\text{Inv}}^{\text{3prong}} = \sqrt{(E_1 + E_2 + E_3)^2 - (\vec{p}_1 + \vec{p}_2 + \vec{p}_3)^2}.$$

- $\text{E}_{CMS}^{\text{3prong}}$: Total energy in the CMS frame of the "3-prong" system.

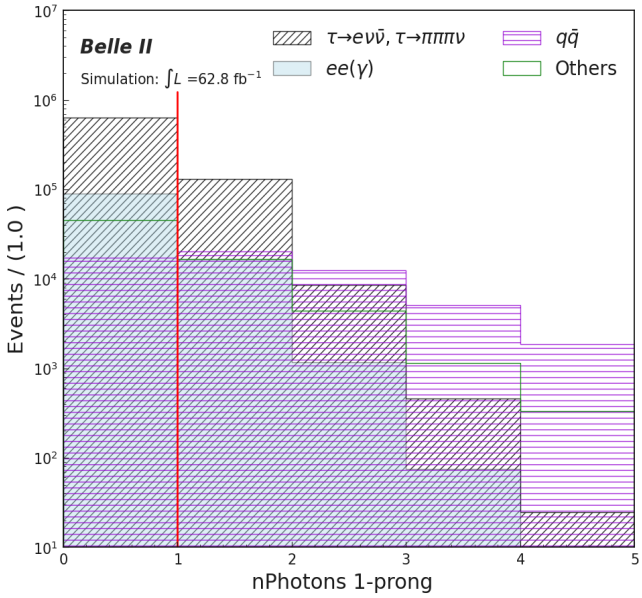
4.3.2 Cut-based selection

Neutral's veto

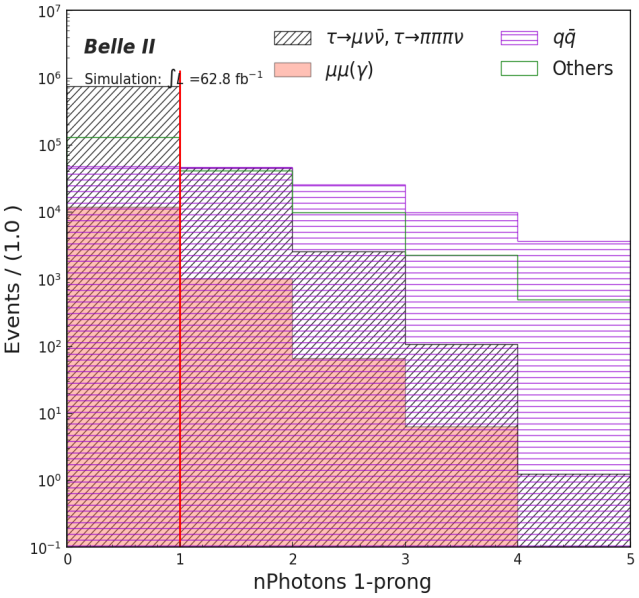
In this work we focus on attaining a pure sample of $\tau^{\pm}(\pi^{\pm}\pi^{-}\pi^{+}\nu)\tau^{\mp}(\ell^{\mp} + \nu\bar{\nu})$ decays. We perform a veto on reconstructed neutral particles in the event:

- $\#\gamma_{\text{1prong}} = \#\gamma_{\text{3prong}} = 0$,
- $\#\pi^0_{\text{1prong}} = \#\pi^0_{\text{3prong}} = 0$.

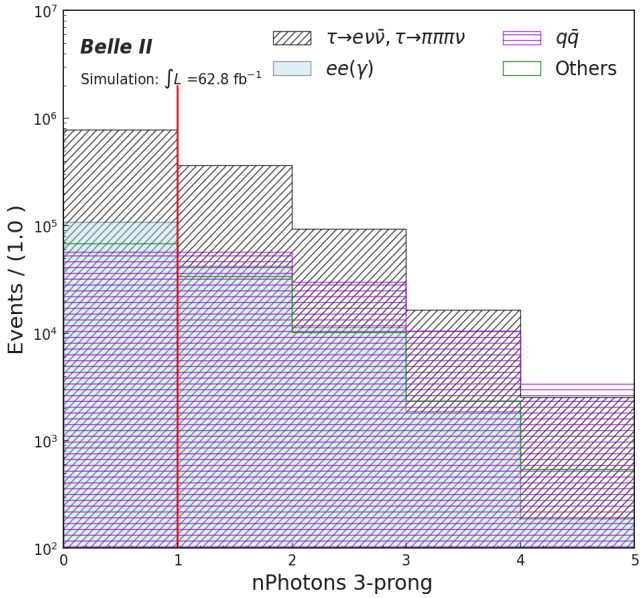
As can be seen from the distribution of the number of photons, $\#\gamma$ in Figure 4.2, and neutral pions, $\#\pi^0$ in Figure 4.3, for 1- and 3-prong sides, these requirements reduce the $q\bar{q}$ contamination. These requirements also prevent radiative τ decays in the signal side and avoid the precise modeling of $\tau^{\pm} \to 3\pi^{\pm}(n\pi^0)\nu$ decays with $n > 0$.



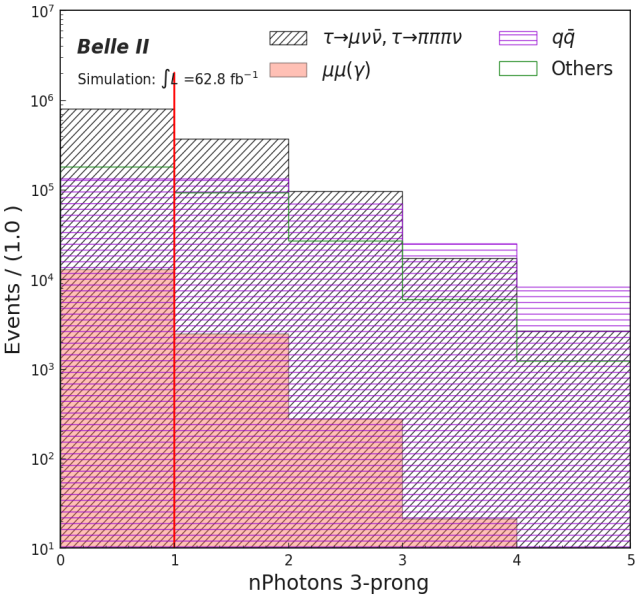
(a) Number of photons on 1-prong side.



(b) Number of photons on 1-prong side.

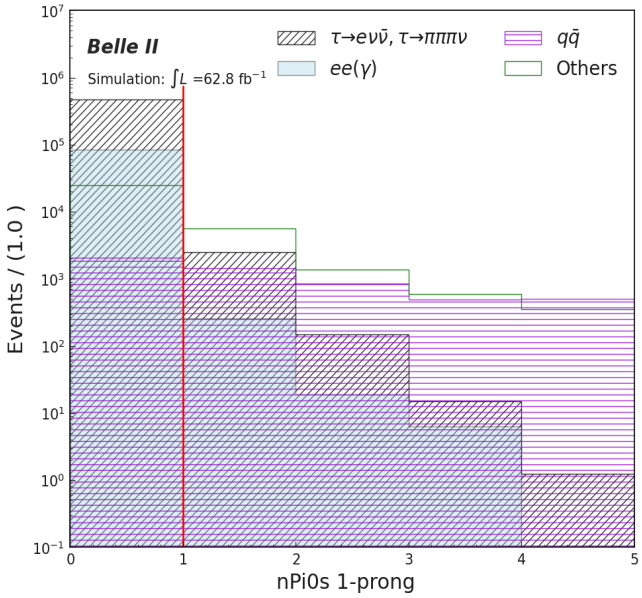


(c) Number of photons on 3-prong side.

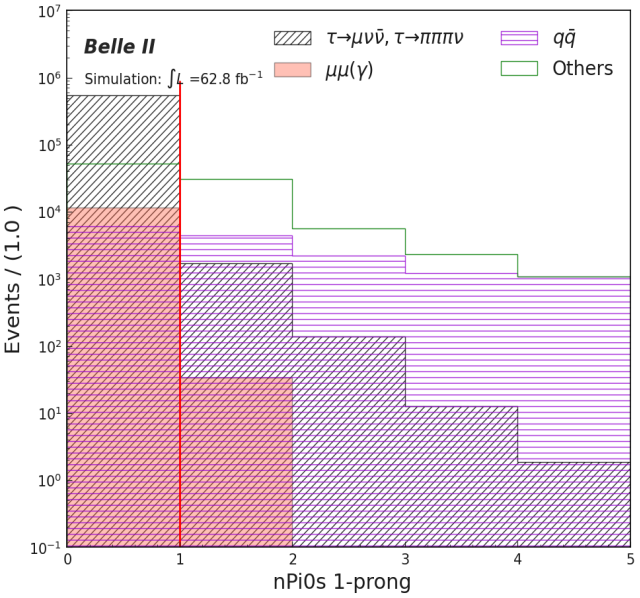


(d) Number of photons on 3-prong side.

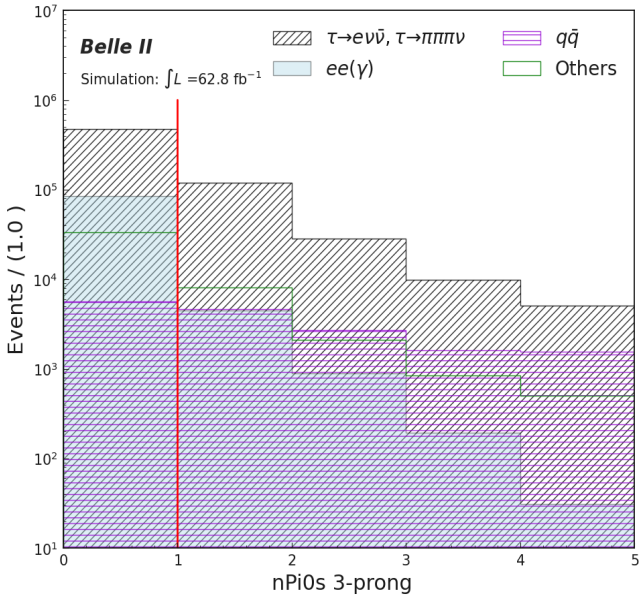
Figure 4.2: Number of photons distributions on the 1-prong and 3-prong side, (a), (c) for $\tau^{\pm} \to e^{\pm}\nu\bar{\nu}$ (left) and (b), (d) $\tau^{\pm} \to \mu^{\pm}\nu\bar{\nu}$ (right). The requirement on these variables is indicated with a solid red line, events to the right are rejected [11].



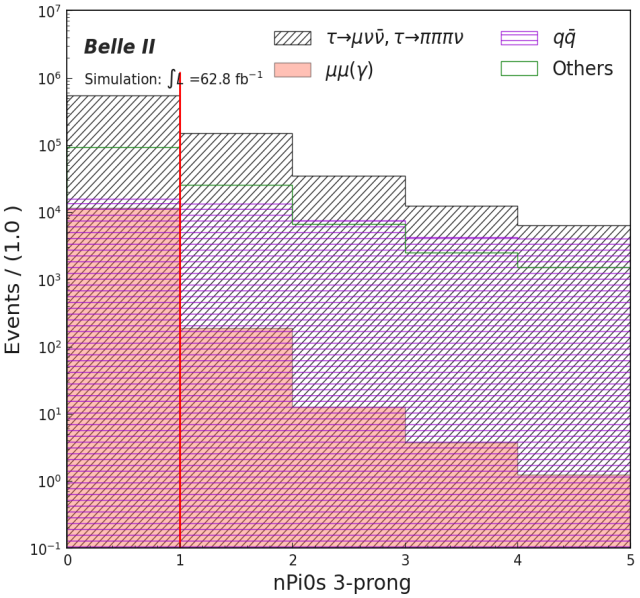
(a) Number of neutral pions on 1-prong side.



(b) Number of neutral pions on 1-prong side.



(c) Number of neutral pions on 3-prong side.



(d) Number of neutral pions on 3-prong side.

Figure 4.3: Number of π^0 mesons distributions on the 1-prong and 3-prong side, (a), (c) for $\tau^{\pm} \to e^{\pm}\nu\bar{\nu}$ (left) and (b), (d) $\tau^{\pm} \to \mu^{\pm}\nu\bar{\nu}$ (right). The requirement on these variables is indicated with a solid red line, events to the right are rejected [11].

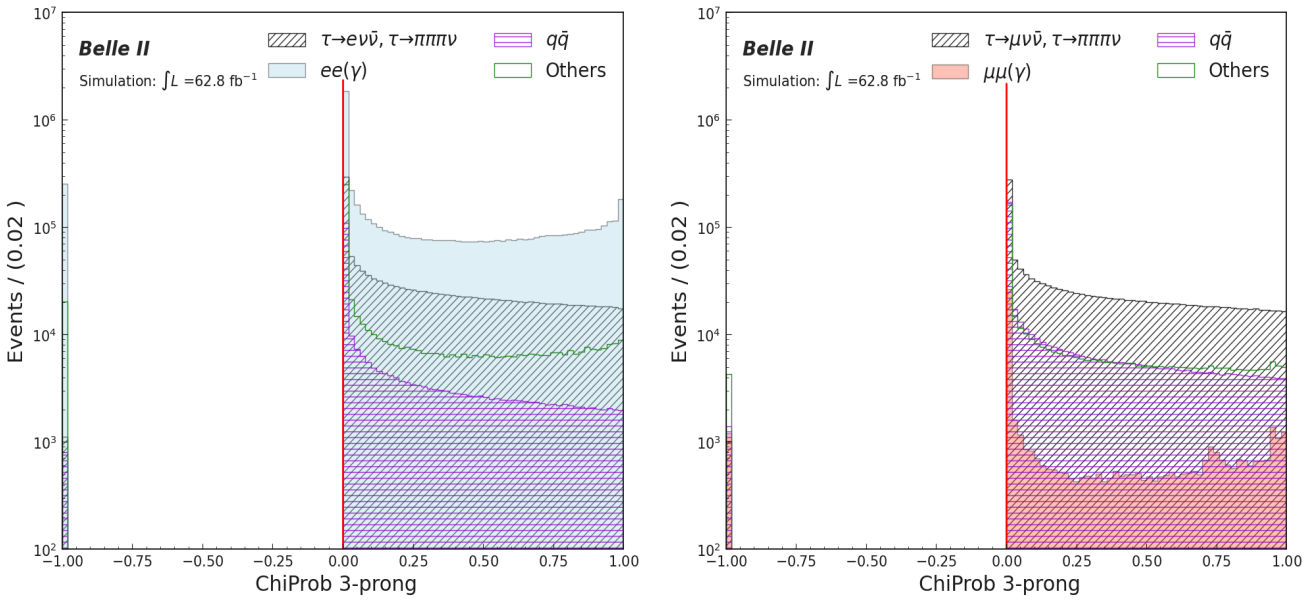
Vertex fit

We reject events where the tag side "3-prong" vertex fit fails to converge, using the distribution of the χ^2 probability of the fit. We require

- χ^2 probability > 0 .

The distribution is shown in Figure 4.4.

The events that we want to reject are either composed of decay chains with long-lived intermediate particles or containing tracks from the underlying beam background (not from collisions) or with actual 3-track vertices (regardless of whether coming from τ pair events or not) but displaying poor reconstruction quality. In the case of long-lived intermediates, this rejection also suppresses a portion of τ tags decaying into K_S, resulting in a slight reduction in reconstruction efficiency. Failed vertex fits are rejected for simplicity, otherwise, the event would require exclusive treatment.



(a) χ^2 probability of the 3-prong side

(b) χ^2 probability of the $\tau \to \pi\pi\pi\nu$ decay.

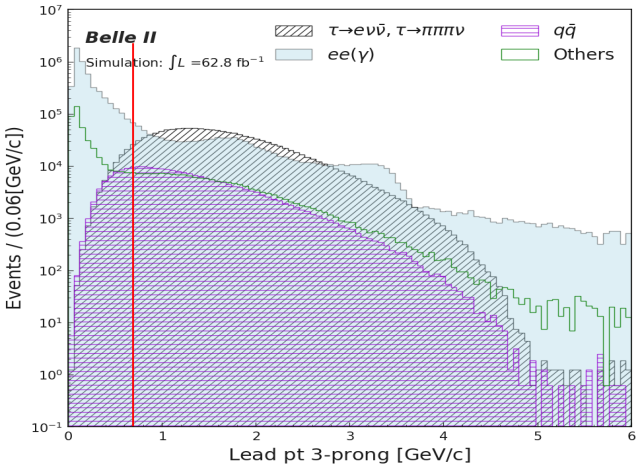
Figure 4.4: Common vertex probability for the "3-prong" fit, (a) the electron channel and b) the muon channel. A value of -1 corresponds to a failed vertex fit convergence. The requirement on these variables is indicated with a solid red line, events to the left are rejected [11].

Transverse momentum

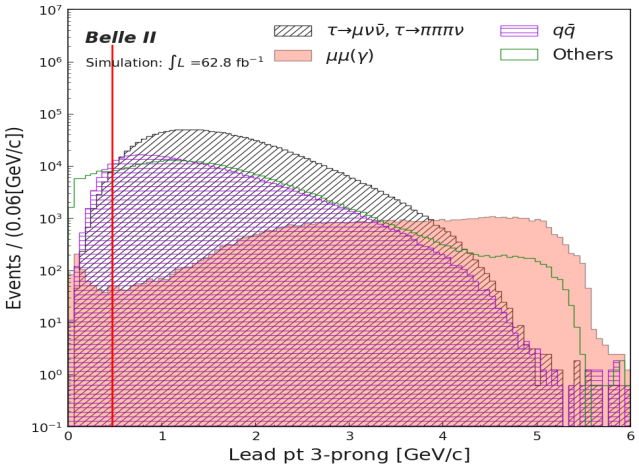
After ranking the transverse momenta of the three tracks on the tag side, we obtain the following requirements by maximizing Equation 4.3 for the electron (muon) sample:

- p_T of the leading track > 0.69 (0.47) GeV/c,
- p_T of the sub-leading track > 0.29 (0.17) GeV/c, and
- p_T of the third track > 0.08 (0.04) GeV/c.

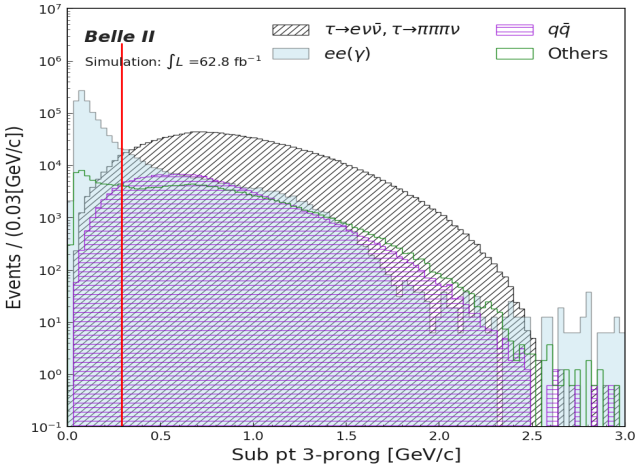
Figure 4.5 shows that these requirements mainly help to reduce the significant contamination coming from the $e^+e^-(\gamma)$ and $\mu^+\mu^-(\gamma)$ final states found in the electron and muon channel, respectively.



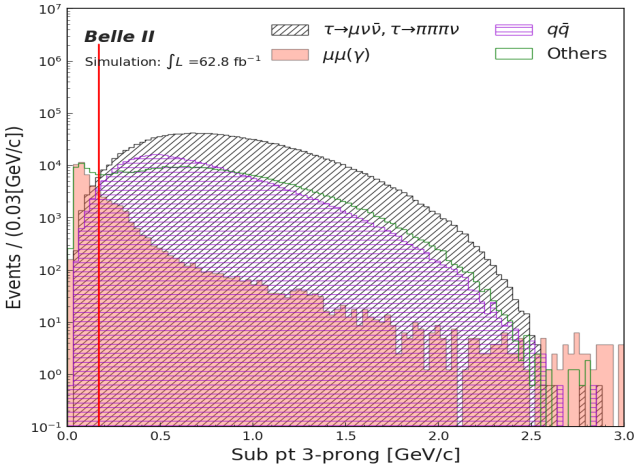
(a) p_T of the leading track.



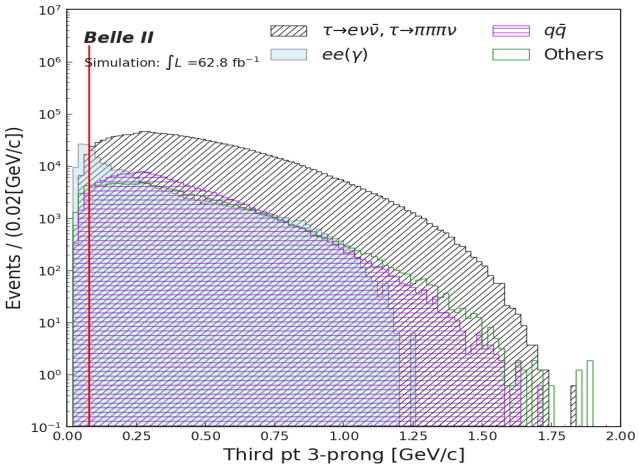
(b) p_T of the leading track.



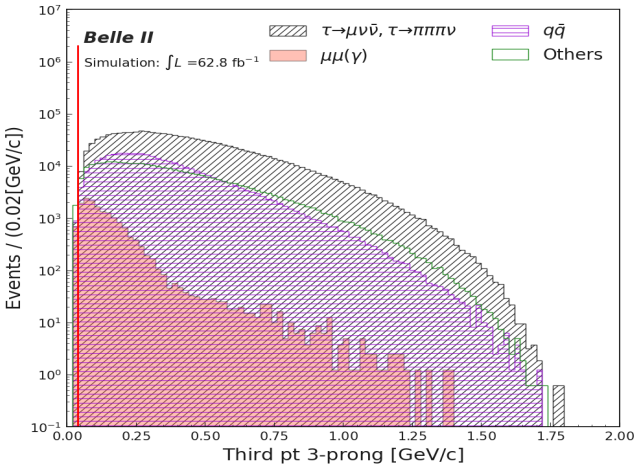
(c) p_T of the sub-leading track.



(d) p_T of the sub-leading track.



(e) p_T of the third track.



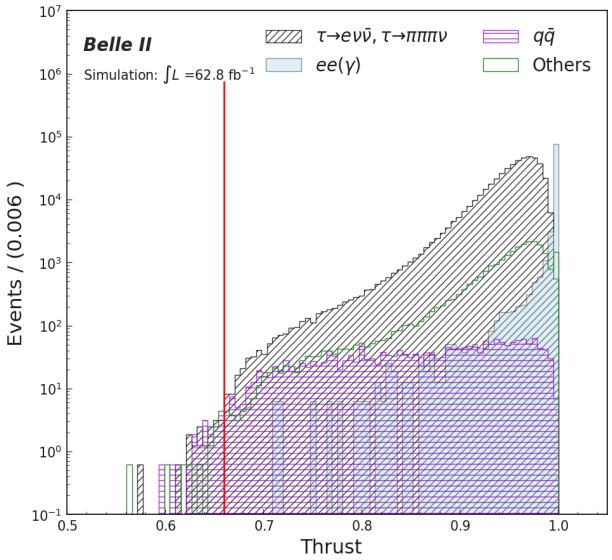
(f) p_T of the third track.

Figure 4.5: Transverse momentum distributions of the track tags for the electron (left) and muon (right) channels. The requirement on these variables is indicated with a solid red line, events to the left are rejected.

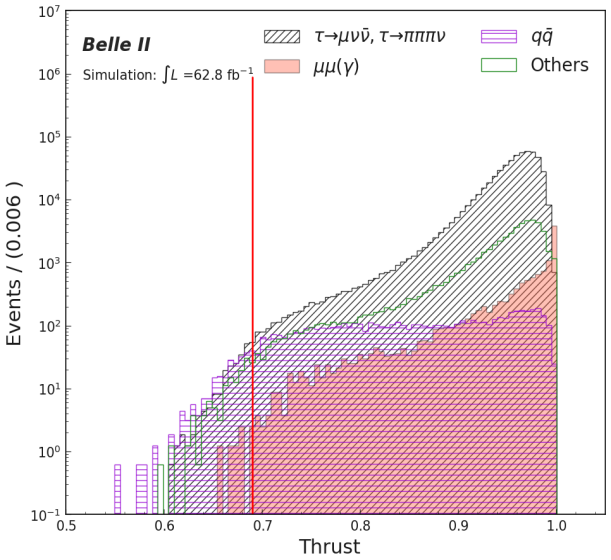
Further selections

We finally use the *Thrust* value (V_{thrust} in Equation 4.1), $E_{\mathrm{CMS}}^{\mathrm{3prong}}$ and $M_{\mathrm{Inv}}^{\mathrm{3prong}}$ (see Section 4.3.1) to suppress the background processes further. The distributions of these variables, with the above selection already applied, are shown in Figure 4.6, 4.8, and 4.7, respectively. After the FOM optimization on these variables, we obtain the following requirements for the electron (muon) channel:

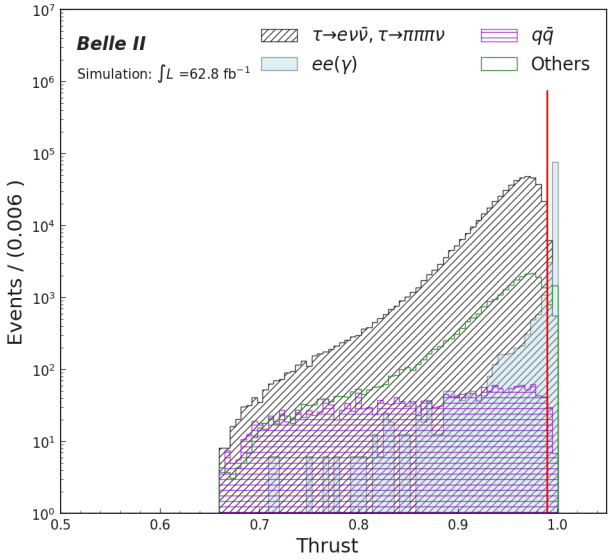
- 0.66 (0.69)$< Thrust <$ 0.99 (1.0) ;



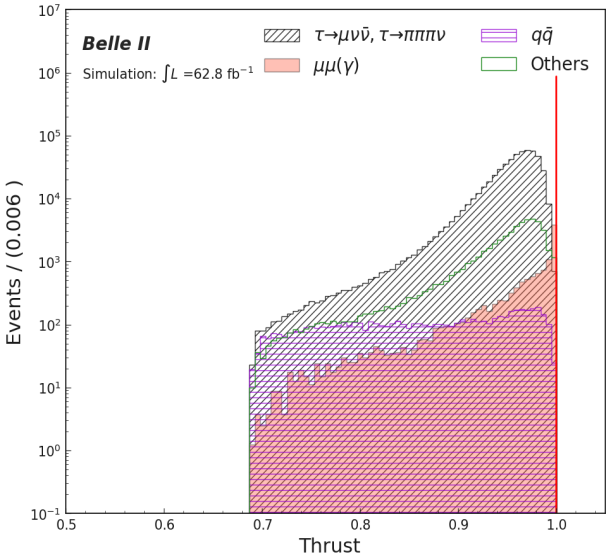
(a) Thrust, lower cut.



(b) Thrust, lower cut.



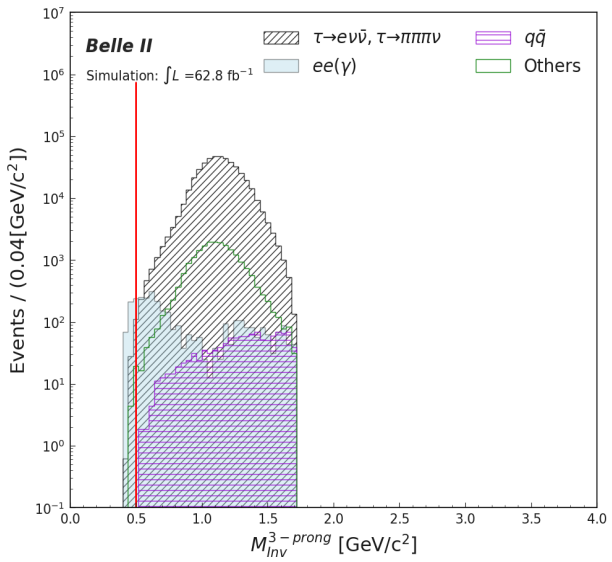
(c) Thrust, upper cut.



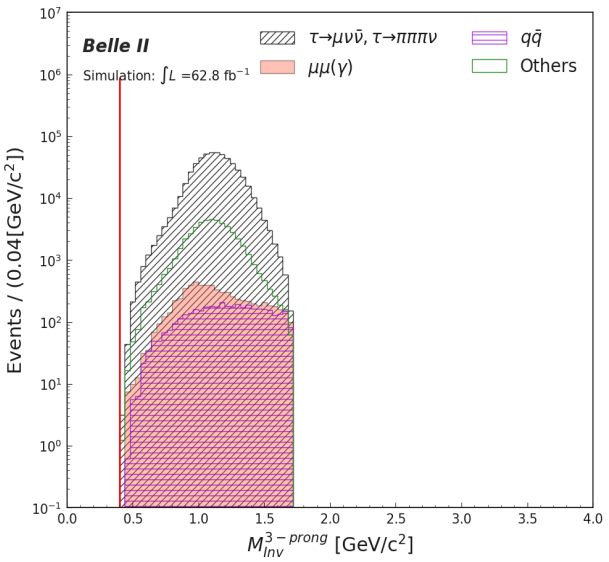
(d) Thrust, upper cut.

Figure 4.6: Distributions for the *Thrust* variable, with the requirements of the previous variables. Lower (top) and upper (bottom) cuts are shown for (a), (c) the electron and (b), (d) muon channels. The requirements are indicated with solid red lines, we keep events between these lines.

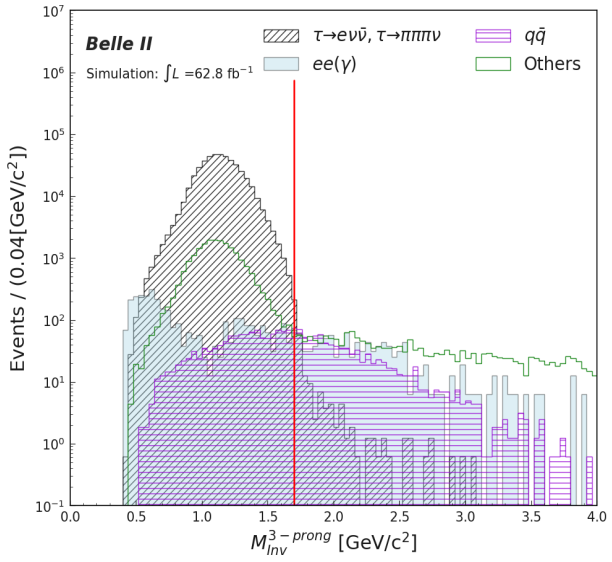
- $0.5(0.4)\ \mathrm{Gev}/c^2 < M_{\mathrm{Inv}}^{\mathrm{3prong}} < 1.7(1.7)\ \mathrm{Gev}/c^2$;



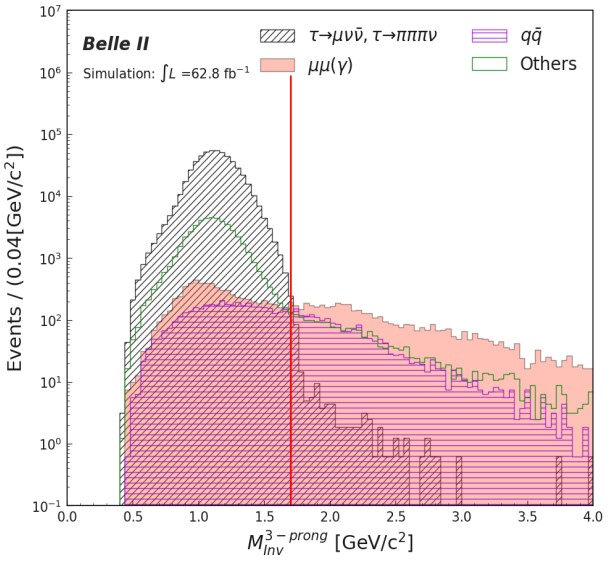
(a) $M_{\mathrm{Inv}}^{\mathrm{3prong}}$, lower cut.



(b) $M_{\mathrm{Inv}}^{\mathrm{3prong}}$, lower cut.



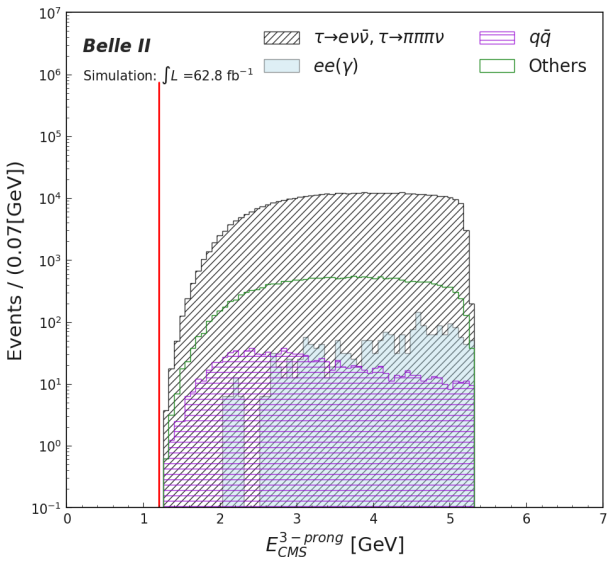
(c) $M_{\mathrm{Inv}}^{\mathrm{3prong}}$, upper cut.



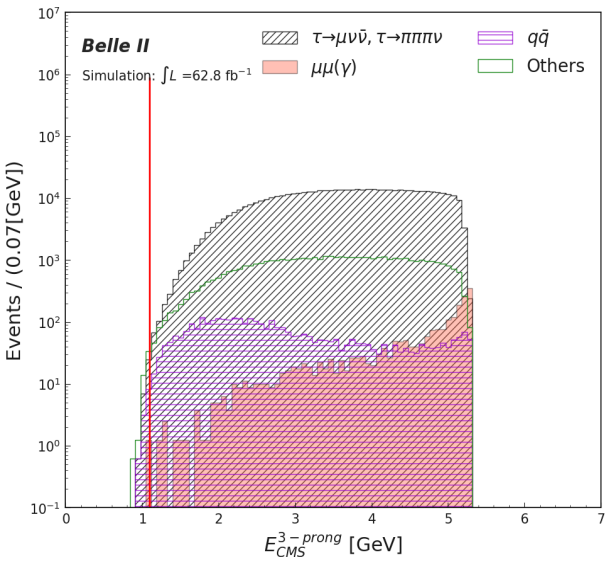
(d) $M_{\mathrm{Inv}}^{\mathrm{3prong}}$, upper cut.

Figure 4.7: Distributions for the $M_{\mathrm{Inv}}^{\mathrm{3prong}}$ variable, with the requirements of the previous variables. Lower (top) and upper (bottom) cuts are shown for (a), (c) the electron and (b), (d) muon channels. The requirements are indicated with solid red lines, we keep events between these lines.

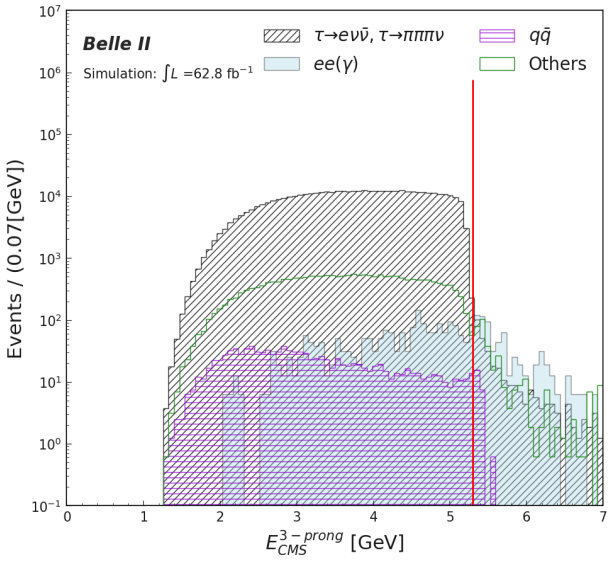
- 1.2 (1.1) GeV $< \mathrm{E}_{CMS}^{\mathrm{3prong}} <$ 5.3 (5.3) GeV .



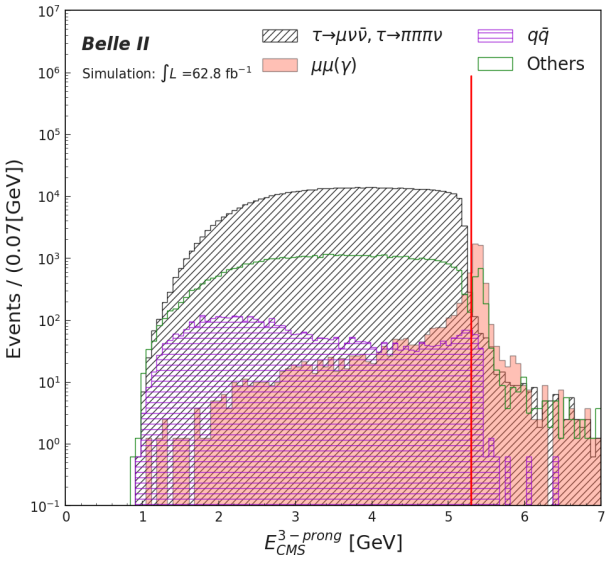
(a) $E_{\mathrm{CMS}}^{\mathrm{3prong}}$, lower cut.



(b) $E_{\mathrm{CMS}}^{\mathrm{3prong}}$, lower cut.



(c) $E_{\mathrm{CMS}}^{\mathrm{3prong}}$, upper cut.



(d) $E_{\mathrm{CMS}}^{\mathrm{3prong}}$, upper cut.

Figure 4.8: Distributions for the $E_{\mathrm{CMS}}^{\mathrm{3prong}}$ variable, with the requirements of the previous variables. Lower (top) and upper (bottom) cuts are shown for (a), (c) the electron and (b), (d) muon channels. The requirements are indicated with solid red lines, we keep events between these lines.

Furthermore, for reasons we will explain shortly, we impose the following requirements:

- $Thrust > 0.9$,
- lepton ID weights > -1.

The first is a data-driven requirement to reject the remaining 2-photon exchange processes that are not simulated [43] (discrepancies between data and MC for thrust less than 0.9 are shared among all τ analyses in Belle II and should be worked out in the future). The second cut restricts the lepton kinematics to the $p - \theta$ region where lepton ID corrections are available, as mentioned in Section 4.2.3.

4.3.3 Remaining events after cut-based selection

Table 4.3 summarizes the number of background events after the selection, while the number of signal events, purities, and efficiencies of the selection are given in Table 4.4. The definition of efficiency and purity is described in detail in Appendix A.

BKG Samples	e	μ
Generic		
$\tau^+\tau^-$ (other)	15773.5	39506.9
$u\bar{u}$	129.5	660.0
$d\bar{d}$	27.0	125.6
$s\bar{s}$	28.3	226.7
$c\bar{c}$	249.9	398.2
B^+B^-	3.8	7.5
$B^0\bar{B}^0$	10.7	15.1
Low multiplicity		
$\mu^+\mu^-(\gamma)$	0.0	1602.7
$e^+e^-(\gamma)$	841.5	0.0
$e^+e^-\mu^+\mu^-$	21.4	692.1
$e^+e^-e^+e^-$	17.0	0.0
$e^+e^-k^+k^-$	0.0	0.1
$e^+e^-\pi^+\pi^-$	0.1	4.0
$e^+e^-p^+p^-$	0.0	0.1
TOTAL BKG	17102.7	43239.0

Table 4.3: Remaining background MC events after cut-based selection. Here, $\tau^+\tau^-$ (other) refers to $\tau-$pair events that do not decay to $\tau^{\pm} \to \ell^{\pm} + \nu\nu$ (see Appendix A). The number of remaining MC events has been scaled to 62.8 fb^{-1} [11].

	e			μ		
Sample $\tau \to \ell\alpha$	Events	Efficiency	Purity	Events	Efficiency	Purity
$M_\alpha = 0$	407469	13.40%	-	529168	17.41%	
$M_\alpha = 0.5$	410354	13.50%	-	525852	17.30%	
$M_\alpha = 0.7$	411825	13.55%	-	525049	17.27%	
$M_\alpha = 1.0$	416663	13.71%	-	513444	16.89%	
$M_\alpha = 1.2$	420930	13.85%	-	494769	16.28%	
$M_\alpha = 1.4$	396060	13.03%	-	448015	14.74%	
$M_\alpha = 1.6$	285526	9.39%	-	276482	9.09%	
$\tau \to \ell + \nu\bar{\nu}$	395140	12.65%	95.85%	493722	16.18%	91.95%

Table 4.4: Efficiencies and purities of signal samples [11].

4.3.4 Summary of the cut-based selection

Table 4.5 is a summary of the cut-based selection criteria.

Selection criteria	$\tau \to e\nu\bar{\nu}$	$\tau \to \mu\nu\bar{\nu}$
Event	$0.9 < Thrust < 0.99$	$0.9 < Thrust < 1.0$
Signal (1 prong) $\tau^\pm \to \ell^\pm\alpha$	electronID > 0.9 $\#\gamma_{1prong} = 0$ $\#\pi^0_{1prong} = 0$ in the acceptance of e-ID performance measurements	muonID > 0.9 $\#\gamma_{1prong} = 0$ $\#\pi^0_{1prong} = 0$ in the acceptance of μ-ID performance measurements
Tag (3 prong) $\tau^\pm \to 3\pi^\pm(\nu)$	clusterE/p ≤ 0.8 vertex χ^2 probability > 0 leading $p_t > 0.69$ GeV/c sub-leading $p_t > 0.29$ GeV/c third $p_t > 0.08$ GeV/c $\#\pi^0_{3\text{prong}} = 0$ $\#\gamma_{3\text{prong}} = 0$ $0.5 < M^{3\text{prong}}_{\text{Inv}} < 1.7$ GeV/c^2 $1.2 < E^{3\text{prong}}_{CMS} < 5.3$ GeV	clusterE/p ≤ 0.8 vertex χ^2 probability > 0 leading $p_t > 0.47$ GeV/c sub-leading $p_t > 0.17$ GeV/c third $p_t > 0.04$ GeV/c $\#\pi^0_{3\text{prong}} = 0$ $\#\gamma_{3\text{prong}} = 0$ $0.4 < M^{3\text{prong}}_{\text{Inv}} < 1.7$ GeV/c^2 $1.1 < E^{3\text{prong}}_{\text{CMS}} < 5.3$ GeV

Table 4.5: Summary of the cut-based selection criteria for the electron and the muon channels [11].

4.4 Measurement strategy

In this section, we present a first sensitivity analysis to measure

$$R = \mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \alpha)/\mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \nu\bar{\nu}) \ . \tag{4.4}$$

If from the data we obtain a value significantly different from zero, we will report it with its corresponding confidence interval. If, on the other hand, the data yields a value of R that is consistent with zero, then the result of this search will be reported in terms of an upper limit at 95 % C.L. The statistical analysis is based on a frequentist approach using the asymptotic CLs technique [44] implemented using the RooStats and HistFactory statistical tools of the ROOT data analysis framework from CERN [45].

4.4.1 Statistical model

To measure the branching ratio of $\tau^{\pm} \to \ell^{\pm} + \alpha$ decays in data, we model the data taking into account all possible physics processes. For any discriminating variable, x, the data can be modeled as

$$F(x) = N_{\alpha} \times f_{\alpha}(x) + N_{\ell\nu\nu} \times f_{\ell\nu\nu}(x) + N_{bkg} \times f_{bkg}(x), \tag{4.5}$$

where f_i denotes the corresponding probability density function (PDF) of the i-th physics contribution:

- f_{α}: $\tau^{\pm} \to \ell^{\pm}\alpha$ (NP signal) ;
- $f_{\ell\nu\nu}$: $\tau^{\pm} \to \ell^{\pm}\nu\nu$ (SM signal) ,
- f_{bkg}: $\tau^{\pm}\tau^{\pm}$, $q\bar{q}$, $B^{+}B^{-}$, $B^{0}\bar{B}^{0}$, and other low multiplicity processes.

N_i is the number of events of the corresponding process, which can be calculated from MC for a given signal channel i with tag channel j as:

$$N_i = 2L\sigma(e^{+}e^{-} \to \tau^{+}\tau^{-})\mathcal{B}(\tau \to signal_i)\mathcal{B}(\tau \to tag_j)\epsilon_{ij}, \tag{4.6}$$

where ϵ_{ij} encodes efficiencies and detector acceptances.

The model satisfies $\int F(x) = N_T = N_{\alpha} + N_{\ell\nu\nu} + N_{bkg}$. Here N_T is the total number of observed events.

From Equation 4.6, we can easily obtain the parameter of interest (POI) of this analysis:

$$R \equiv \frac{\mathcal{B}(\tau \to \ell\alpha)}{\mathcal{B}(\tau \to \ell\nu\nu)} = \frac{\epsilon_{\ell\nu\nu}}{\epsilon_{\alpha}} \frac{N_{\alpha}}{N_{\ell\nu\nu}} \ . \tag{4.7}$$

Moreover, we can rewrite Equation 4.5 as follows:

$$F(x) = \frac{\epsilon_{\alpha}}{\epsilon_{\ell\nu\nu}} \times N_{\ell\nu\nu} \times R \times f_{\alpha}(x) + N_{\ell\nu\nu} \times f_{\ell\nu\nu}(x) + N_{bkg} \times f_{bkg}(x) \ . \tag{4.8}$$

In this way, we can extract the value of R (relative branching ratio) directly by fitting the model to the data with $N_{\ell\nu\nu}$ and N_{bkg} allowed to float in the fit. We expect common systematic uncertainties of the NP and the SM modes will largely cancel out in the ratio defined in Equation 4.8. The level of cancellation depends on how similar are the NP and SM channel distributions.

Table 4.6 shows the relative efficiencies and uncertainties that we use as input values in our model.

M_α (GeV/c^2)	$\epsilon_\alpha/\epsilon_{e\nu\nu}$	$\epsilon_\alpha/\epsilon_{\mu\nu\nu}$
0	1.06 ± 0.002	1.08 ± 0.002
0.5	1.07 ± 0.002	1.07 ± 0.002
0.7	1.07 ± 0.002	1.07 ± 0.002
1.0	1.08 ± 0.002	1.04 ± 0.002
1.2	1.09 ± 0.002	1.01 ± 0.002
1.4	1.03 ± 0.002	0.91 ± 0.002
1.6	0.74 ± 0.011	0.56 ± 0.001

Table 4.6: Relative efficiencies for different α masses [11].

4.4.2 Corrections

To ensure the correct calculation of the measurement of R, the following corrections are applied to MC:

- *lmlX* trigger efficiency. For the *lmlX* trigger used in this analysis, we compute the trigger efficiency in data with a tag-and-probe approach in bins of x_{prf}, with a small correction from TSIM to bring the relative to absolute trigger efficiency. More details are given in Appendix B.

- The electron and muon ID efficiencies. As the performance of the lepton ID is different in MC simulation and data, correction factors are needed to take the discrepancies in lepton identification efficiency into account. These factors are taken from the tables provided by the Belle II Lepton ID Performance Group [46].

- The $\pi \to e$ and $\pi \to \mu$ fake rates. For the difference performance in data and MC, we apply the corrections due to the probability that a charged pion ($\pi^\pm$) is misidentified as a charged lepton ($\ell^\pm$). These factors are taken from the tables provided by the Belle II Lepton ID Performance Group [46].

For these corrections we follow the official recommendations of the Belle II Performance Group [47]. The corrections are applied to the binned MC distributions, as weights in each bin.

4.4.3 Measurement of R in pseudo-data

To test the measurement strategy on R, templates obtained from MC simulations are used to model the probability density functions $f_i(x)$ in Equation (4.5). The templates are shown in Figure 4.9. The templates were produced from independent samples of those used for the test measurement. As a discriminating variable (instead of the lepton momentum) we use

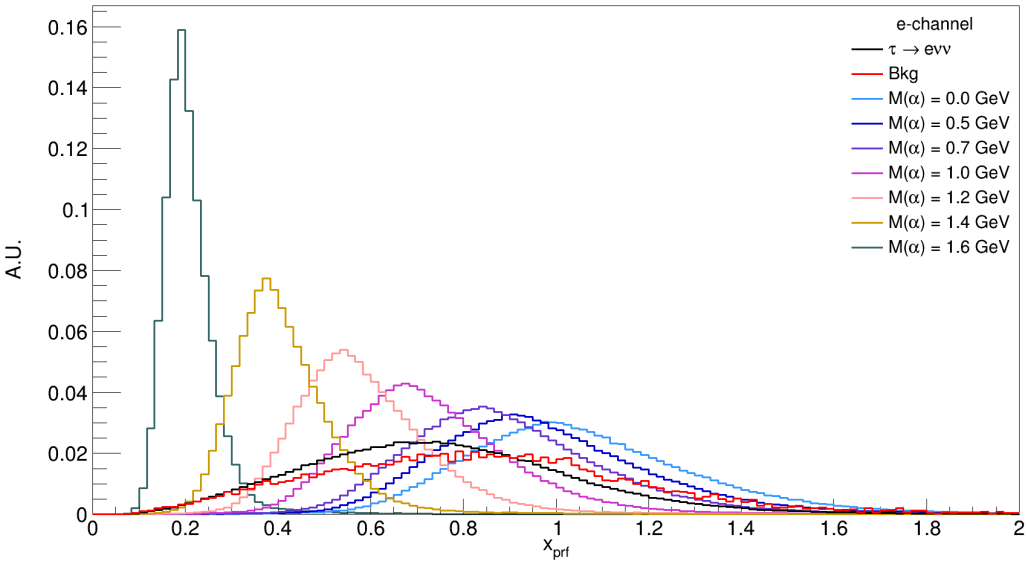
$$x = x_{prf} = 2E_\ell/m_\tau \ , \tag{4.9}$$

where E_ℓ is the energy of the final state lepton in the τ pseudo-rest frame and m_τ is the mass of the τ lepton [13]. This variable is often used to present the theoretical distribution of $\tau^\pm \to \ell^\pm\nu\nu$ and it has a defined range of $0 \leq x_{prf} \leq 2$.

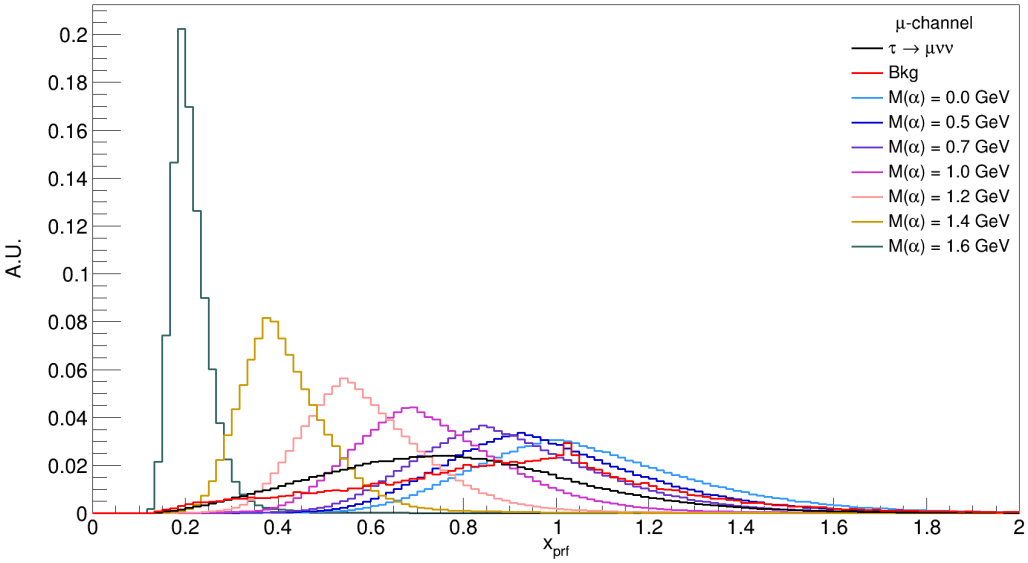
With the templates that model the probability density functions $f_i(x)$ in Equation 4.5, we produce a set of 1000 pseudo samples (MC toys) corresponding to the background only hypothesis, each with $N_{\ell\nu\nu}$ and N_{bkg} reported in Table 4.3 and Table 4.4, corresponding to 62.8 fb^{-1}.

We fit the x distribution without signal and with several levels of signal injected and α masses. All the results were consistent with the inputs, as an example, the results for the case of signal injected according to the ARGUS bounds are shown in Table 4.7.

Assuming no signal is found, we compute the upper limit of R for each sample MC toy, without injecting signal. This way, we can determine the mean and standard deviation of the upper limit results under the background-only hypothesis. In Figure 4.10 and Table 4.8 are presented the results for each mass.



(a) Electron channel.

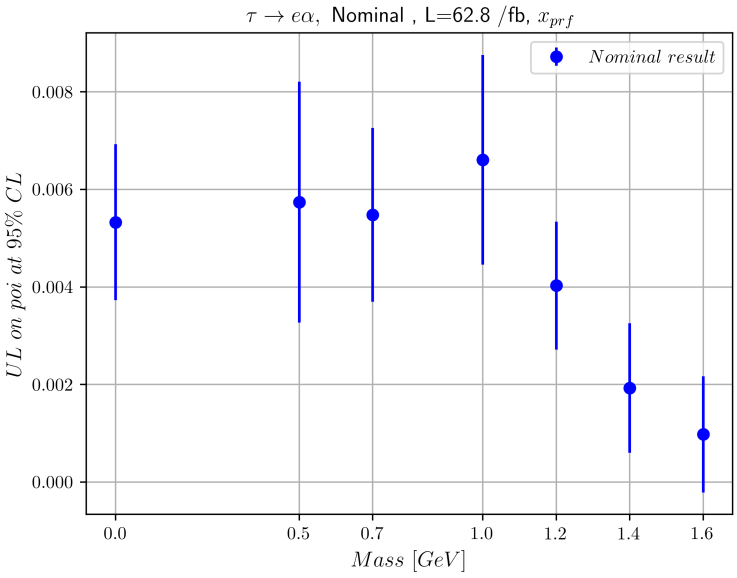


(b) Muon channel.

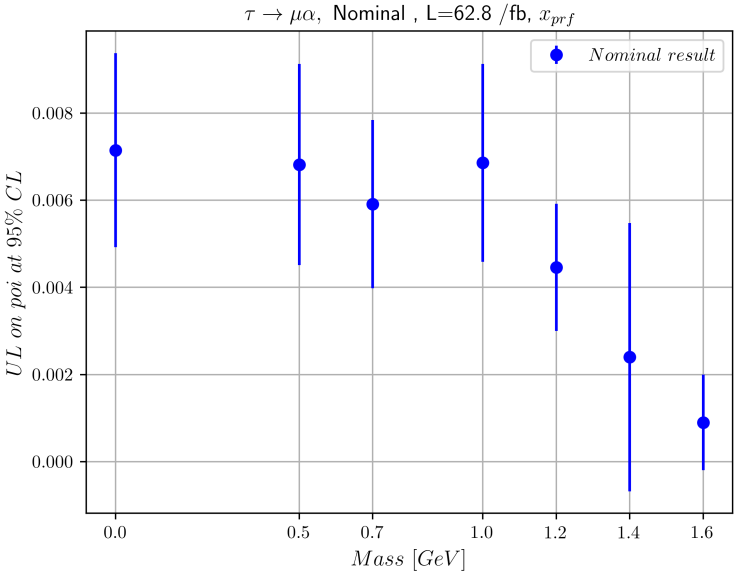
Figure 4.9: Normalized signal and background MC templates for the variable x_{prf} for the (a) electron and (b) muon channels [11].

M_α (GeV/c^2)	$\tau \to e + \alpha$		$\tau \to \mu + \alpha$	
	ARGUS	Fit result	ARGUS	Fit result
0	0.015	0.0169 ± 0.0026	0.026	0.0306 ± 0.0034
0.5	0.017	0.0164 ± 0.0026	0.029	0.0304 ± 0.0033
0.7	0.024	0.0216 ± 0.0027	0.027	0.0266 ± 0.0029
1.0	0.036	0.0314 ± 0.0034	0.03	0.0349 ± 0.0036
1.2	0.034	0.0356 ± 0.0023	0.022	0.023 ± 0.0023
1.4	0.025	0.0262 ± 0.0011	0.033	0.0315 ± 0.0011
1.6	0.006	0.0062 ± 0.0004	0.003	0.0032 ± 0.0004

Table 4.7: Fit results for the extraction of $Br(\tau \to \ell\alpha)/Br(\tau \to \ell\nu\bar{\nu})$ with signal injected according to the ARGUS bounds [11].



(a) Electron channel.



(b) Muon channel.

Figure 4.10: Expected upper limit on R at 95% C.L. for the (a) electron and (b) muon channels [11].

M_α (GeV/c^2)	$\mathcal{B}(\tau^{\pm} \to e^{\pm}\alpha)/\mathcal{B}(\tau^{\pm} \to e^{\pm}\nu\bar{\nu})$	$\mathcal{B}(\tau^{\pm} \to \mu^{\pm}\alpha)/\mathcal{B}(\tau^{\pm} \to \mu^{\pm}\nu\bar{\nu})$
0	0.0053 ± 0.0016	0.0072 ± 0.0022
0.5	0.0057 ± 0.0025	0.0068 ± 0.0023
0.7	0.0055 ± 0.0018	0.0059 ± 0.0019
1.0	0.0066 ± 0.0022	0.0069 ± 0.0023
1.2	0.004 ± 0.0013	0.0045 ± 0.0015
1.4	0.0019 ± 0.0013	0.0024 ± 0.0031
1.6	0.001 ± 0.0012	0.0009 ± 0.0011

Table 4.8: Expected upper limit on R at 95% C.L. for the different mass values of the α particle [11].

Furthermore, we performed additional tests:

- A comparison of the upper limit calculation was made using different software implementations (RooStats, pyhf, and BAT), and the results were consistent.
- For the case of an injected signal, a calculation of the significance of the observed signal at different integrated luminosities was performed. We conclude that the current implementation is stable and sensitive to the different components of the model.

4.5 Systematic uncertainties

The systematic uncertainties are implemented in our model as shape-correlated nuisance parameters that follow a Gaussian distribution. This way, we include the systematic uncertainties in a unified approach that works either for measuring or setting an upper limit. We have considered the following sources:

- **Lepton ID efficiency**:
  The lepton ID corrections factors provided by the Belle II Lepton ID efficiency performance group have associated statistical and systematic uncertainties binned in the momentum p_{lab} and polar angle θ_{lab} [42], whose sum in quadrature is taken here as the systematic uncertainty due to the lepton ID efficiency corrections.
- **Lepton ID fake rate**:
  Similarly, the lepton ID $\pi^{\pm} \to \ell^{\pm}$ fake rate corrections have associated statistical and systematic uncertainties binned in the momentum p_{lab} and polar angle θ_{lab} [42]. The typical ranges of both ID corrections are summarized in Table 4.9.

	Corr. range	Uncertainty range	Average uncertainty
Electron	$0.84 - 1.06$	$0.9\% - 12.6\%$	$+5.3\%, -2.9\%$
Muon	$0.63 - 1.02$	$1.3\% - 32.8\%$	$+11.7\%, -1.6\%$
Electron fake	$0.6 - 6.0$	$4.3\% - 34.6\%$	$+17.6\%, -14.7\%$
Muon fake	$0.3 - 1.5$	$1.4\% - 37.0\%$	$+18.0\%, -18.2\%$

Table 4.9: Typical ranges for corrections to the lepton-identification efficiencies and misidentification rates, together with ranges for their respective uncertainties and their average values [12].

- **Trigger efficiency**:
  This efficiency has an associated statistical uncertainty. The trigger uncertainties ranges are around $0.1\% - 4\%$ for the electron channel and $0.2\% - 1.5\%$ for the muon channel, depending on the value of x_{prf}.

- π^0 **efficiency**:
  The overall MC π^0 reconstruction efficiency correction (see Section 4.6.2.) provided by the Belle II Neutrals Performance Group has an associated total uncertainty [41]. The neutral pion reconstruction efficiency is evaluated from these studies on independent data samples to be 0.914 ± 0.020.

- **Tracking efficiency**:
  The overall data/MC discrepancy in the tracking efficiency has an associated total uncertainty provided by the Belle II Tracking Group [40].

- **Track momentum scale** (to be evaluated on data):
  The overall momentum scale correction applied to tracks in data, which is evaluated from independent data samples, has an associated total uncertainty [40].

- **Beam energy**:
  We account for the effects of the spread in the measured beam energy following the recommendations of the Belle II Performance Group [48].

- **Relative efficiency**:
  We take into account the fact that the statistics of the MC samples used for this analysis are finite.

The first three items above represent the dominant sources of systematic uncertainties (the resulting uncertainties are symmetrized as mentioned in Section 4.6.6.), while the rest are identified to have a negligible effect on the upper limit estimation, through direct inspections of the R results in the previous generated pseudo data, fitted with alternative templates and efficiencies in a range allowed by each systematic uncertainty source.

4.6 Data-MC validation

4.6.1 Strategy

An important aspect of this analysis is to ensure that we have confidence in the modeling of data, without looking at the actual data[2]. Therefore, we have developed a validation strategy to compare the data and MC, which will be described and justified in this subsection.

The general idea is to compare MC and data with two methods:

1. Look at 10% of data ($6.3\ \text{fb}^{-1}$) in the signal region but compare only the variables where there is no discernible difference in shape between the background and signal+background hypotheses. We refer to such variables as "safe". The signal region is defined by the selections listed in Table 4.5.

[2]The analysis follows a blind procedure.

2. Look at the full data sample (62.8 fb^{-1}) in a "sideband" region that is background enriched. The sideband is defined by inverting a looser version of the signal region selections. This is illustrated in Figure 4.11. The exact set of selections is summarized in Table 4.11.

These methods are complementary. The first allows us to check the data-MC agreement for the exact topology and kinematics of the signal region, although for a restricted set of safe variables. The second method allows for a more detailed study of the main backgrounds with no restrictions on the variables.

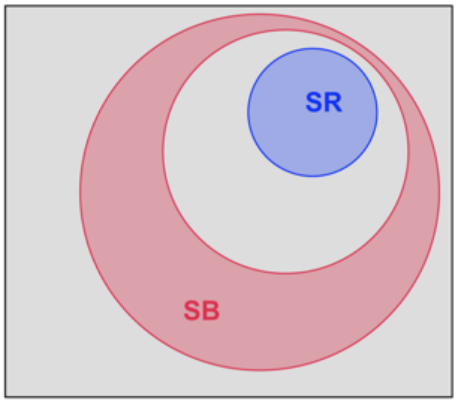


Figure 4.11: Simplified diagram showing the full phase space (grey box), signal region (SR, blue circle) and sideband (SB, red circle). Here, the SR and SB do not overlap since the latter is defined by inverting a looser version of the SR cuts [11].

Validation in the signal region with safe variables

When comparing data to MC in the signal region, it is crucial to examine only those variables that do not indicate the presence or absence of the signal process. This will typically be the case, for example, when looking at tag side variables that are independent of what happens on the signal side. Essentially, we want to focus on variables that are "safe" with respect to the signal process to avoid any bias in our analysis.

In order to decide whether or not a variable is "safe", we perform a statistical test of compatibility in shape between the background and signal+background distributions where the signal is injected according to the upper limits from the ARGUS Collaboration [25]. This is done using the so-called Kolmogorov test (KS test) [49].

If the KS test returns 1 or very close to 1, then, there is no discernible difference in shape, and in this case, we deem the variables as "safe" to look at in the data. The full set of KS test results are summarized in Table 4.10, and two examples of the input distributions and KS test results can be found in Figure 4.12.

Then, the following variables are identified as "safe":

Tag side variables

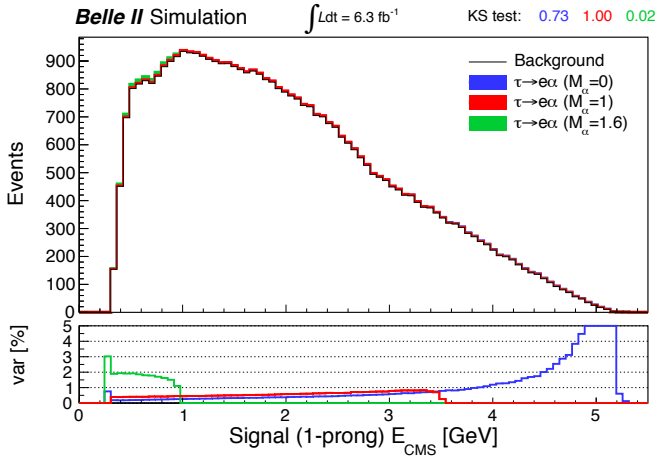
- E_{CMS}^{3prong} : Total energy in the CMS frame of the "3-prong" system.
- $M_{\text{Inv}}^{3\text{prong}}$: Invariant mass of the "3-prong" system.
- χ^2 probability: χ^2 probability of the "3-prong" vertex fit.

- $\#\gamma_{\text{3prong}}$: Number of photons in the event on the tag side ("3-prong").
- $\#\pi^0_{\text{3prong}}$: Number of neutral pions in the event on the tag side ("3-prong").
- p_T leading: Transverse momentum (p_T) of the leading track on the tag side.
- p_T sub-leading: Transverse momentum (p_T) of the sub-leading track on the tag side.
- p_T third: Transverse momentum (p_T) of the third track on the tag side.

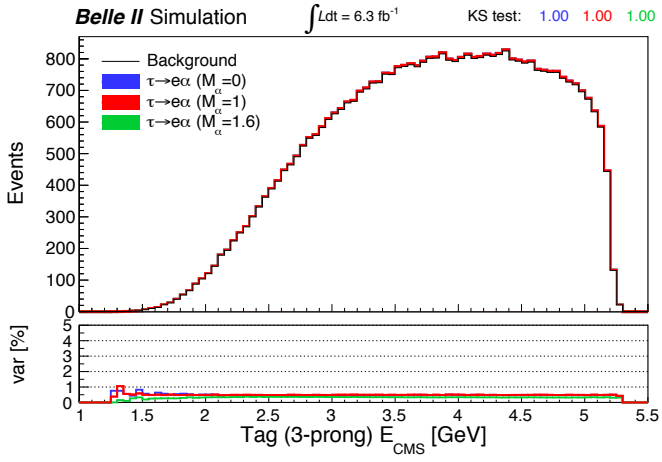
Event shape/kinematic variables:

- *Thrust*: Thrust magnitude, as calculated in Equation 4.1.
- missing momentum θ: Polar angle of the missing momentum of the event.

If we observe any data-MC disagreement, as there is no noticeable distinction between the background and signal+background distributions in MC, this disagreement can be attributed to either MC mismodeling or processes that are not present in the simulation.



(a)



(b)

Figure 4.12: Comparison of the expected background and signal+background distributions in the $\tau \to e\nu\bar{\nu}$ channel for (a) the signal-side E_{CMS} and (b) the tag-side E_{CMS}. The MC is normalized to 10% of the data luminosity (6.3 fb^{-1}), and the signal is injected according to the current upper limits from the ARGUS Collaboration. The results from the Kolmogorov test are indicated in each plot on the top right. The lower panel shows the % variation of the signal+background with respect to the background-only distribution [11]

Variable	Channel	M_α (GeV/c^2) 0	0.5	0.7	1.0	1.2	1.4	1.6	Safe to look at?
E_{CMS}^{1prong}	e	0.73	0.92	1.00	1.00	0.81	0.11	0.02	No
	μ	0.06	0.24	0.65	0.67	0.03	0.00	0.00	
E_{CMS}^{3prong}	e	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Yes
	μ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
M_{Inv}^{3prong}	e	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Yes
	μ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
$Thrust$	e	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Yes
	μ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
E_{CMS}^{Total} of event	e	0.88	0.99	1.00	1.00	1.00	0.68	0.47	No
	μ	0.18	0.48	0.92	1.00	0.56	0.05	0.05	
χ^2 probability	e	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Yes
	μ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
Missing momentum	e	1.00	1.00	1.00	1.00	1.00	0.94	0.78	No
	μ	0.99	0.98	0.99	1.00	0.98	0.23	0.18	
Missing momentum θ	e	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Yes
	μ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
Missing M^2	e	0.96	1.00	1.00	1.00	1.00	0.97	0.97	No
	μ	0.31	0.74	1.00	1.00	0.81	0.35	0.58	
$\#\gamma_{3prong}$	e	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Yes
	μ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
$\#\pi^0_{3prong}$	e	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Yes
	μ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
Signal p_T^{1prong}	e	0.76	0.93	1.00	1.00	0.94	0.23	0.04	No
	μ	0.10	0.30	0.73	0.99	0.21	0.00	0.00	
p_T leading	e	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Yes
	μ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
p_T sub-leading	e	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Yes
	μ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
p_T third	e	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Yes
	μ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	

Table 4.10: Kolmogorov test results for $\tau \to e\nu\bar{\nu}$ and $\tau \to \mu\nu\bar{\nu}$ channels [11].

Validation in sideband

For a sidebands region, the following conditions are desirable:

1. An enriched region with the main backgrounds, mainly $q\bar{q}$ and $\tau^{\pm} \to \pi^{\pm}(n\pi^0)\nu$ with a charged pion faking a lepton.
2. Orthogonality with the signal region (SR), selecting events kinematically similar to the SR events.
3. Minimal signal leaking into the SB region.

Therefore, events entering the sideband must satisfy most (but not all) of the SR selections. Events are explicitly vetoed if they pass the complete SR selections with looser neutrals cuts. The selections that define the sideband are summarised in Table 4.11.

Selection criteria	$\tau \to e\nu\bar{\nu}$	$\tau \to \mu\nu\bar{\nu}$
Event	$0.9 <$ thrust < 0.99	$0.9 <$ thrust < 1.0
Signal (1 prong)	electronID > 0.9 in the acceptance of lepton e-ID performance measurements	muonID > 0.9 in the acceptance of lepton μ-ID performance measurements
Tag (3 prong) $\tau \to 3h(\nu)$	clusterE/p ≤ 0.8 vertex χ^2 probability > 0 leading $p_t > 0.69$ GeV/c sub-leading $p_t > 0.29$ GeV/c third $p_t > 0.08$ GeV/c $1.2 < E^{\text{3prong}}_{CMS} < 5.3$ GeV	clusterE/p ≤ 0.8 vertex χ^2 probability > 0 leading $p_t > 0.47$ GeV/c sub-leading $p_t > 0.17$ GeV/c third $p_t > 0.04$ GeV/c $1.1 < E^{\text{3prong}}_{\text{CMS}} < 5.3$ GeV
SR orthogonality	veto events that fall into the SR with looser neutrals selections: $\#\pi^0_{\text{1prong}} = 0, \quad \#\pi^0_{\text{3prong}} \leq 2$ $\#\gamma_{\text{1prong}} \leq 1, \quad \#\gamma_{\text{3prong}} \leq 1$	veto events that fall into the SR with looser neutrals selections: $\#\pi^0_{\text{1prong}} = 0, \quad \#\pi^0_{\text{3prong}} \leq 2$ $\#\gamma_{\text{1prong}} \leq 1, \quad \#\gamma_{\text{3prong}} \leq 1$

Table 4.11: Summary of the sideband selection criteria [11]. The signal region (SR) refers to the selections defined earlier in Table 4.5.

4.6.2 Corrections

An additional correction is applied to MC in this step:

- The π^0 reconstruction efficiency. These are taken from the tables provided by the Belle II Neutrals Performance Group [50].

For data, we apply to each track a global momentum scale correction. This accounts for imperfections in the magnetic field map used for data reconstruction. When applying this correction, we follow the recommendations of the Belle II Tracking Group [51].

4.6.3 Improvements

Photon isolation

There are several kinds of background ECL clusters present in data that are difficult to model precisely in MC. These include hadronic split-off clusters, several clusters originating from a single low momentum track, and those associated with beam background. From preliminary Data-MC comparisons on the number of photons and neutral pions in the tag side, it was clear that we needed to solve first this issue, as can be seen in Figure 4.14 (top).

In order to suppress these kinds of background clusters, we now include two additional requirements to the photon reconstruction criteria described in Section 4.2.1. We use the minimum distance between a cluster and the nearest track (clusterC2TDist) to ensure the cluster isolation of our photon candidates. This helps to reject hadronic split-offs and multiple clusters from a single track that tends to mimic low-energy photons, as illustrated in Figure 4.13.

- [clusterC2TDist $>$ 40 cm] or [E(γ) $>$ 400 MeV] .

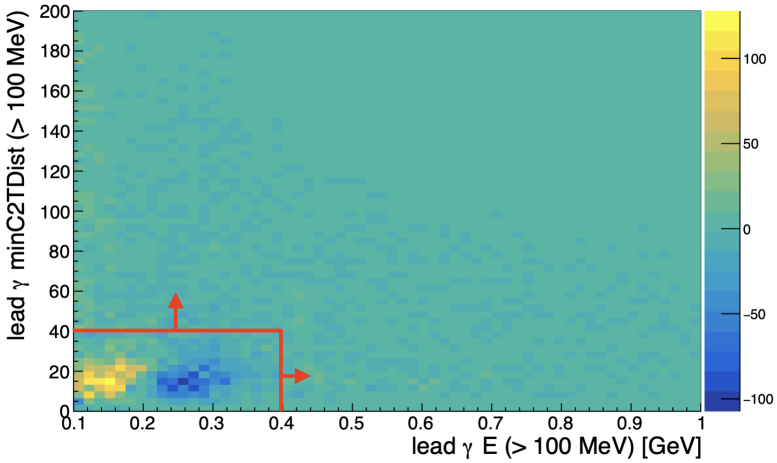


Figure 4.13: Data−MC distribution of the isolation vs. energy for the leading (highest energy) photon. The photon isolation requirement is indicated by the red box and arrows [11].

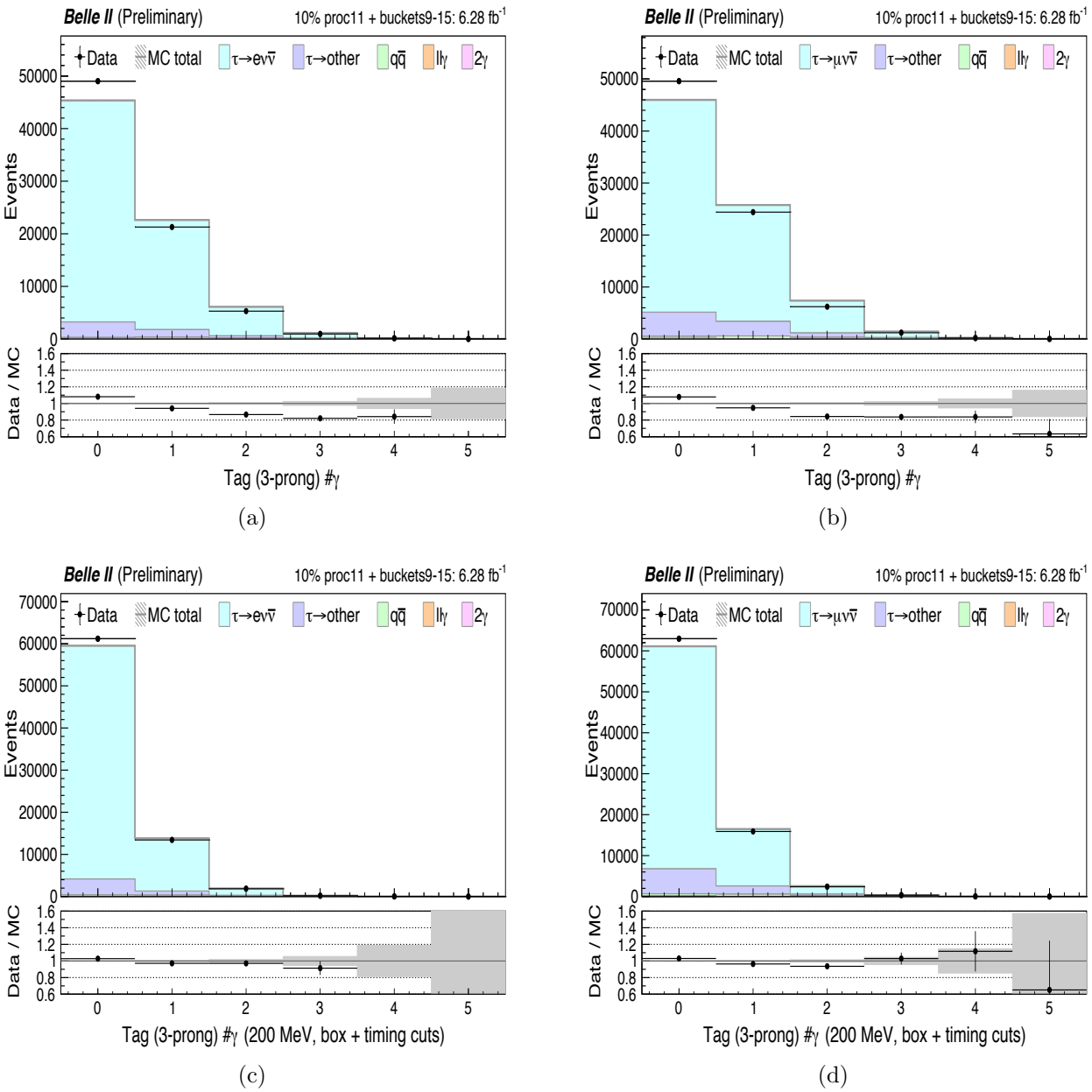


Figure 4.14: Data vs MC distributions of the number of photons on the tag-side (top) before and (bottom) after including the photon isolation and timing criteria. The (left) electron and (right) muon channel signal regions are shown. The error band includes the statistical and trigger efficiency related uncertainties [11].

Timing

In addition, very out-of-time clusters from beam background are rejected with a loose timing requirement, which follows the official recommendation of the Belle II Neutrals Performance Group:

- abs(clusterTiming) < 200 ns .

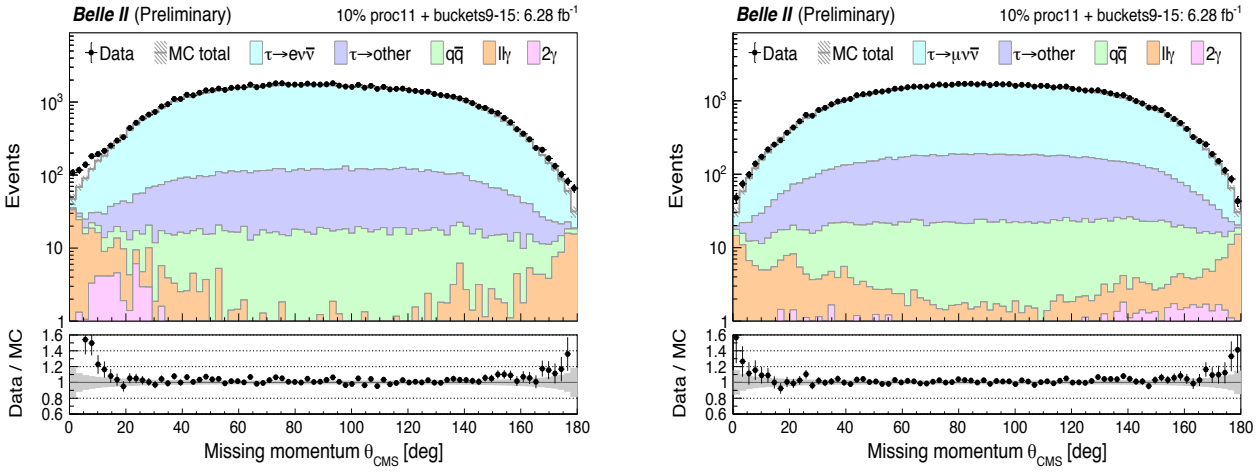
The MC-data comparison distributions considering both the photon isolation and timing criteria are shown at the bottom of Figure 4.14.

Missing momentum polar angle

We observe an excess in data when the missing momentum vector is pointing very forward or very backward. This can be seen in Figure 4.15.

This excess is most likely due to 2-photon processes that are not included or are miss-modeled in our MC samples. This kind of excess has been reported recently by other Belle II tau physics analyses (e.g. tau mass measurement) [43], and we propose to follow a similar strategy by requiring events to satisfy:

- $20° <$ Missing momentum θ in CMS $< 160°$.



(a) (b)

Figure 4.15: Data vs MC distributions of the missing momentum polar angle in CMS for (a) the electron channel and (b) the muon channel signal regions. The error band includes the statistical and trigger efficiency related uncertainties [11].

Impact on the selection

The selection of this analysis will remain unchanged. The reason is that after the incorporation of the photon isolation, timing, and missing momentum polar angle requirements, the optimized cuts in Table 4.5 do not charge significantly.

However, after the new requirements, a significant improvement in signal efficiency has been obtained [3] as can be seen in Table 4.12.

[3]Now, more events with low-quality photons that do not meet our requirements are considered as signal.

		e			μ	
Sample $\tau \to \ell + \alpha$	Events	Efficiency	Purity	Events	Efficiency	Purity
$M_\alpha = 0$	551115	18.13%	-	679090	22.34%	
$M_\alpha = 0.5$	556678	18.31%	-	675761	22.23%	
$M_\alpha = 0.7$	562005	18.49%	-	676363	22.25%	
$M_\alpha = 1.0$	574226	18.89%	-	668749	22.00%	
$M_\alpha = 1.2$	581010	19.11%	-	652625	21.47%	
$M_\alpha = 1.4$	546159	17.97%	-	598855	19.70%	
$M_\alpha = 1.6$	389773	12.82%	-	373125	12.27%	
$\tau \to l\nu\bar{\nu}$	547103	17.50%	95.93%	643186	21.08%	91.93%

Table 4.12: New efficiencies and purities for signal samples [11].

Impact on the limit

After applying the π^0 efficiency correction and improvements, which were described in this and the previous section, changes in the upper limit estimation are negligible.

4.6.4 Data-MC plots in signal region

In this section, we show the data/MC agreement in the signal regions. As described in Section 4.6, only 10% of the total data is considered. All of the corrections described in Section 4.4.2 and 4.6.2, and the new selection criteria of Section 4.6.3 are included.

In these plots, the MC error band includes the following uncertainties summed in quadrature:

- MC statistical uncertainty.
- Trigger efficiency correction uncertainty.
- Lepton ID efficiency correction uncertainties.
- $\pi \to \ell$ fake-rate correction uncertainties.
- π^0 efficiency correction uncertainty.

The uncertainty in data is the combination of statistical and momentum scale uncertainty.

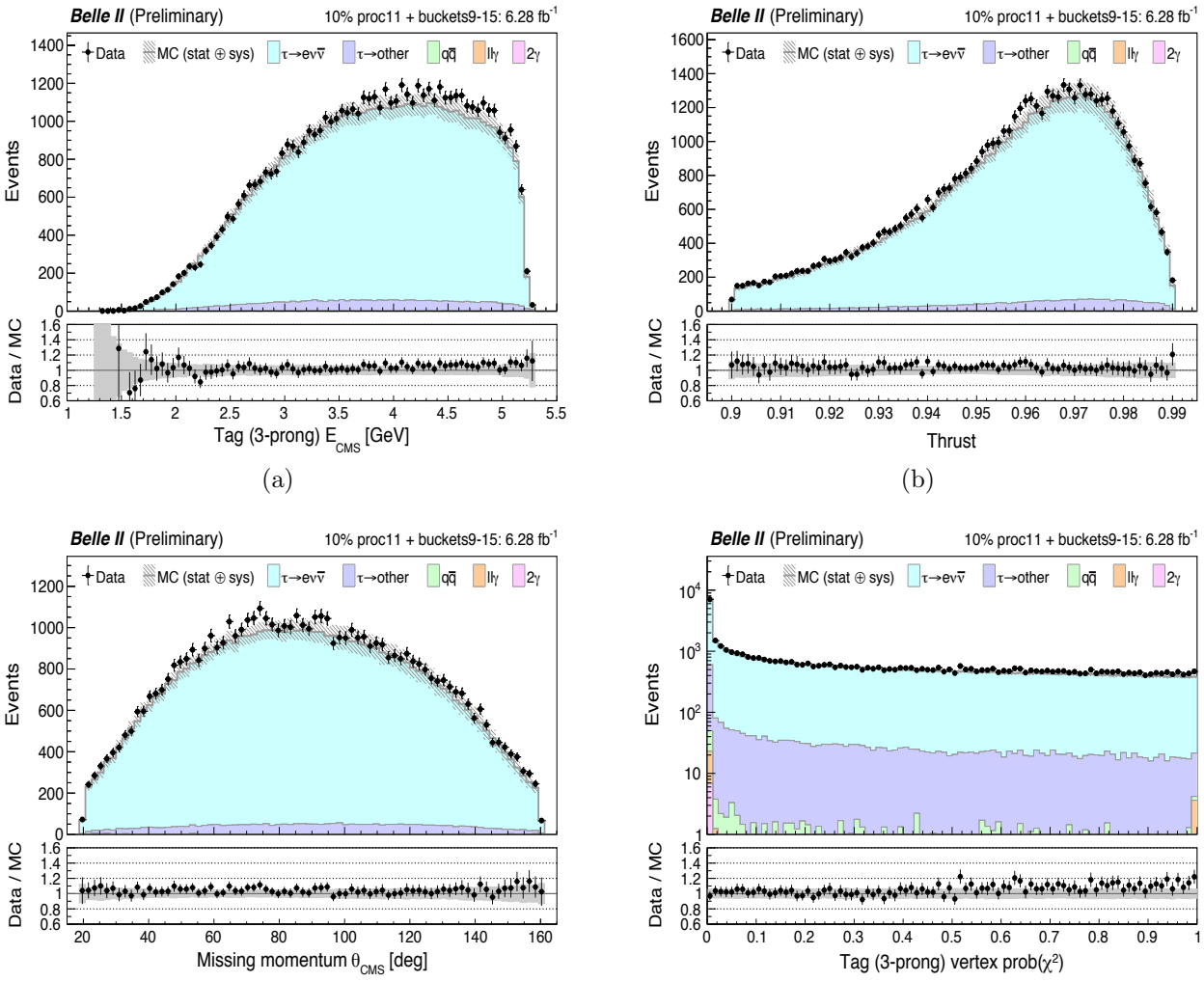
The electron channel distributions are shown in figures 4.16-4.17. The muon channel distributions are shown in figures 4.18-4.19

4.6.5 Data-MC plots in side-bands

In this section, we show the data/MC agreement in the sidebands. As described in Section 4.6, the full data sample is considered.

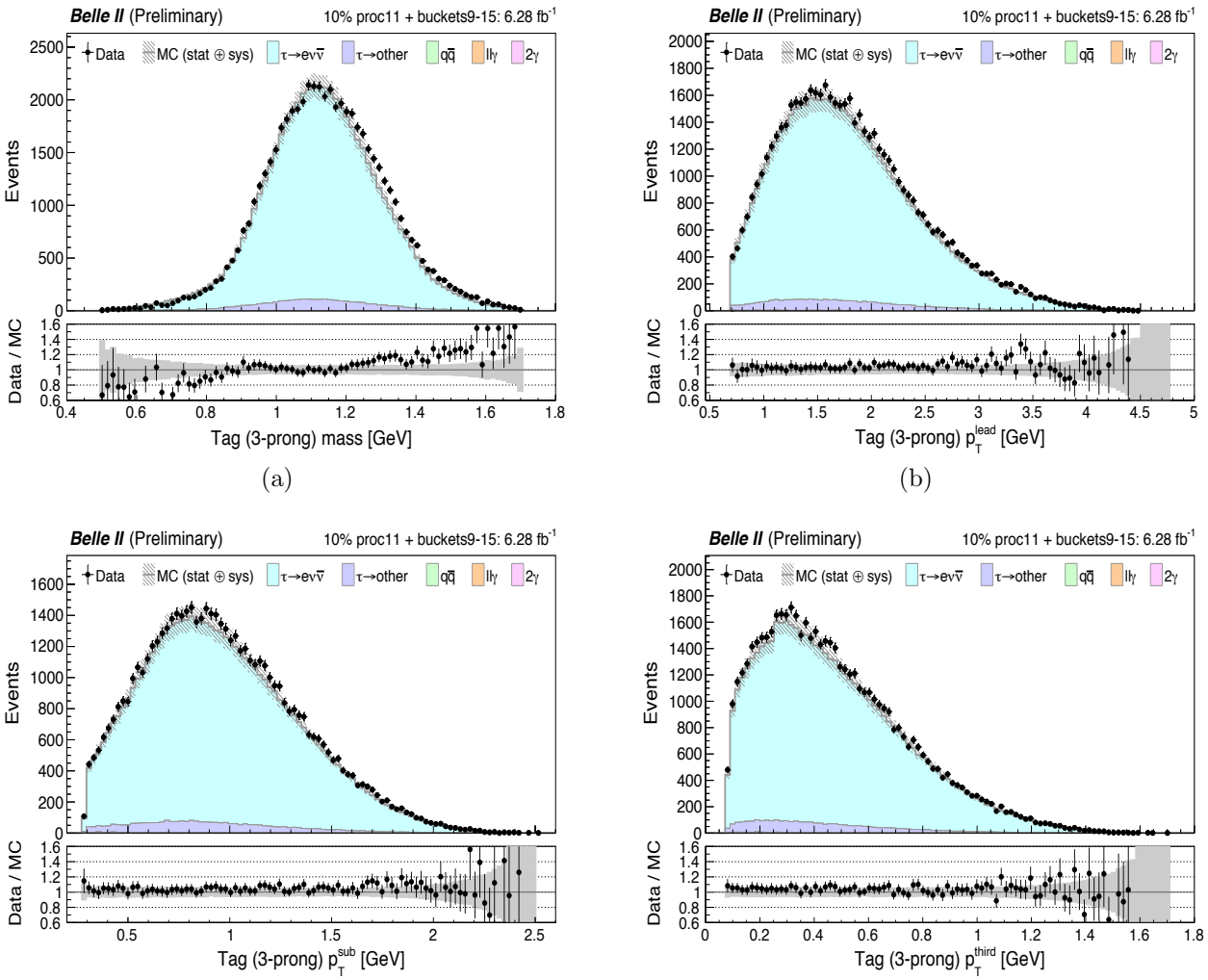
As in signal region strategy, all of the corrections described in Section 4.4.2 and 4.6.2, and the new selection criteria of Section 4.6.3 are included, and we use the same uncertainties calculation.

The electron channel distributions are shown in Figures 4.20-4.21. The muon channel distributions are shown in Figures 4.22-4.23.



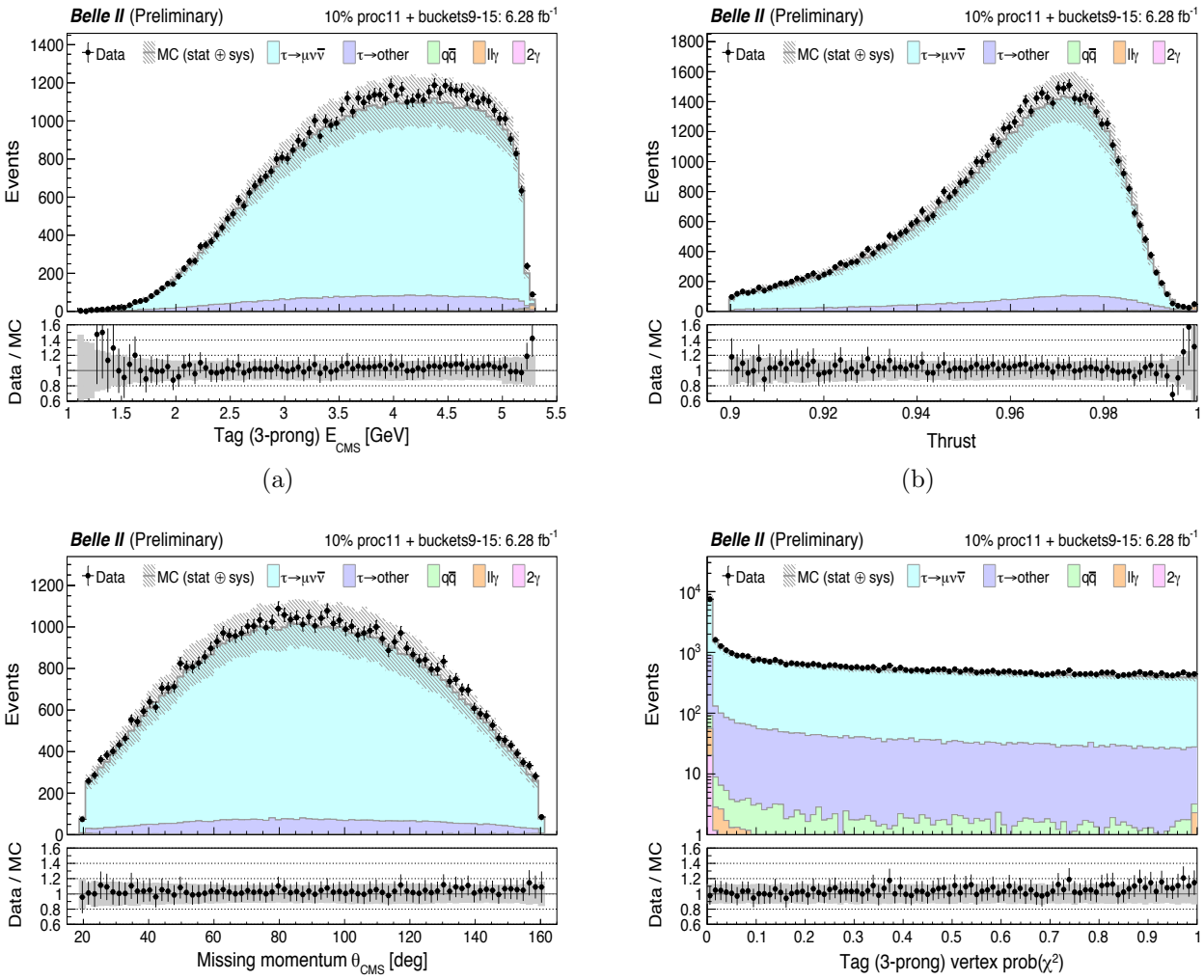
(a) (b) (c) (d)

Figure 4.16: Data vs MC distributions in the electron channel signal region for (a) the tag E_{CMS}, (b) thrust, (c) missing momentum polar angle in CMS, and (d) tag vertex χ^2 probability [11]. The MC error band includes all the uncertainties listed at the start of Section 4.6.4.



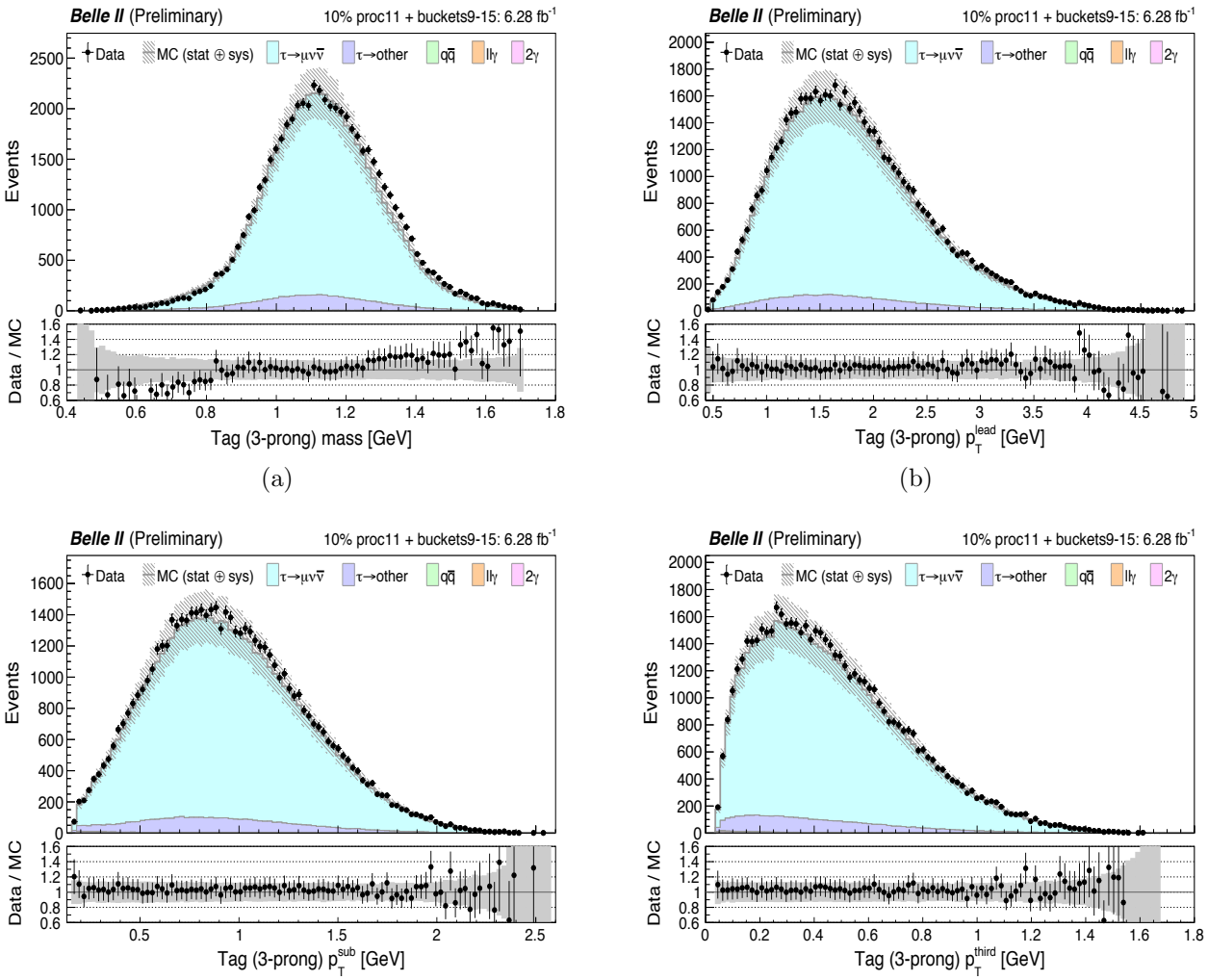
(a) (b) (c) (d)

Figure 4.17: Data vs MC distributions in the electron channel signal region for (a) the tag mass, (b) tag leading track p_T, (c) tag subleading track p_T and (d) tag third track p_T [11]. The MC error band includes all the uncertainties listed at the start of Section4.6.4.



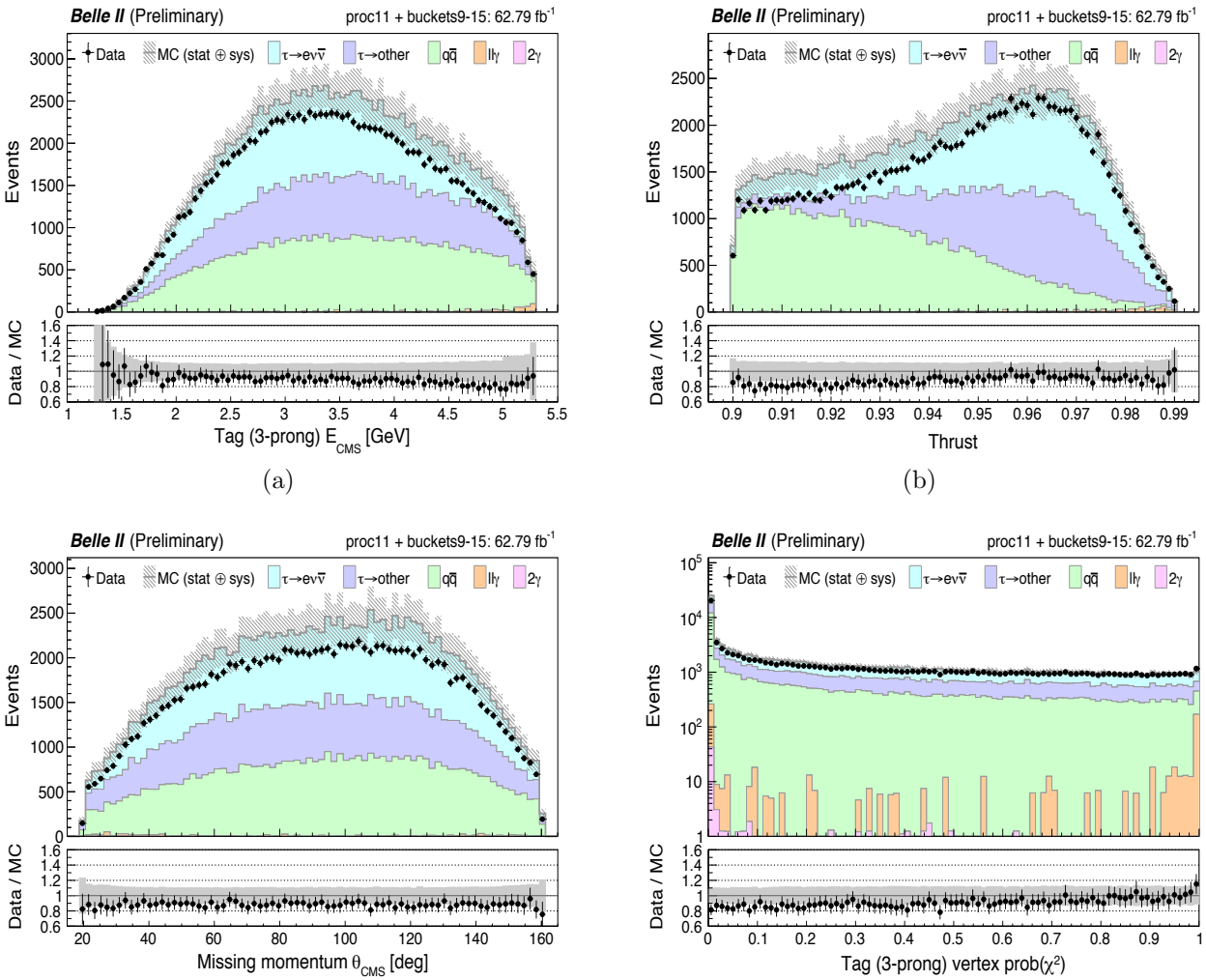
(a) (b) (c) (d)

Figure 4.18: Data vs MC distributions in the muon channel signal region for (a) tag E_{CMS}, (b) thrust, (c) missing momentum polar angle in CMS, and (d) tag vertex χ^2 probability [11]. The MC error band includes all the uncertainties listed at the start of Section 4.6.4.



(a) (b) (c) (d)

Figure 4.19: Data vs MC distributions in the muon channel signal region for (a) the tag mass, (b) tag leading track p_T, (c) tag subleading track p_T and (d) tag third track p_T [11]. The MC error band includes all the uncertainties listed at the start of Section 4.6.4.



(a) (b) (c) (d)

Figure 4.20: Data vs MC distributions in the electron channel sideband for (a) tag E_{CMS}, (b) thrust, (c) missing momentum polar angle in CMS, and (d) tag vertex χ^2 probability [11]. The MC error band includes all the uncertainties listed at the start of Section 4.6.4.

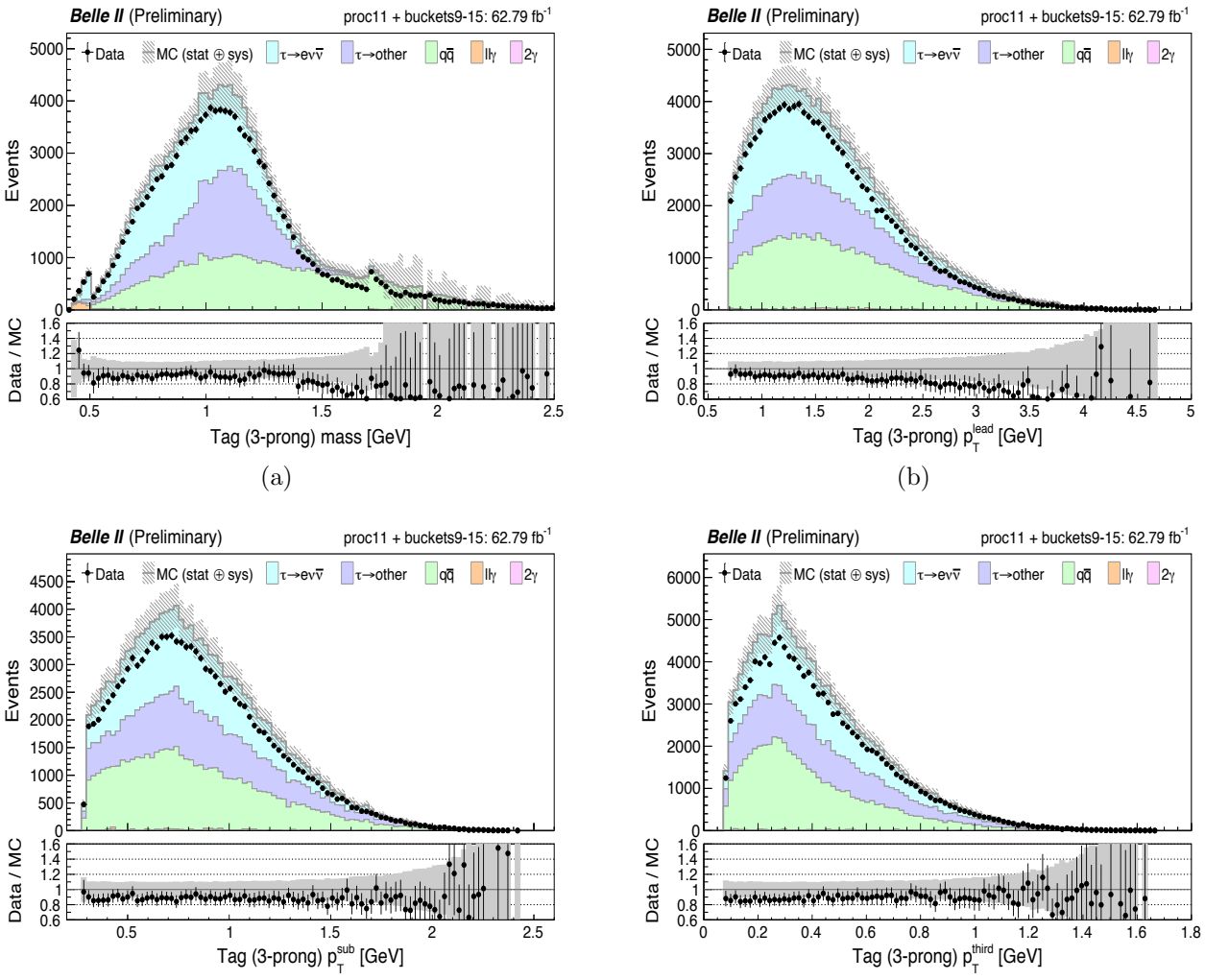
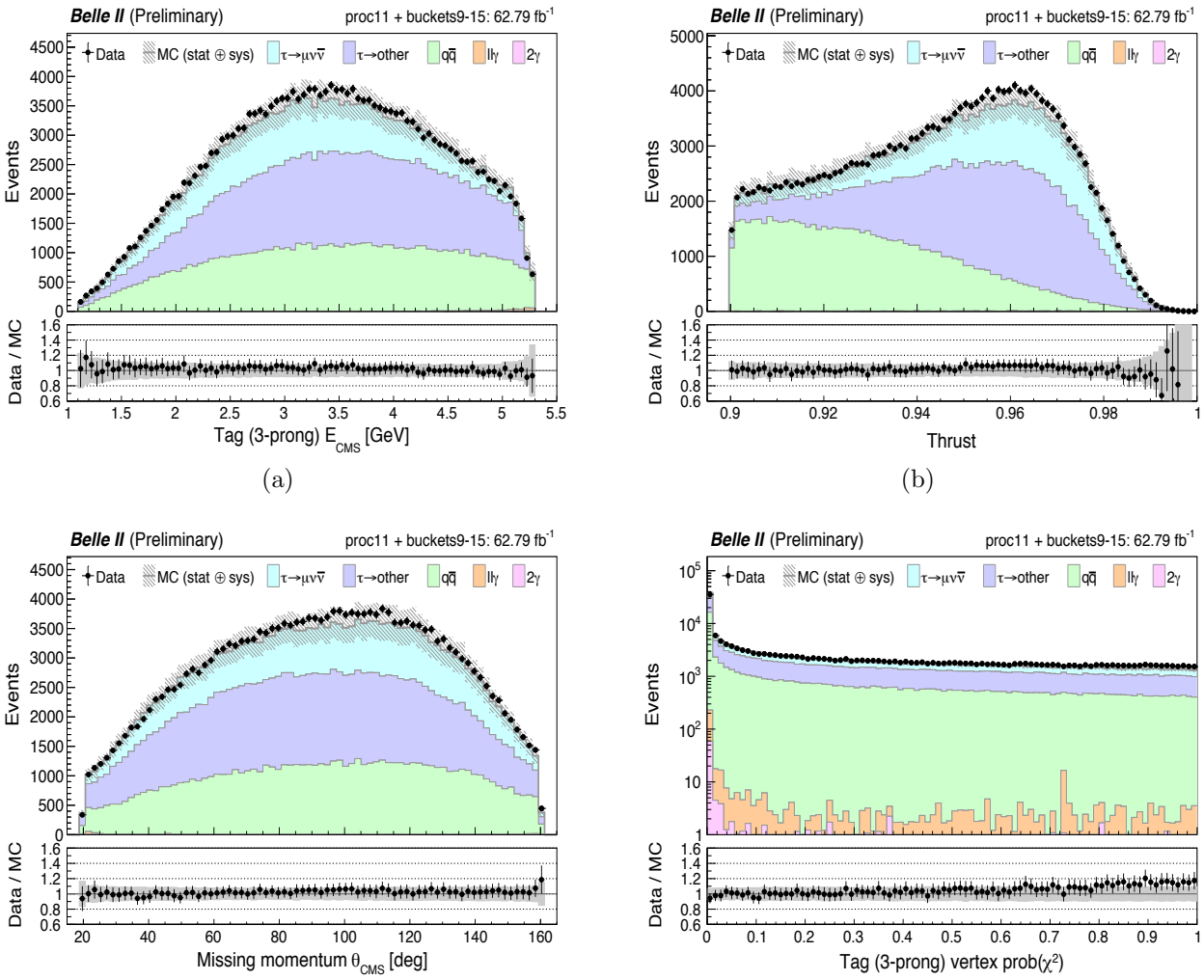
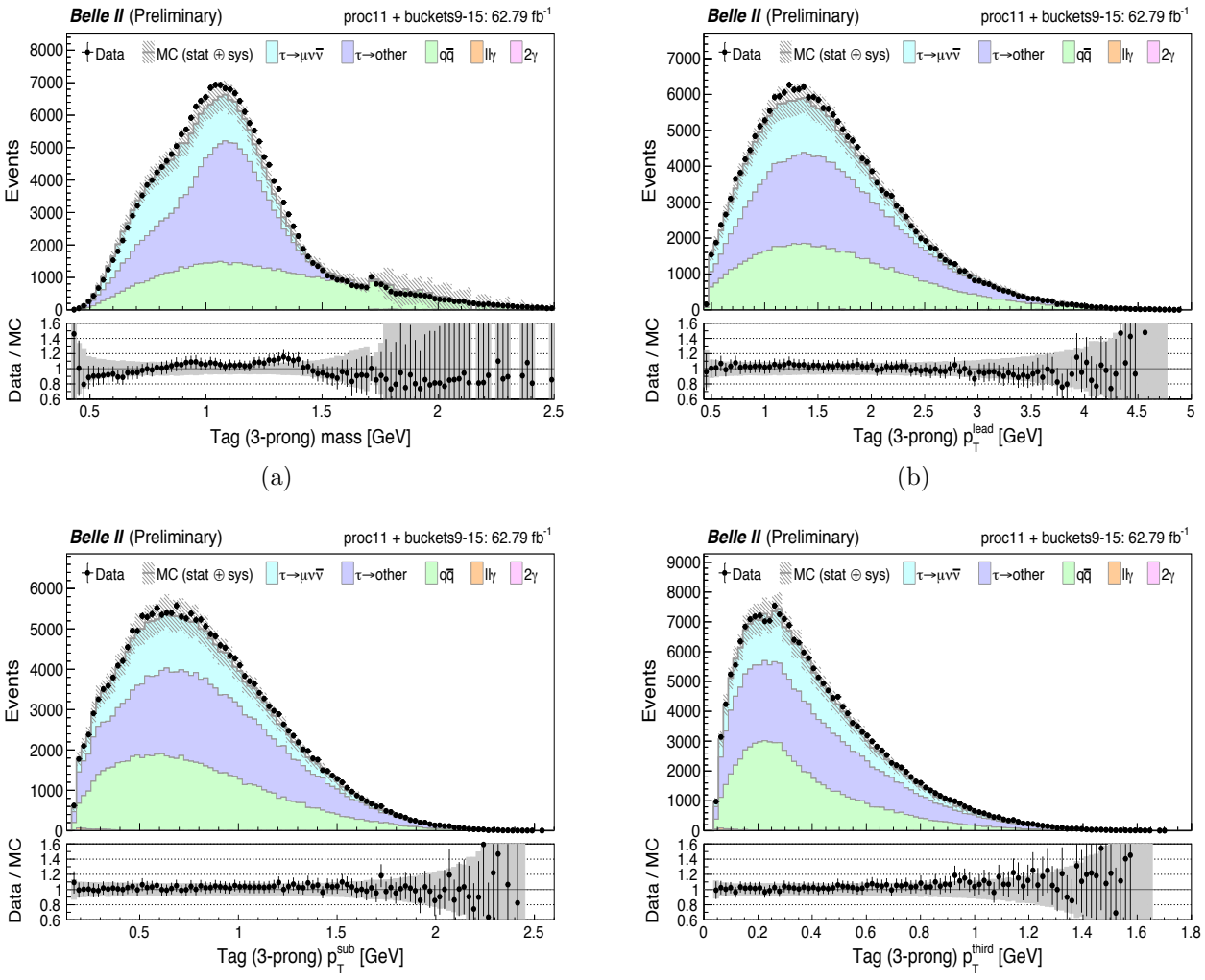


Figure 4.21: Data vs MC distributions in the electron channel sideband for (a) the tag mass, (b) tag leading track p_T, (c) tag subleading track p_T and (d) tag third track p_T [11]. The MC error band includes all the uncertainties listed at the start of Section 4.6.4.



(a) (b) (c) (d)

Figure 4.22: Data vs MC distributions in the muon channel sideband for (a) tag E_{CMS}, (b) thrust, (c) missing momentum polar angle in CMS, and (d) tag vertex χ^2 probability [11]. The MC error band includes all the uncertainties listed at the start of Section 4.6.4.



(a) (b) (c) (d)

Figure 4.23: Data vs MC distributions in the muon channel sideband for (a) the tag mass, (b) tag leading track p_T, (c) tag subleading track p_T and (d) tag third track p_T [11]. The MC error band includes all the uncertainties listed at the start of Section 4.6.4.

$q\bar{q}$ normalization

From the previous figures, we observe overall good data/MC agreement in both signal regions and in the μ-channel sideband. However, as can be observed in Figure 4.24, there is an overall deficit in the e-channel sideband. From other ongoing analyses in Belle II (e.g. m_τ measurement [43]) and the past experience from τ-physics analyses performed at Belle and BABAR, the $q\bar{q}$ background is not well reproduced by MC. Luckily, the discrepancy is an overall deficit.

To study this deficit, we define a "control region" enriched in $q\bar{q}$ by the requirement that the tag (3-prong) mass in the sideband is greater than m_τ. Figure 4.26(a) shows the mass distribution before this requirement while (b) shows the composition of the control region.

From the control region, we can derive a data-driven $q\bar{q}$ normalisation factor (NF) as follows:

$$\mathrm{NF} = \frac{N_{q\bar{q}}^{data}}{N_{q\bar{q}}^{MC}} \,, \tag{4.10}$$

where

$$N_{q\bar{q}}^{data} = N^{data} - N_{non-q\bar{q}}^{MC} \,. \tag{4.11}$$

Following this procedure, we compute $\mathrm{NF} = 0.6748 \pm 0.0691$.

After applying the data-driven normalization factor to the $q\bar{q}$ contribution in MC, we observe significantly better data-MC agreement in the sideband. This is illustrated in Figure 4.25.

Although $q\bar{q}$ has a significant contribution in the sideband, its contribution in the signal regions is very small ($< 0.1\%$). Hence, we do not expect a large impact on the UL due to the $q\bar{q}$ normalization factor. Nevertheless, we checked this, and confirmed that the impact on the UL is negligible.

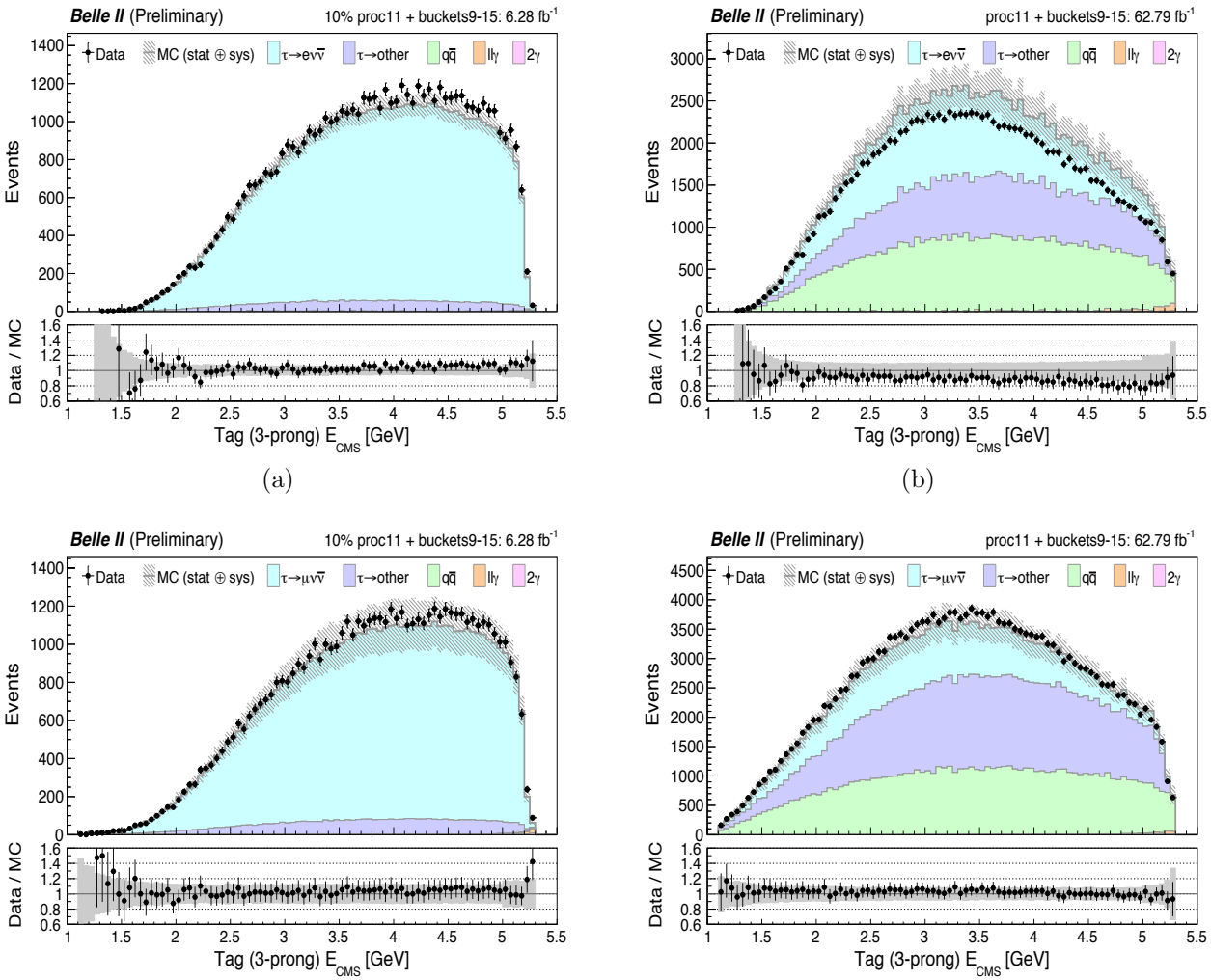
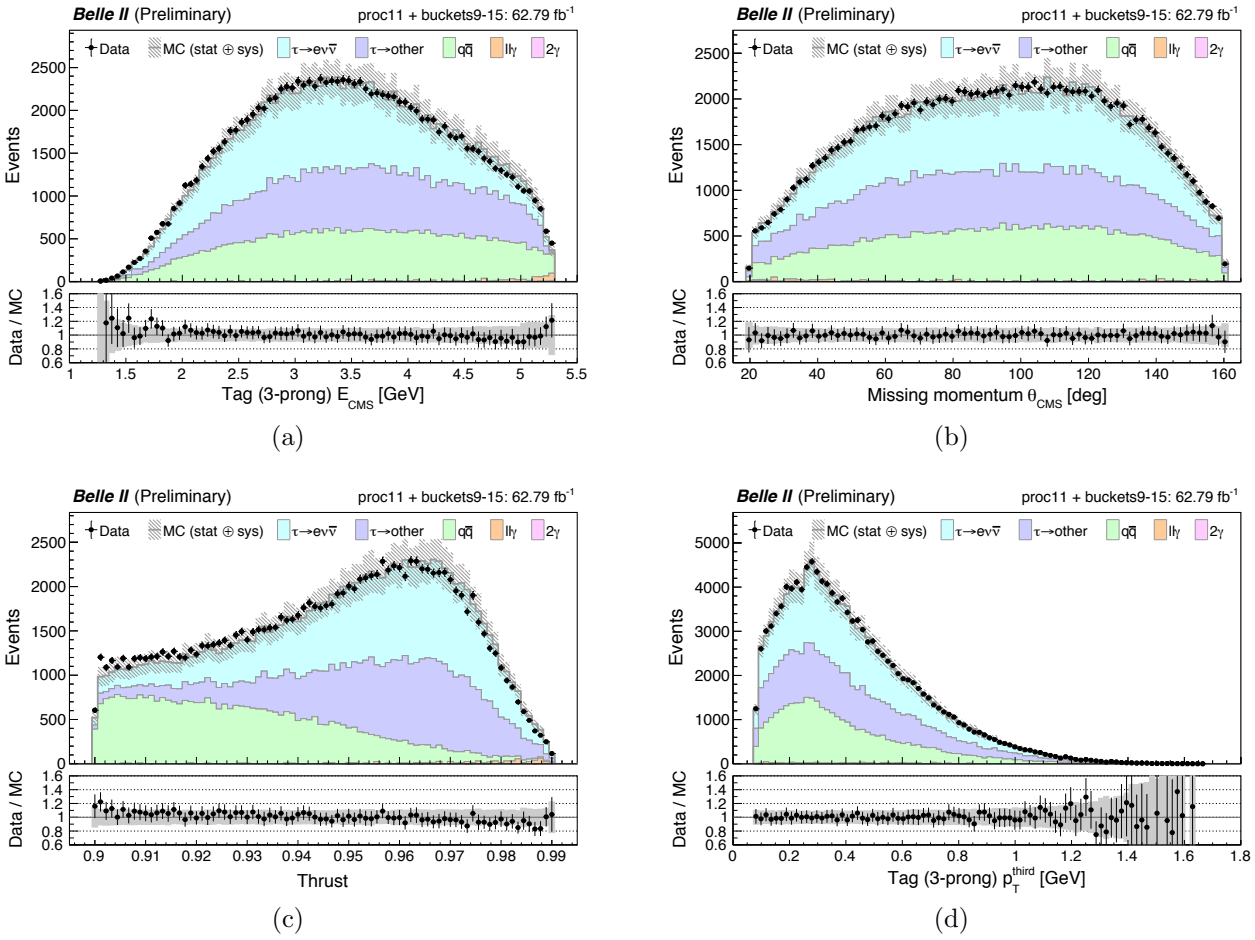
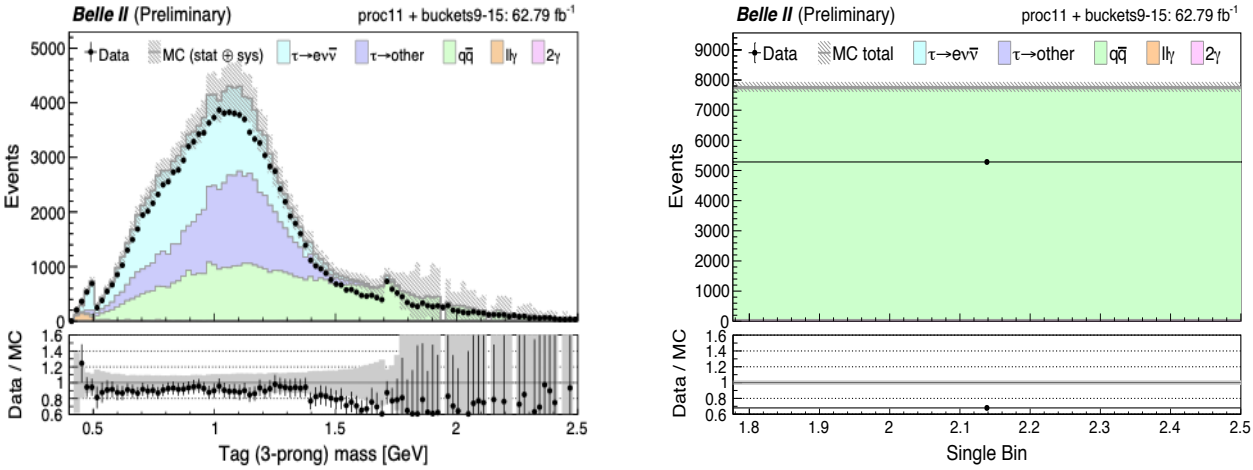


Figure 4.24: Data vs MC distributions of the tag E_{CMS} in the (a) e-channel signal region, (b) e-channel sideband, (c) μ-channel signal region and (d) μ-channel sideband [11]. The MC error band includes all the uncertainties listed at the start of Section 4.6.4.



(a) (b) (c) (d)

Figure 4.25: Data vs MC distributions in the e-channel sideband after applying the data-driven normalization factor to the $q\bar{q}$ contribution in MC (green). Here is shown (a) the tag E_{CMS}, (b) missing momentum polar angle in CMS, (c) thrust, and (d) tag third track p_T [11].



(a) (b)

Figure 4.26: (a) Distribution of the tag (3-prong) mass in the e-channel sideband.(b) Composition of the $q\bar{q}$ control region with the requirement of $m_{3-prong} > m_{\tau}$ [11].

4.6.6 Data vs MC in the partially unblinded data

In this section, we present the last check of data-MC validation.

Since we do not find any unexpected behavior in data-MC distributions in the previous strategies, then, we will show the data vs MC comparison for "unsafe" variables, looking at the partially unblinded data (10%, 6.28 fb^{-1}) in the signal regions.

The uncertainties in MC and data are calculated in the same way as described above in previous sections.

The x_{prf} distribution is shown in Figure 4.28 and 4.28. Other "unsafe" variables are shown in Figure 4.29.

Zeroth order polynomial fit check

Looking at Figures 4.27 and 4.28, data tends to be systematically above MC by a small amount, particularly in the electron channel.

We checked whether or not an overall normalization correction to the total SM MC (zeroth order polynomial fit to data/MC ratio) can resolve this issue. The p0 fit in the electron (muon) channel returns 1.040 ± 0.011 (1.019 ± 0.016). After applying these normalization corrections to the total MC distributions we get the plots shown in Figure 4.30.

The data/MC agreement is significantly improved. Thus, we expect the small overall excess to be absorbed by the SM template for which the normalization is allowed to float in the fit.

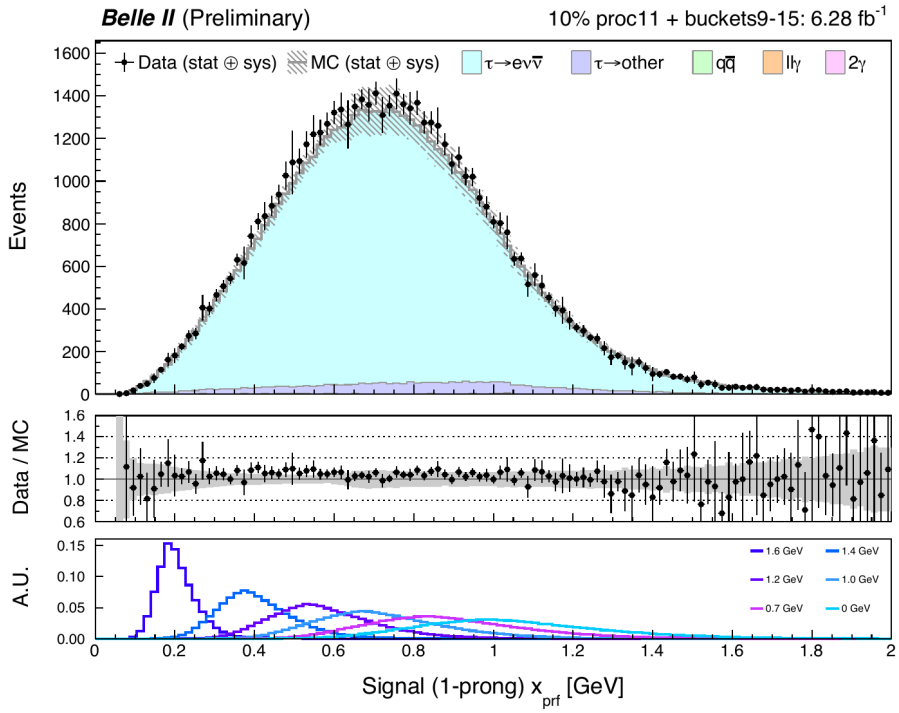


Figure 4.27: Data vs MC distribution of x_{prf} for the e-channel. The middle panel shows the data/MC ratio, with MC uncertainty shown by the grey band and data uncertainty by the error bars on the black points. The lower panel shows the signal distributions for different α masses normalized to 1 [11].

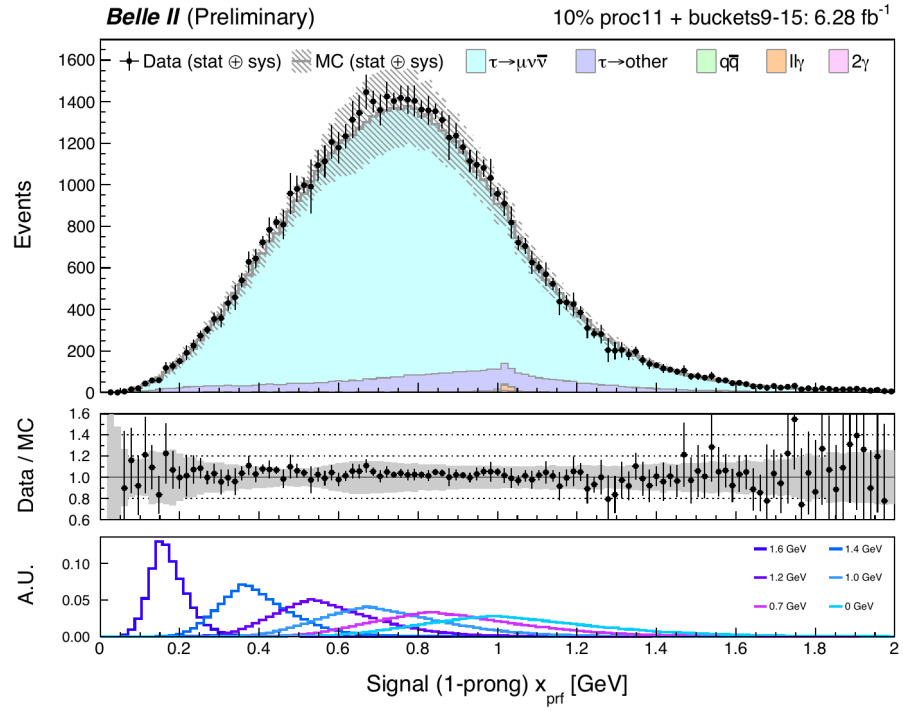


Figure 4.28: Data vs MC distribution of x_{prf} for the μ-channel. The middle panel shows the data/MC ratio, with MC uncertainty shown by the grey band and data uncertainty by the error bars on the black points. The lower panel shows the signal distributions for different α masses normalized to 1 [11].

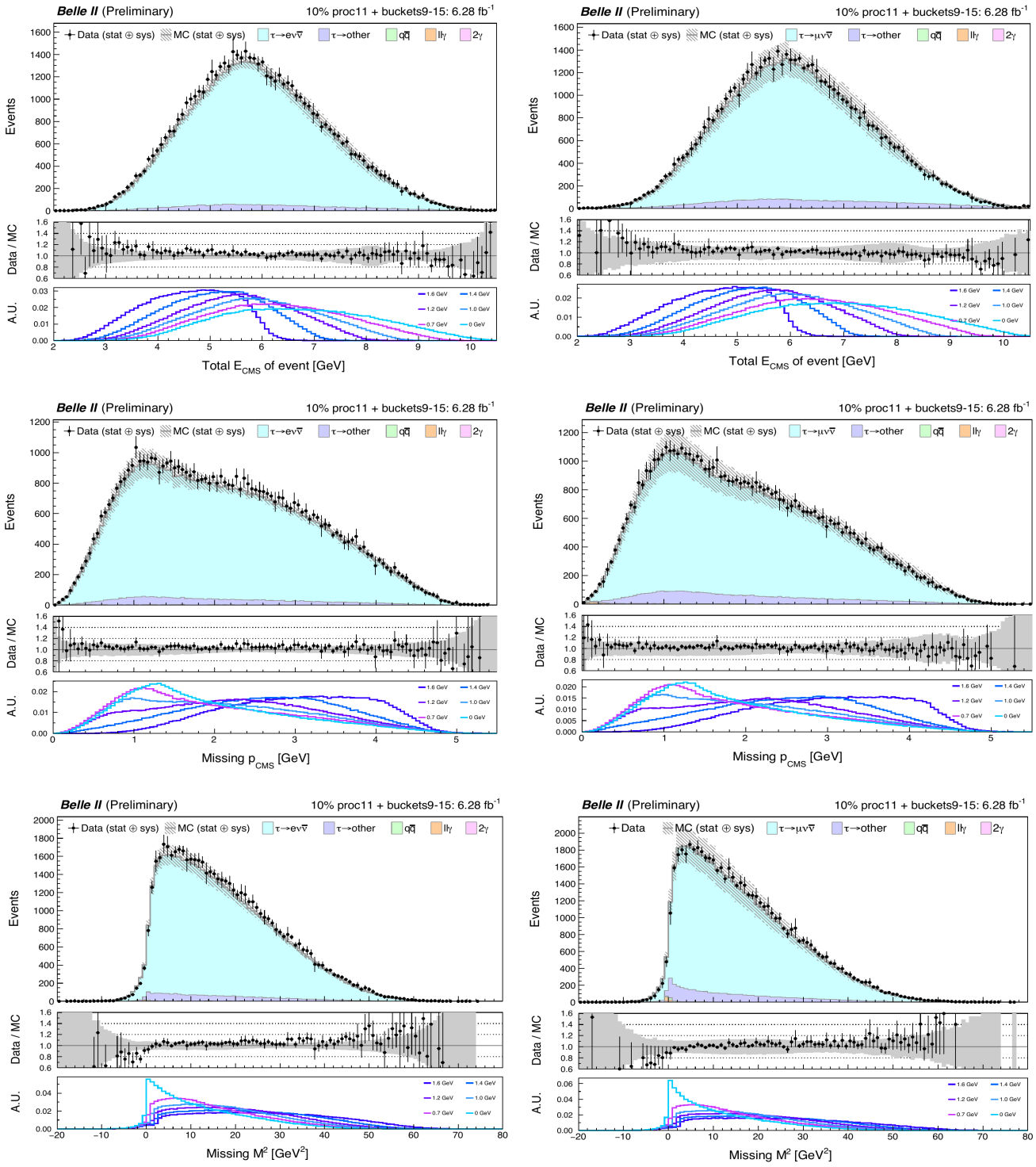
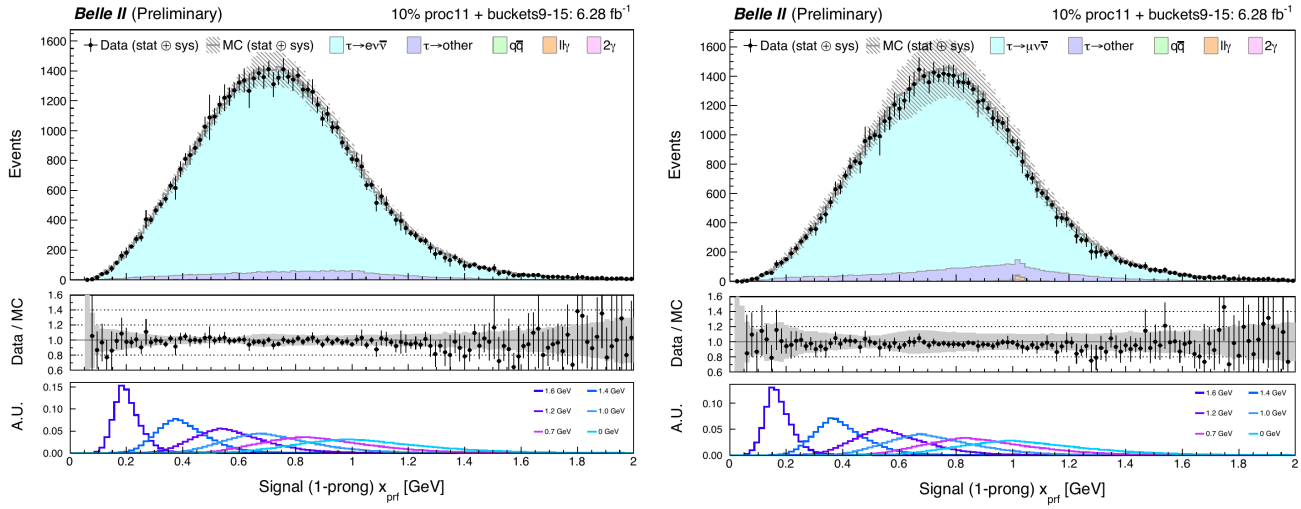


Figure 4.29: Data vs MC distribution in the (left) e-channel and (right) μ-channel for (top) the total E_{cms}, (middle) missing p_{cms} and (bottom) missing M^2. The middle panel shows the data/MC ratio, with MC uncertainty shown by the grey band and data uncertainty by the error bars on the black points. The lower panel shows the signal distributions for different α masses normalized to 1 [11].



(a) (b)

Figure 4.30: Data vs MC distributions of x_{prf}, after applying an overall normalization correction. The (a) electron and (b) muon channels are shown. The middle panel shows the data/MC ratio, with MC uncertainty shown by the grey band and data uncertainty by the error bars on the black points. The lower panel shows the signal distributions for different α masses normalized to 1 [11].

4.7 Final results

In this section, as we do not observe any significant excess in data compared to the SM prediction, we proceed to determine upper limits results following the methodology explained in Section 4.4.

First, we calculate an upper limit with the 10 % of unblinded data, and then, with the full data set corresponding to 62.8 fb^{-1}.

4.7.1 Upper limit on a Partially Unblinded Data

We use the same data sample as in Section 4.6.4 and 4.6.6. Figure 4.30 shows the Data-MC distribution of x_{prf}.

The procedure for the upper limit calculation was explained in Section 4.4. Figure 4.31 reports the expected upper limit under the background-only hypothesis (Brazilian plot) and the observed upper limit (red dots). The Brazilian plot shows the expected values for the upper limit (black dashed line) and the uncertainty bands correspond to $\pm 1\sigma$ and $\pm 2\sigma$, colored in green and yellow, respectively. We note the observed upper limits fluctuate within the expected band. The results published by the ARGUS collaboration (blue dots) are also included.

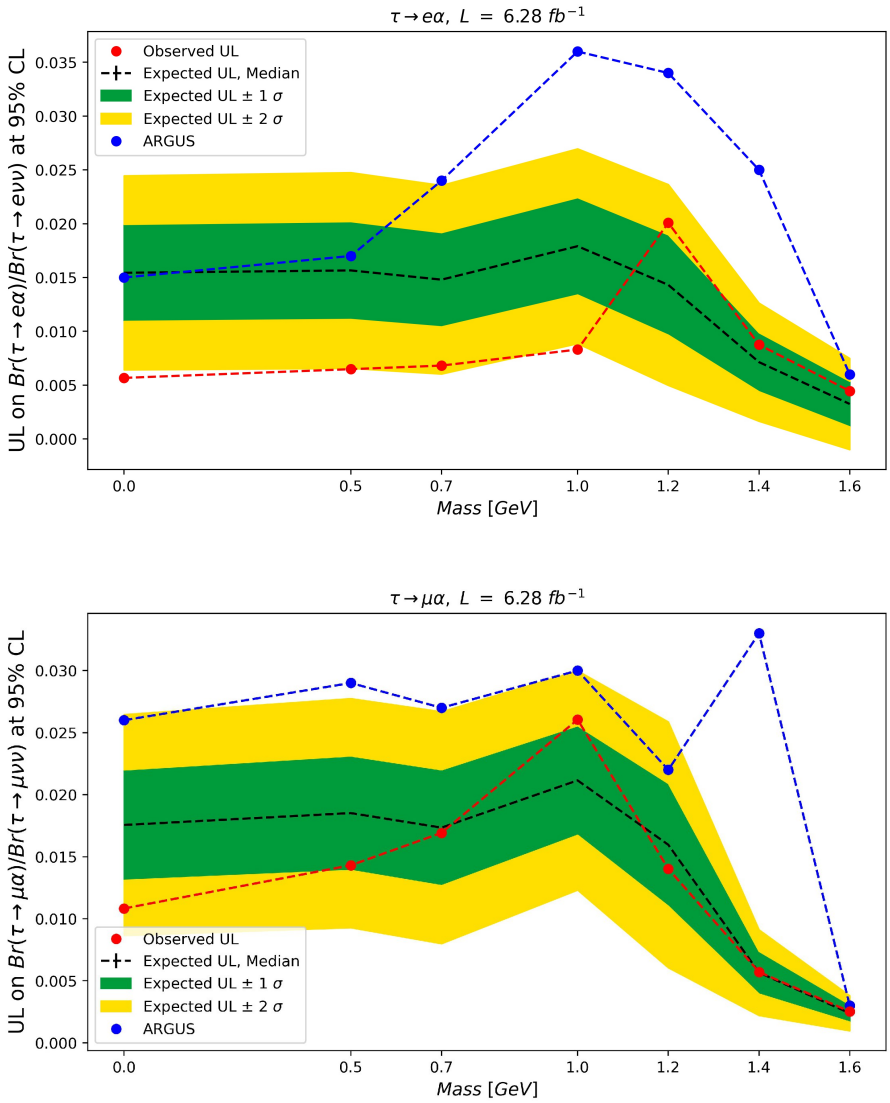


Figure 4.31: Results for a) the electron channel and b) the muon channel. The expected upper limits on the ratio $\mathcal{B}(\tau \to \ell\alpha)/\mathcal{B}(\tau \to \ell\nu\bar{\nu})$ at 95% C.L. are shown as a black dashed line, surrounded by the $\pm 1\sigma$ (green) and $\pm 2\sigma$ (yellow) bands. The observed upper limits are shown in red. The ARGUS collaboration results are shown in blue [11].

4.7.2 Box opening

After the previous validations, we proceed to look at the full dataset.

In this step, we notice that the likelihoods of the nuisance parameters showed instabilities, e.g. double minima or discontinuities in the likelihood function when it is profiled over each nuisance parameter. We observe that the fit is unstable if a shape variation that enters the likelihood function has a) large statistical fluctuations in the upward and the downward variations, or b) the upward and downward variations are very asymmetric or have the same sign. An essential component in ensuring that the likelihood fit is well-behaved involves smoothing [52] and symmetrizing the upward and the downward variations. For the latter, a conservative approach is to compare the upward and the downward variations in each bin with respect to the nominal template and take the largest of the two variations as the symmetric $\pm 1\sigma$ uncertainty for the bin in question. When the likelihood is profiled over each nuisance parameter, this

procedure removes double minima or discontinuities in the likelihood function.

Then, it was necessary to make the following improvements:

- increase the amount of the background MC samples listed in Table 4.1 to 1ab^{-1};
- re-bin the trigger efficiency correction to reduce the statistical fluctuations in low and high regions of x_{prf} ;
- smooth the relative up and down systematic variations caused by lepton ID efficiency, fake rate, and π^0 efficiency corrections;
- symmetrize the templates of the up and downward variations for all nuisance parameters.

Figure 4.32 exemplifies the procedure by illustrating the electron ID correction before and after the smoothing and symmetrization.

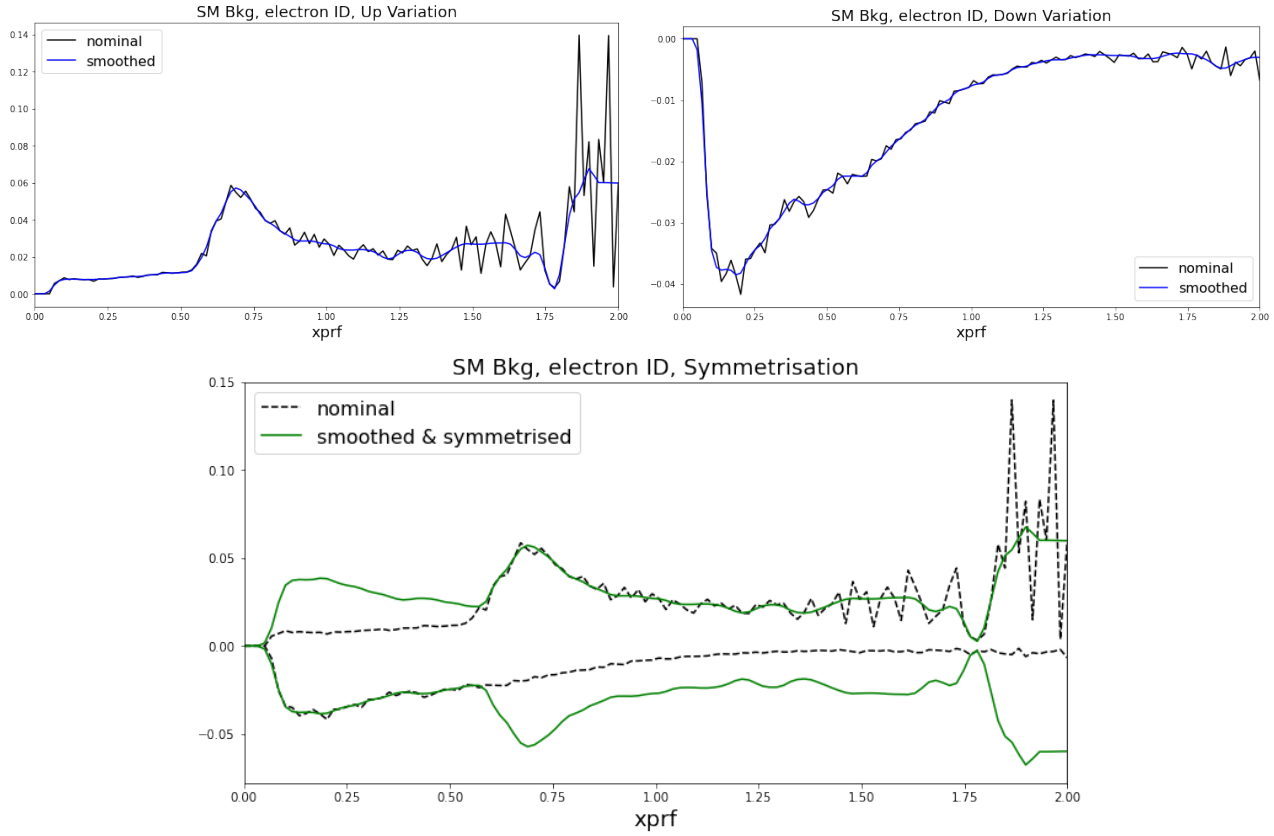


Figure 4.32: Shown are the relative variations of the electron ID for the background template before and after the smoothing procedure (top) and symmetrization (bottom). This particular correction variation has been chosen for the sake of example [11].

4.7.3 Upper limit on Data

The final fit and upper limit results are shown in Table 4.13. Results at 90 and 95 % C.L. are calculated.

Figure 4.13 presents the final results at 95 % C.L. The plot reports the expected upper limit under the background-only hypothesis (Brazilian plot) and the observed upper limit (red dots). The ranges of the upper limits are $(1.1 - 9.7) \times 10^{-3}$ for the electron channel and $(0.7 - 12.2) \times 10^{-3}$ for the muon channel. The systematic uncertainties (see Section 4.5) degrade on average our upper limit sensitivity by approximately 35% in both channels.

Our 95% CL limits are 2.2 to 14 times more stringent than the best previous bounds in [25], depending on the value of the α mass.

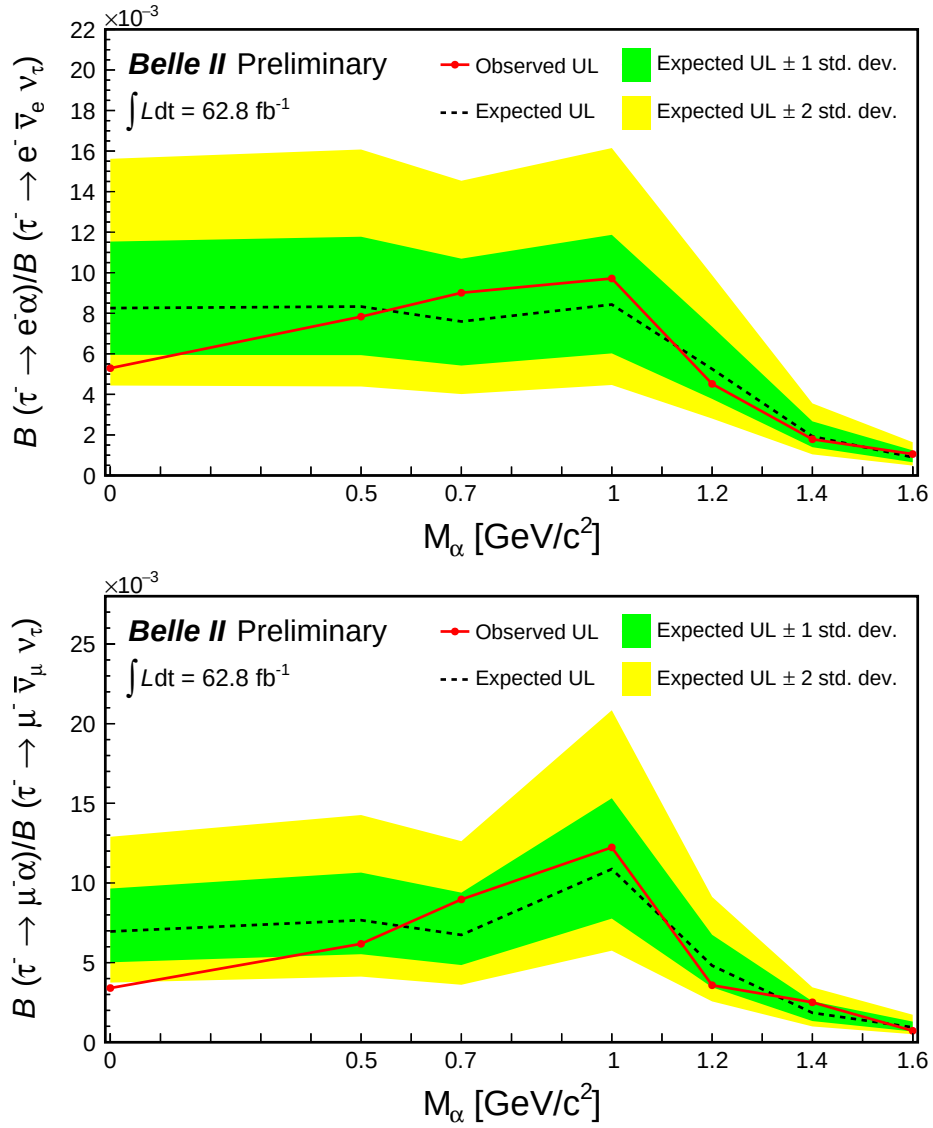


Figure 4.33: Upper limits at 95% C.L. on the branching-fraction ratios $\mathcal{B}_{e\alpha}/\mathcal{B}_{e\nu\bar{\nu}}$ (top) and $\mathcal{B}_{\mu\alpha}/\mathcal{B}_{\mu\nu\bar{\nu}}$(*bottom*) as a function of the α mass, as well as their expectations from the background-only hypothesis. All values are linearly interpolated between mass points [12].

M_α (GeV/c^2)	$\mathcal{B}_{e\alpha}/\mathcal{B}_{e\nu\bar{\nu}}$ ($\times 10^{-3}$)	UL at 95% C.L. ($\times 10^{-3}$)	UL at 90% C.L. ($\times 10^{-3}$)
0.0	-8.1 ± 3.9	5.3 (0.94)	4.3 (0.76)
0.5	-0.9 ± 4.3	7.8 (1.40)	6.5 (1.15)
0.7	1.7 ± 4.0	9.0 (1.61)	7.6 (1.36)
1.0	1.7 ± 4.2	9.7 (1.73)	8.2 (1.47)
1.2	-1.1 ± 2.6	4.5 (0.80)	3.7 (0.66)
1.4	-0.3 ± 1.0	1.8 (0.32)	1.5 (0.26)
1.6	0.2 ± 0.5	1.1 (0.19)	0.9 (0.16)
M_α (GeV/c^2)	$\mathcal{B}_{\mu\alpha}/\mathcal{B}_{\mu\nu\bar{\nu}}$ ($\times 10^{-3}$)	UL at 95% C.L. ($\times 10^{-3}$)	UL at 90% C.L. ($\times 10^{-3}$)
0.0	-9.4 ± 3.7	3.4 (0.59)	2.7 (0.47)
0.5	-3.2 ± 3.9	6.2 (1.07)	5.1 (0.88)
0.7	2.7 ± 3.4	9.0 (1.56)	7.8 (1.35)
1.0	1.7 ± 5.4	12.2 (2.13)	10.3 (1.80)
1.2	-0.2 ± 2.4	3.6 (0.62)	2.9 (0.51)
1.4	0.9 ± 0.9	2.5 (0.44)	2.2 (0.38)
1.6	-0.3 ± 0.5	0.7 (0.13)	0.6 (0.10)

Table 4.13: Final results, central values, and upper limits at 90 and 95 % C.L. In the parentheses are shown the corresponding UL for $\mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \alpha)$ computed using the PDG [13] branching fractions for $\mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \nu\bar{\nu})$ [11].

Chapter 5

Search of the decay $\tau^{\pm} \to \ell^{\pm} + \gamma$ at Belle II

In this chapter, we present the analysis for the search of the decay $\tau^{\pm} \to \ell^{\pm} + \gamma$.

The $\tau^{\pm} \to \ell^{\pm} + \gamma$ decay is a Lepton Flavor Violating (LFV) process and is one of the golden channels for searching new physics beyond the Standard Model of particles physics.

The most stringent limits were presented by Babar and Belle, with integrated luminosities of 516 fb^{-1} and 988 fb^{-1}, respectively. More details are reviewed in Chapter 2.

Here, we perform a sensitivity analysis to measure the upper limit on the branching ratio of $\tau^{\pm} \to \ell^{\pm} + \gamma$ with the Belle II experiment assuming 400 fb^{-1} (almost the sample size collected by the Belle II detector before LS1, as mentioned in Chapter 3).

Our results represent an update on the search for $\tau^{\pm} \to \ell^{\pm} + \gamma$ within Belle II. In this case, the background ratios are much larger than in Belle. However, we are working to improve the current searching and statistical methods with 400 fb^{-1}. This analysis is still in a blinded stage.

5.1 Data and Monte Carlo simulation samples

5.1.1 Monte Carlo samples

To generate signal Monte Carlo (MC) samples of $\tau^{\pm} \to \ell^{\pm} + \gamma$, we use the official (centralized) MC production of the Belle II experiment. We generate 4 million signal events for both charge signs. The samples correspond to the "Early Phase 3 MC14a" production with nominal, run-independent beam background overlays (`'BGx1'`).

To quantify the background composition after the event reconstruction and selection, we generate generic and low multiplicity samples with the same beam background and data taking conditions as the signal samples. The samples are listed in Table 5.1 (in this analysis $e^+e^-K^-K^+$, $e^+e^-\pi^-\pi^+$, and $e^+e^-p^-p^+$ are not considered as they are negligible). It is convenient to split the MC samples into two sets, "train" and "test". The total integrated luminosity of these samples is 800 fb^{-1} (some samples require an appropriate rescaling). The samples are simulated using the production cross sections listed in Table 3.1.

5.1.2 Data

The data samples considered here were collected by the Belle II experiment from e^+e^- collisions at a center-of-mass energy of 10.58 GeV during 2019. Three different data sets were collected this year under different circumstances, as the SuperKEKB instantaneous luminosity gradually

Process	Name of production	Scale factor	"train" sample $\int \mathcal{L}$ dt (fb^{-1})	"test" sample $\int \mathcal{L}$ dt (fb^{-1})
Generic				
$e^+e^- \to B\bar{B}$	mixed /charged	1	400	400
$e^+e^- \to \tau^+\tau^-$	taupair	1	400	400
$q\bar{q}$				
$e^+e^- \to u\bar{u}$	uubar	1	400	400
$e^+e^- \to d\bar{d}$	ddbar	1	400	400
$e^+e^- \to s\bar{s}$	ssbar	1	400	400
$e^+e^- \to c\bar{c}$	ccbar	1	400	400
Low multiplicity				
$e^+e^- \to e^+e^-(\gamma)$	eeg	8	400	400
$e^+e^- \to \mu^+\mu^-(\gamma)$	mmg	1	400	400
4 leptons process				
$e^+e^- \to e^+e^-e^+e^-$	eeee	4	400	400
$e^+e^- \to e^+e^-\mu^+\mu^-$	eemm	4	400	400

Table 5.1: Process description and integrated luminosity of the Early Phase III MC14 samples used in the analysis.

increased. More details are given in Table 5.2. The total integrated luminosity is 8.76 ± 0.00145 fb^{-1}.

The sample is small but enough to perform a preliminary Data-MC comparison in sidebands as shown in Section 5.7.

Data sets	Processing Label	Experiment	$\int$ L dt (pb^{-1})
Data set 1	Proc11	7	425.5 ± 0.3
Data set 2	Proc11	8	4597.4 ± 0.9
Data set 3	Proc11	10	3741.3 ± 1.1

Table 5.2: Belle II experiment data sets used in this work. The processing and experiment number labels are shown, as well as, the corresponding integrated luminosity.

The MC and data samples were processed without any pre-selection of events (a process named "skim").

5.2 Event Selection

5.2.1 Event reconstruction

We reconstruct $e^+e^- \to \tau^+\tau^-$ events where one tau lepton decays to $\tau^{\pm} \to \ell^{\pm} + \gamma$ and the other to a charged particle (denoted as $\pi^{\pm 1}$) and missing energy, $\tau^{\pm} \to \pi^{\pm} + missing\ energy$. When reconstructing these "1x1 prong topology" events, we require exactly two tracks with zero net charge, originating close to the interaction region:

- any displacement from the interaction point should be less than 3 cm in the z-axis,

[1]For the reconstruction of any charged particle, in this chapter, we use the notation of "$\pi^{\pm}$", where $\pi^{\pm}$ can be an electron e, a muon μ, a charged pion $\pi^{\pm}$, a kaon $k^{\pm}$, etc. For leptons, only "$\ell^{\pm}$" is used.

- and less than 1 cm in the transverse plane.

The photon candidates are reconstructed by satisfying the following requirements:

- ECL clusters with energy deposits of at least 0.2 GeV ($E(\gamma) > 0.2$ GeV) ;
- should be within the CDC acceptance;
- the sum of the number of crystals in an ECL cluster should be greater than 1.5 ;
- the time of the ECL cluster (calculated as the photon timing minus the event t_0) should be less than 200 ns .

The neutral pion candidates are constructed by combining two photons satisfying the conditions above but with a lower energy threshold, with an additional requirement on the invariant mass of the two photon system:

- $E(\gamma) > 0.1$ GeV ;
- $115 < M_{\gamma\gamma} < 152$ MeV/c^2 .

Additionally, photons used for the neutral pion reconstruction are not considered candidates for the $\tau^{\pm} \to \ell^{\pm}\gamma$ signal decay.

5.2.2 1x1 prong topology

We use the thrust vector (see Equation 4.1) -computed from the two tracks, the photons candidates, and the neutral pions- to separate the events into signal and tag hemispheres.

As shown in Figure 5.1, the hemisphere corresponding to the signal side, $\tau^{\pm} \to \ell^{\pm}+\gamma$, should contain a single track ("1-prong") plus a photon candidate; while the hemisphere corresponding to the tag side, $\tau^{\pm} \to \pi^{\pm} + missing\ energy$, should contain only a single track ("1-prong"). The 1x1 prong topology is imposed by requiring

$$(\overrightarrow{p}^{CMS}_{signal} \cdot \widehat{n}_{trust}) \cdot (\overrightarrow{p}^{CMS}_{tag} \cdot \widehat{n}_{trust}) < 0 \ . \tag{5.1}$$

5.2.3 Particle identification

Signal side

For the signal identification, we require the charged track to be either an electron or a muon, by imposing the following particle identification (PID) conditions:

- Global electron "ID_noSVD_noTOP" likelihood[2]> 0.95,

  or

- Global muon ID likelihood > 0.95 .

Similar to the $\tau^{\pm} \to \ell^{\pm}+\alpha$ analysis (in Chapter 4), we apply correction factors to MC associated to both the lepton efficiency and the probability that a charged pion ($\pi^{\pm}$) is misidentified as a charged lepton ($\ell^{\pm}$) (or $\pi^{\pm} \to \ell^{\pm}$ fake rate). These correction factors are provided by the Lepton ID Group within the Belle II Collaboration and their statistical and systematics uncertainties will be considered as a systematic uncertainty source in Section 5.6.

[2]The information of the SVD and TOP detectors is excluded.

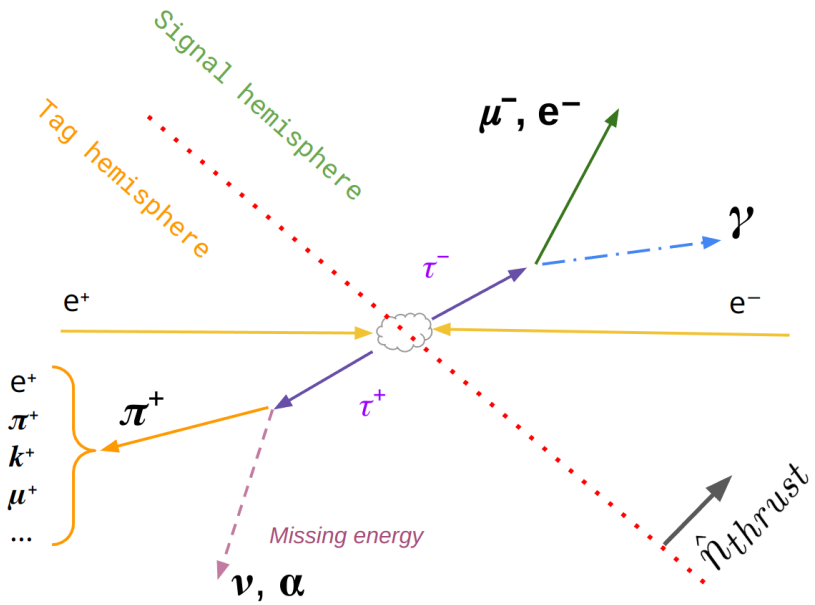


Figure 5.1: Ilustration of the 1×1 topology. The signal $\tau^{\pm} \to \ell^{\pm} + \gamma$ and the tag $\tau^{\pm} \to \pi^{\pm} + missing\ energy$ decays are in opposite hemispheres.

Tag side

On the tag side, we exclude the possibility to have the same lepton candidate on this side by applying a veto condition to the charged track: if on the signal side, we find a muon candidate, on the tag side, we allow any possible charged particle except for a muon (analogously for the electron signal candidate). Then, the tag track should satisfy:

- Global muon ID likelihood<0.5 if the signal side has a muon candidate,

  and

- Global electron ID_noSVD_noTOP likelihood $<$ 0.5 if the signal side has an electron candidate.

These requirements basically reject $e^+e^-(\gamma)$ and $\mu^+\mu^-(\gamma)$ processes.

5.2.4 Trigger

Events in data are required to fire the logical OR of several un-prescaled low-multiplicity ECL (lml) triggers:

- **lml0** or **lml1** or **lml2** or **lml4** or **lml6** or **lml7** or **lml8** or **lml9** or **lml10** .

In order to account for the trigger effect when computing the upper limit in Section 5.5, we require the same trigger bits for the MC samples, making use of the Trigger Simulation (TSIM).

A description of these trigger bits is provided in Appendix B. Hereafter, this combination of triggers will be referred to as *lmlX*.

5.2.5 LFV skim

As was mentioned previously, the MC and data samples were processed without preselection requirements, that is, unskimmed samples are used for this analysis. In order to mimic the conditions of the LFV skim, which was created specifically for lepton flavor violation searches, we apply the following conditions at the reconstruction level:

- $1.0 < M_{\tau^-} < 2.0$ GeV/c^2 ;
- $-1.5 < \Delta E < 0.5$ GeV .

These conditions are applied on the LFV skim by default, with the purpose to keep most of the $\tau^{\pm} \to \ell^{\pm} + \gamma$ signal and to reject SM processes. By using these conditions from now on, we prevent changes in the future, since sooner than later, skim algorithms will be applied by default to all accessible MC and data samples due to disk space constraints.

5.2.6 Additional selections

After the reconstruction of the events with the requirements above, in an offline stage, we apply the following requirements:

- $\mathrm{P}_t > 0.1$ GeV/c, for the two selected tracks;
- $-0.8660 < \cos\theta_{miss} < 0.9565$ (missing energy in the CDC acceptance) .

The P_t transverse momentum condition ensures that the tracks have a transverse component, perpendicular to the beam axis, within the detector fiducial region. The second requirement is imposed to reduce the probability that the missing particle(s) fall outside the acceptance of the detector; then, we increase the probability of the missing particle to be a neutrino, rather than a photon o charged track falling outside the detector acceptance.

Signal photon selection

In this analysis, both the signal efficiency and the achieved purity significantly affect the sensitivity of the search. To increase the signal efficiency, we decided not to veto the neutral particles.

To properly perform the calculation of the upper limit on the branching ratio of the $\tau^{\pm} \to \ell^{\pm} + \gamma$ decay, until Section 5.5, we select events with only one γ in the signal side:

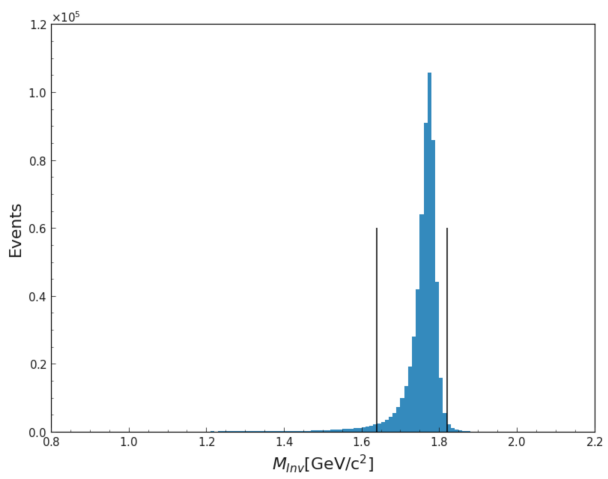
- $\#\gamma_{\mathrm{signal-side}} = 1$.

Invariant mass pre-selection

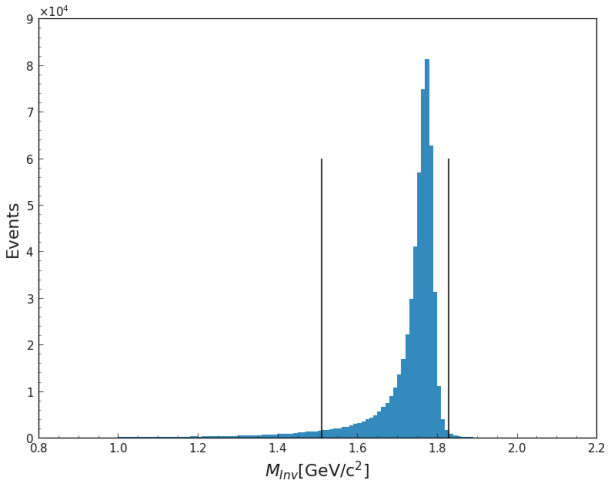
A pre-selection cut on the invariant mass M_{inv},

$$M_{inv} \equiv M_{\ell+\gamma} = \sqrt{E^2_{\ell+\gamma} - P^2_{\ell+\gamma}}\,, \tag{5.2}$$

of the $\ell^{\pm}\gamma$ candidate is applied. Remaining MC background events are used to train background suppression methods. Moreover, a narrower M_{inv} signal region helps to reduce the size of the samples that we are analyzing.



(a) M_{inv} distribution for muon channel.



(b) M_{inv} distribution for muon channel.

Figure 5.2: Invariant mass pre-selection. The requirements are shown with solid black lines. This analysis is performed with signal and background MC events inside this region.

5.3 Background Suppression

The selection is optimized by maximizing the following Punzi figure of merit (FOM):

$$FOM = \frac{S}{\frac{a}{2} + \sqrt{B}} , \tag{5.3}$$

where S is the number of simulated signal events with $\tau^{\pm} \to \ell^{\pm}\gamma$ decays and B is the total number of background events, both after reconstruction and pre-selection requirements. The constant a has been chosen to the value $a = 1.28$ which corresponds to a 90 % C.L. search of NP.

5.3.1 Discriminating variables used for the optimization

The variables used for the optimization of the background suppression are:

1. P_{ℓ}^{CMS}: Momentum of the signal track in the center of mass system (CMS) frame.
2. P_{π}^{CMS}: Momentum of the tag track in the CMS frame.
3. E_{γ}: Photon energy.
4. $\cos\theta_{\ell-\gamma}^{CMS}$: Cosine of the opening angle between the signal track and photon candidates in the CMS frame.
5. $\cos\theta_{miss-\pi}^{CMS}$: Cosine of the angle between the missing momentum (see below) and the track on the tag side in the CMS frame.
6. $\cos\theta_{\ell-\tau}^{R}$: Cosine of the opening angle between the signal lepton ℓ and its mother τ in the τ rest frame.
7. $\cos\theta_{\ell-\pi}^{CMS}$: Cosine of the opening angle between the signal and tag tracks in the CMS frame.
8. $Thrust$: The magnitude of the thrust vector given by Equation 4.1.

9. P_{miss}: Momentum corresponding to the missing energy. The missing momentum P_{miss} is calculated by subtracting the vector sum of all visible momenta (tracks, photon, and neutral pions) in the event from the sum of the e^+ and e^- beam momenta:
$$\overrightarrow{P}_{miss} = \overrightarrow{P}_{e^+} + \overrightarrow{P}_{e^-} - \overrightarrow{P}^{sum}_{visible} \tag{5.4}$$

10. E_{vis}: Visible energy of the event.

11. M^2_ν: Missing mass squared on the tag-side, defined as:
$$M^2_\nu = \left(E^{CMS}_{\mu\gamma} - E^{CMS}_{tag}\right)^2 - \left(P^{CMS}_{miss}\right)^2, \tag{5.5}$$
where $E^{CMS}_{\mu\gamma}$ (E^{CMS}_{tag}) is the sum of the energy on the signal (tag) side in the CMS frame.

In Appendix C we show the distributions of these variables for the muon and electron channels. These distributions are presented after all the requirements in Section 5.2.

5.3.2 Cut-based selection

A cut-based selection is performed by maximizing the FOM described in Equation 5.3 for each variable described in Section 5.3.1. The initial number of events, i.e., after applying all the previous requirements (see Section 5.2), can be seen in Table 5.3 and 5.4. The optimization of the requirements (cuts) is performed with the "train" MC sample. Then, to test the performance of the optimization, the same cuts are applied to an independent "test" MC sample.

MC14 samples	μ channel	e channel
$\tau^\pm \to \ell^\pm\gamma$	$\epsilon = 27.57\%$	$\epsilon = 26.38\%$
$\tau^+\tau^-$	183,146 ± 428	192,043 ± 438
$\mu^+\mu^-(\gamma)$	92,585 ± 304	73 ± 9
$q\bar{q}$	8,808 ± 94	1,578 ± 40
$e^+e^-(\gamma)$	7,936 ± 252	7,725,712 ± 7,861
$\ell^+\ell^-\ell^+\ell^-$	196 ± 28	276 ± 33
$B\bar{B}$	8 ± 3	13 ± 4
Total BKG	292,679 ± 591	7,919,695 ± 7,873

Table 5.3: Initial number of events of the "train" MC samples, for the muon (left) and electron (right) channels. The signal efficiency is also shown. Background samples correspond to 400 fb^{-1}.

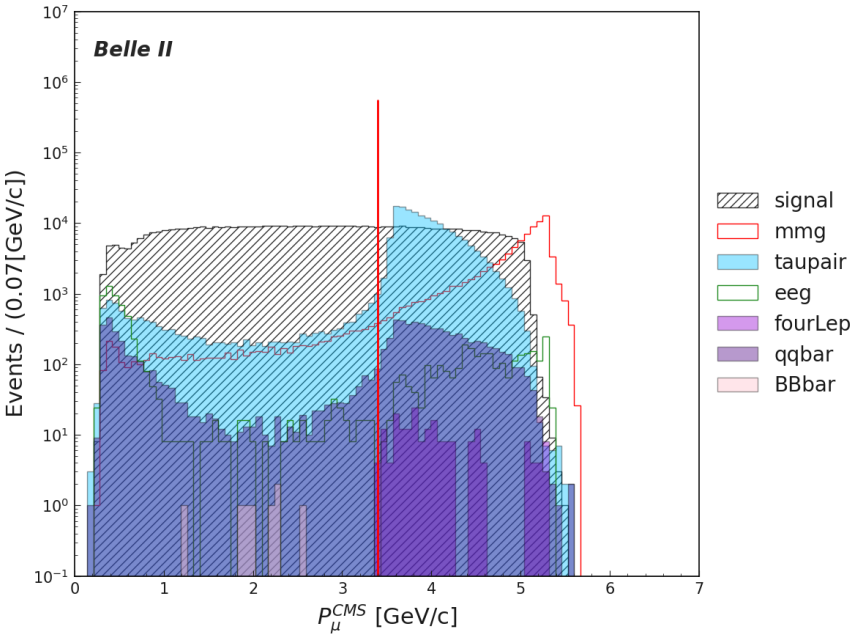
MC14 samples	μ channel	e channel
$\tau^\pm \to \ell^\pm\gamma$	$\epsilon = 27.60\%$	$\epsilon = 26.30\%$
$\tau^+\tau^-$	182,496 ± 427	191,882 ± 438
$\mu^+\mu^-(\gamma)$	92,362 ± 304	79 ± 9
$q\bar{q}$	8,644 ± 93	1,620 ± 40
$e^+e^-(\gamma)$	7,888 ±251	7,728,504 ± 7,863
$\ell^+\ell^-\ell^+\ell^-$	196 ± 28	272 ± 33
$B\bar{B}$	9 ± 3	13 ± 4
Total BKG	291,595 ± 589	7,922,370 ± 7,875

Table 5.4: Initial number of events of the "test" MC samples, for the muon (left) and electron (right) channels. The signal efficiency is also shown. Background samples correspond to 400 fb^{-1}.

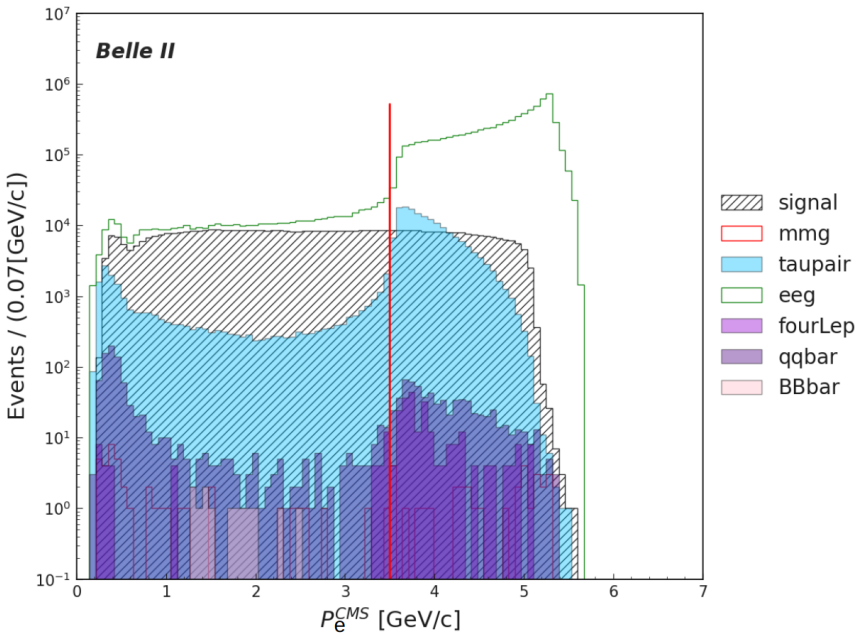
Optimization

1. P_{ℓ}^{CMS}

We apply an upper cut on the signal track momentum to reduce low multiplicity backgrounds, as can be seen in Figure 5.3. Tracks generated from $e^+e^-(\gamma)$ and $\mu^+\mu^-(\gamma)$ backgrounds are expected to have high momenta. $P_{\ell}^{CMS} < 3.4$ (< 3.5) GeV/c is applied for the muon (electron) channel.



(a)

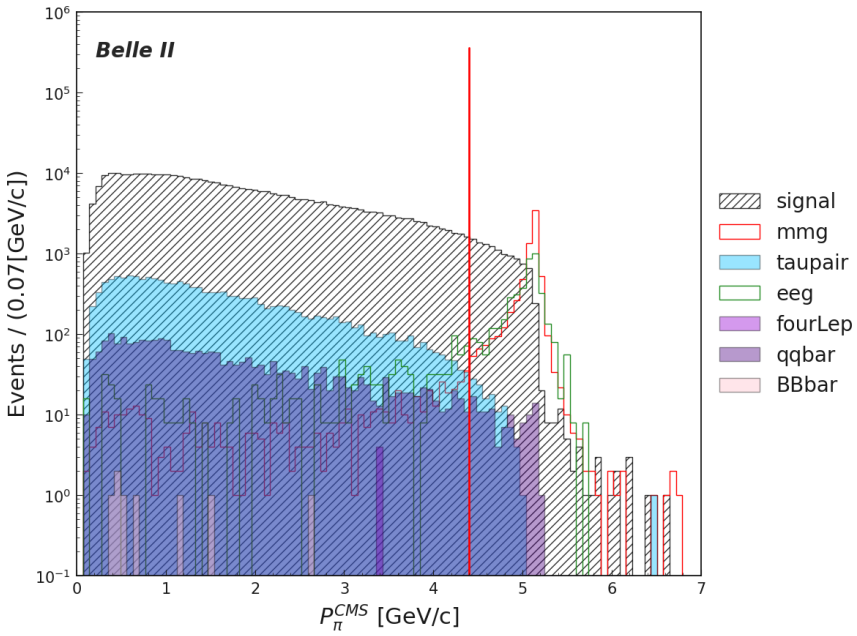


(b)

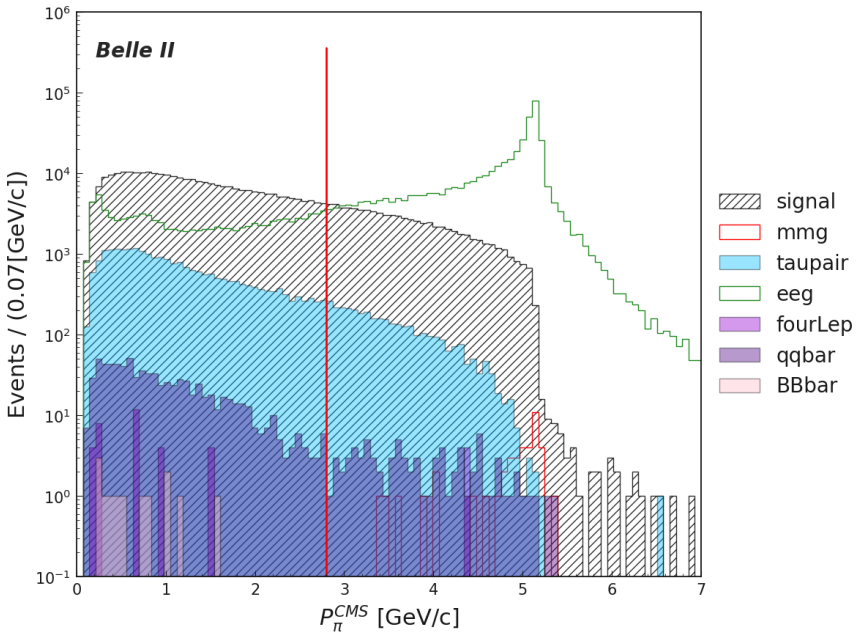
Figure 5.3: Distribution of the P_{ℓ}^{CMS} variable for (a) the muon and (b) the electron channels. The requirement on this variable is indicated with a solid red line, events to the right are rejected.

2. P_π^{CMS}

Under the same previous assumption, we apply an upper cut on the momentum of the tag track $P_\pi^{CMS} < 4.4$ (< 2.8) GeV/c, for the muon (electron) channel. Then, we remove the remaining low multiplicity, as can be seen in Figure 5.4 on the left side. As expected, $e^+e^-(\gamma)$ is the most important contamination for the electron channel.



(a)

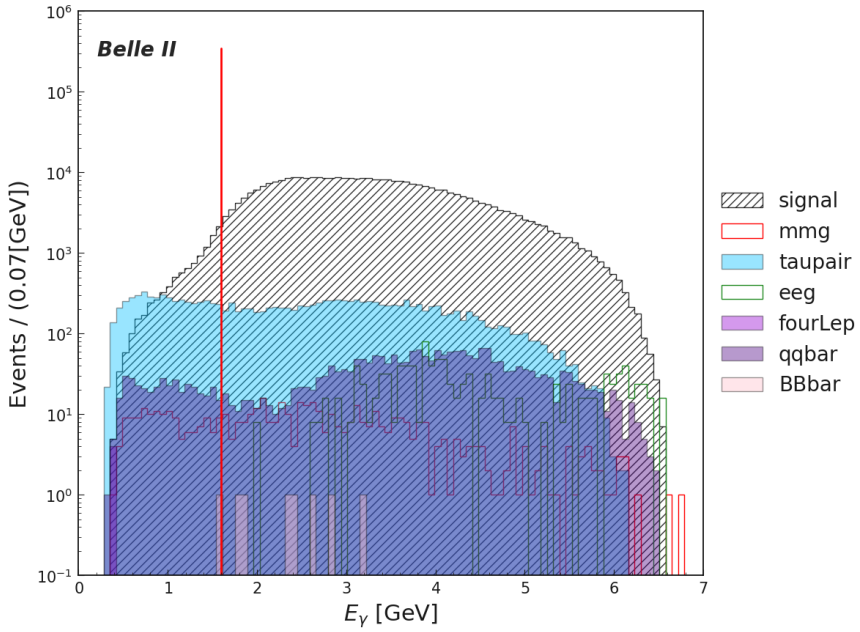


(b)

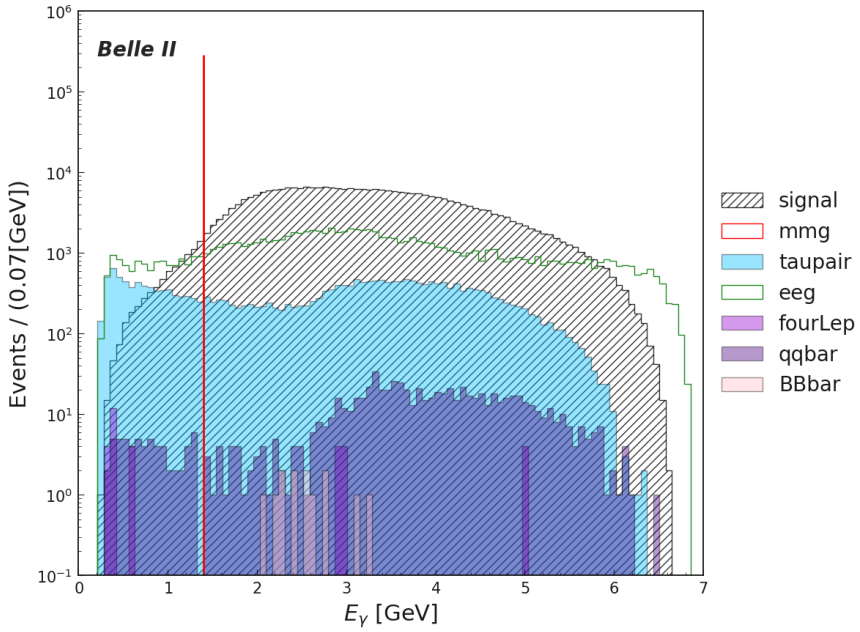
Figure 5.4: Distribution of the P_π^{CMS} variable with the requirement of the previous variable for (a) the muon and (b) the electron channels. The requirement on this variable is indicated with a solid red line, events to the right are rejected.

3. E_γ

The photon candidate in the $\tau^{\pm} \to \ell^{\pm} + \gamma$ decay is expected to be highly energetic, and a lower cut is effective to reject $\tau^+\tau^-$ background processes. The selection 1.6 GeV $< E_\gamma$ (1.4 GeV $<$) is applied for the muon (electron) channel, as can be seen in Figure 5.5.



(a)



(b)

Figure 5.5: Distribution of the E_γ variable, with the requirements of the previous variables for (a) the muon and (b) the electron channels. The requirement on this variable is indicated with a solid red line, events to the left are rejected.

4. $\cos\theta^{CMS}_{\ell-\gamma}$

For signal, the lepton and the photon should be approximately pointing to the same direction in the CMS, thus, we required $0.75 < \cos\theta^{CMS}_{\ell-\gamma} < 0.78$ and $0.24 < \cos\theta^{CMS}_{\ell-\gamma} < 0.78$ for the muon and electron channels, respectively. This cut is useful to reduce $\tau^+\tau^-$ events, as can be seen in Figure 5.6

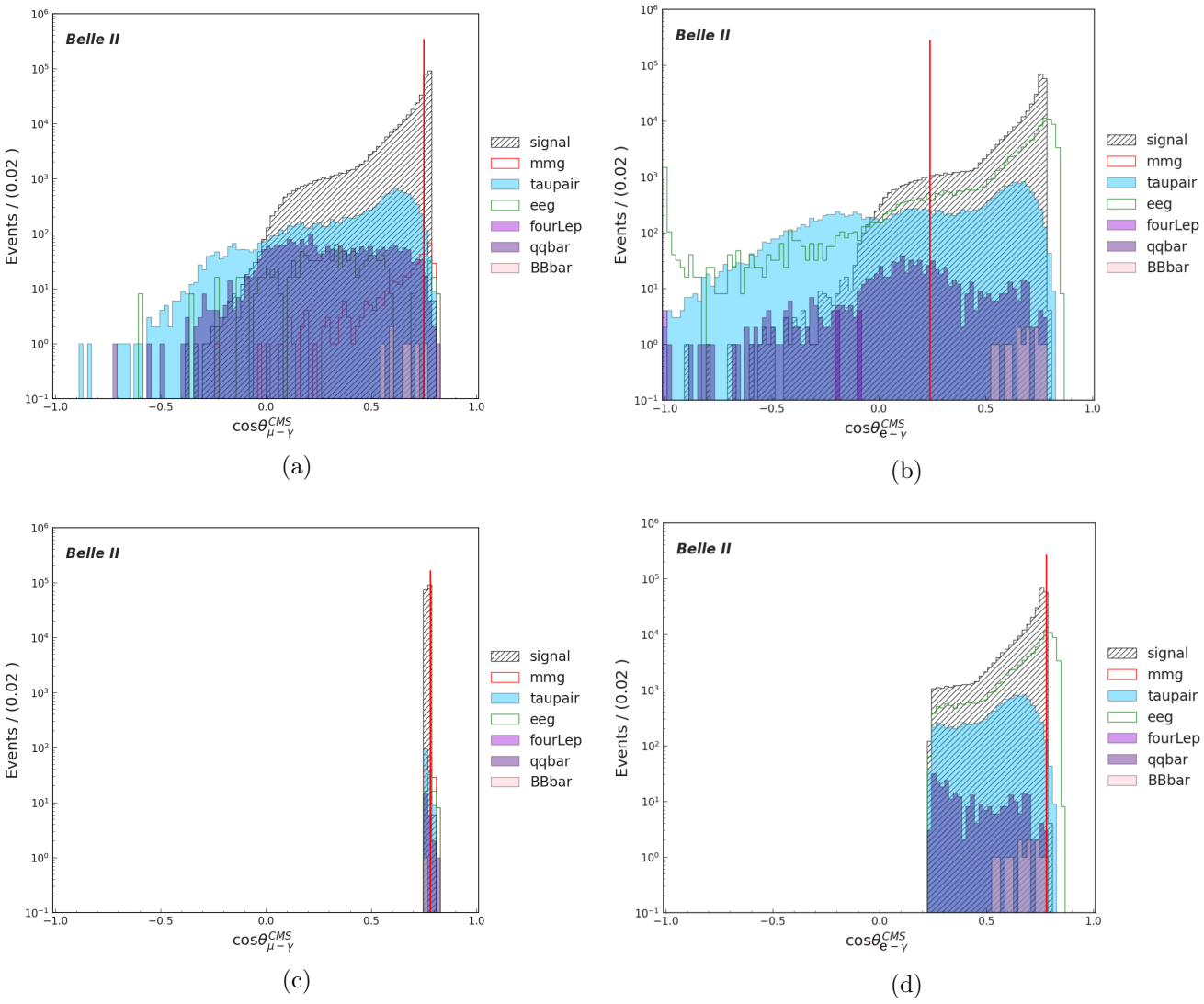


Figure 5.6: Distributions of the $\cos\theta^{CMS}_{\ell-\gamma}$ variable with the requirements of the previous variables. Lower (top) and upper (bottom) cuts are shown for (a), (c) the muon and (b), (d) electron channels. The requirements on these variables are indicated with solid red lines, we keep events between these lines.

5. $\cos\theta^{CMS}_{miss-\pi}$

The cosine of the opening angle between the missing energy and the tag track is required to be $0.11 < \cos\theta^{CMS}_{\ell-\gamma} < 0.98$ and $-0.04 < \cos\theta^{CMS}_{\ell-\gamma} < 0.96$, for the muon and electron channels, respectively. For signal events, we expect that the missing energy and the tag track form a small angle, since no missing energy is expected on the signal side. The selection is shown in Figure 5.7.

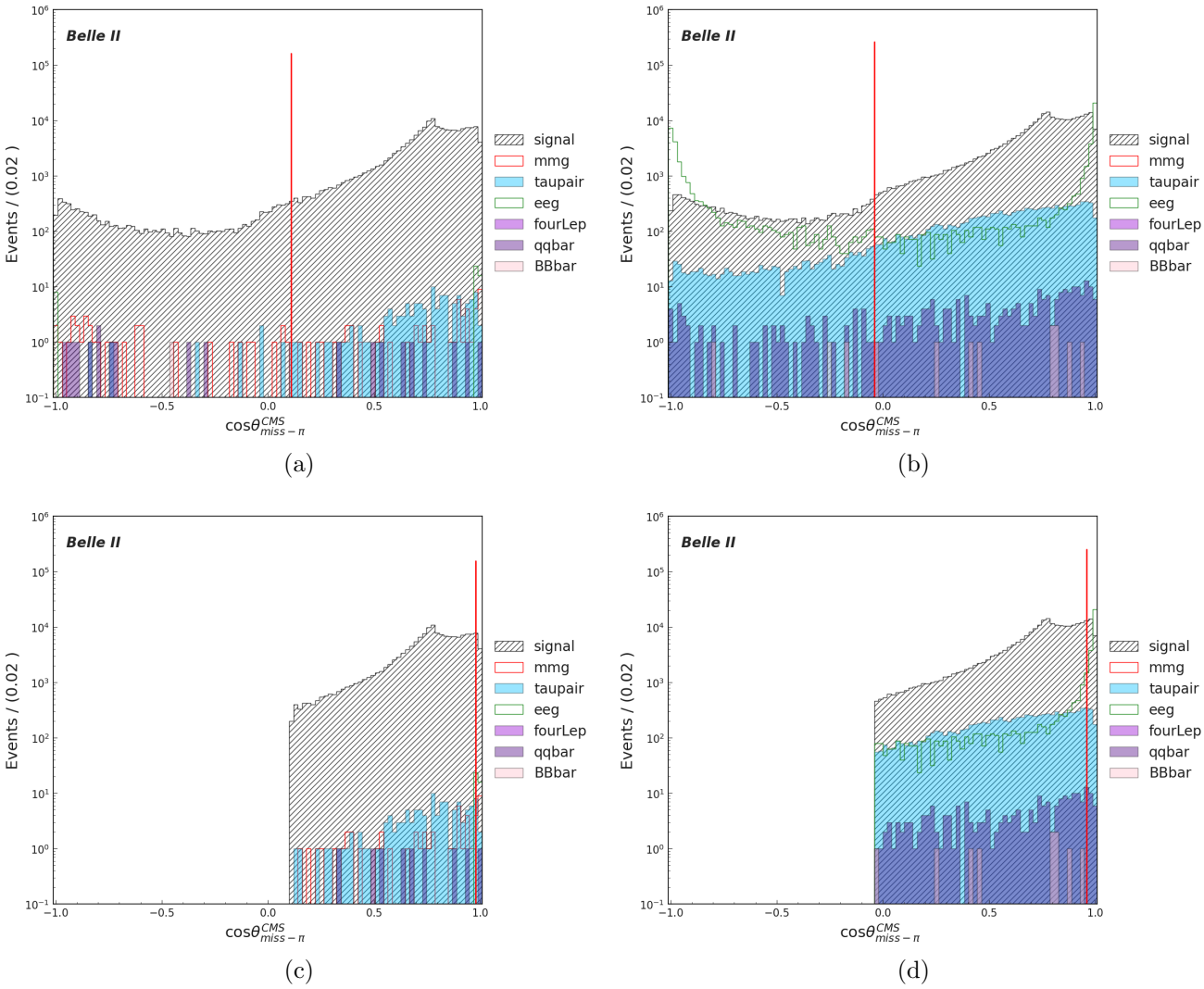
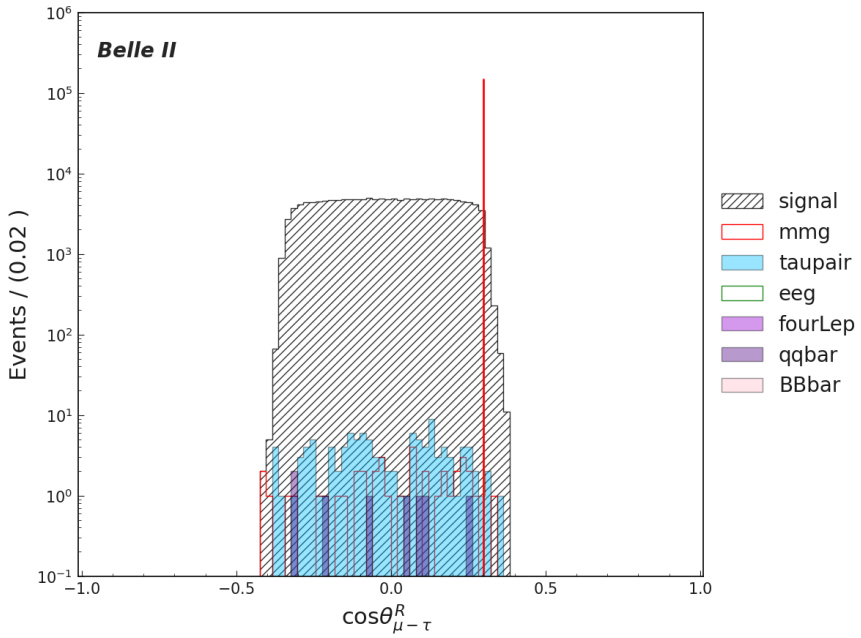


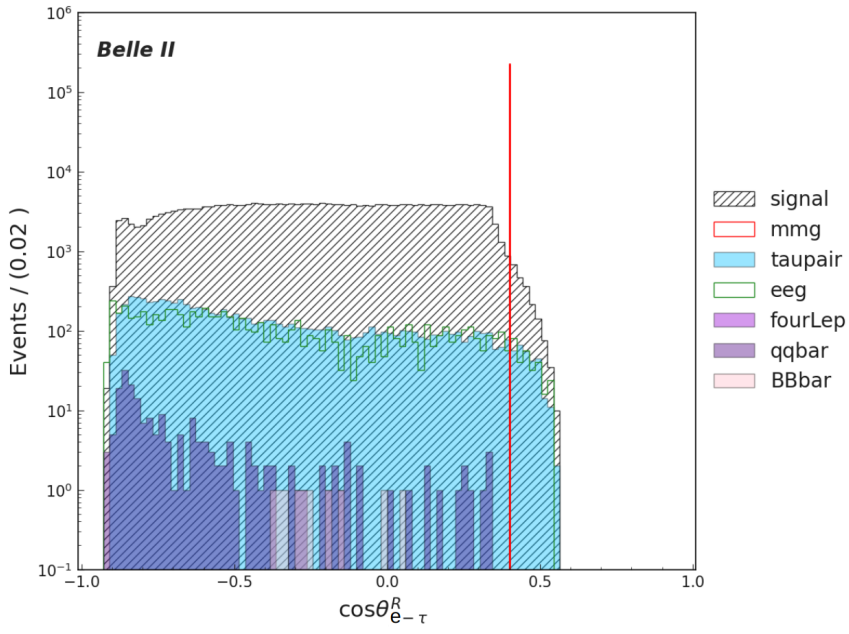
Figure 5.7: Distributions of the $\cos\theta^{CMS}_{miss-\pi}$ variable with the requirements of the previous variables. Lower (top) and upper (bottom) cuts are shown for (a), (c) the muon and (b), (d) electron channels. The requirements on these variables are indicated with solid red lines, we keep events between these lines.

6. $\cos\theta^R_{\ell-\tau}$

The cosine of the opening angle between the signal lepton and the boost direction of the mother τ in the rest frame is required to be $\cos\theta^R_{\ell-\tau} < 0.3$ (< 0.4) for the muon (electron) channel. For signal candidates, we expect to have a uniform distribution in this variable, as can be seen in Figure 5.8.



(a)



(b)

Figure 5.8: Distribution of the $\cos\theta^R_{\ell-\tau}$ variable, with the requirements of the previous variables for (a) the muon and (b) the electron channels. The requirement on this variable is indicated with a solid red line, events to the right are rejected.

7. $\cos\theta^{CMS}_{\ell-\pi}$

The opening angle between the two tracks is expected to be greater than 90°. Then, an upper cut selection $\cos\theta^{CMS}_{\ell-\pi} < 0.02$ (< 0.25) is applied for the muon (electron) channel, as can be seen in Figure 5.9.

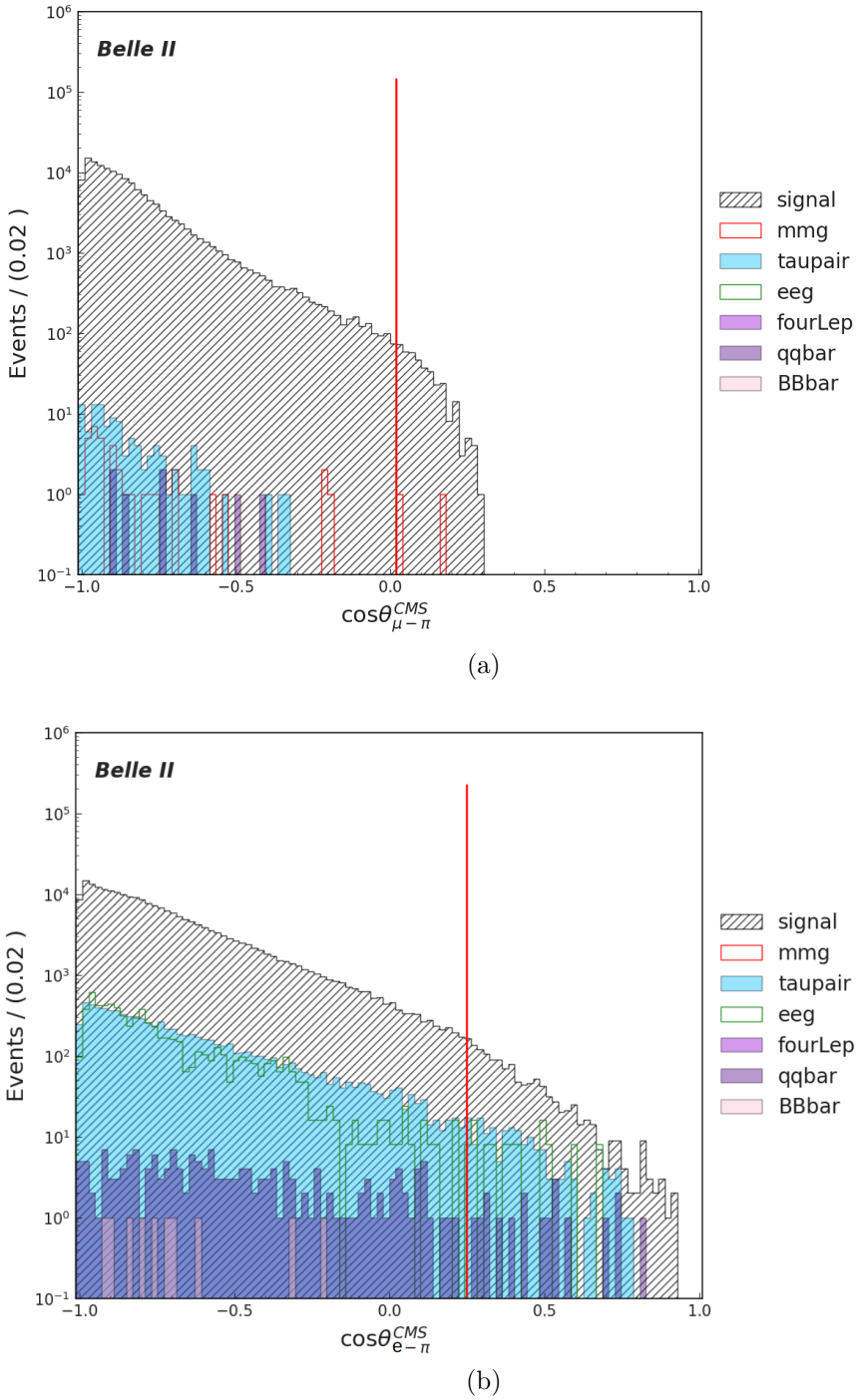


Figure 5.9: Distribution of the $\cos\theta^{CMS}_{\ell-\pi}$ variable, with the requirements of the previous variables for (a) the muon and (b) the electron channels. The requirement on this variable is indicated with a solid red line, events to the right are rejected.

8. *Thrust*

The magnitude of the thrust vector for the signal events should be close to 1. We require $0.9 < thrust < 0.97$ and $0.91 < thrust < 0.96$, for the muon and electron channels, respectively. These conditions help to reduce low multiplicity and remaining $q\bar{q}$ events.

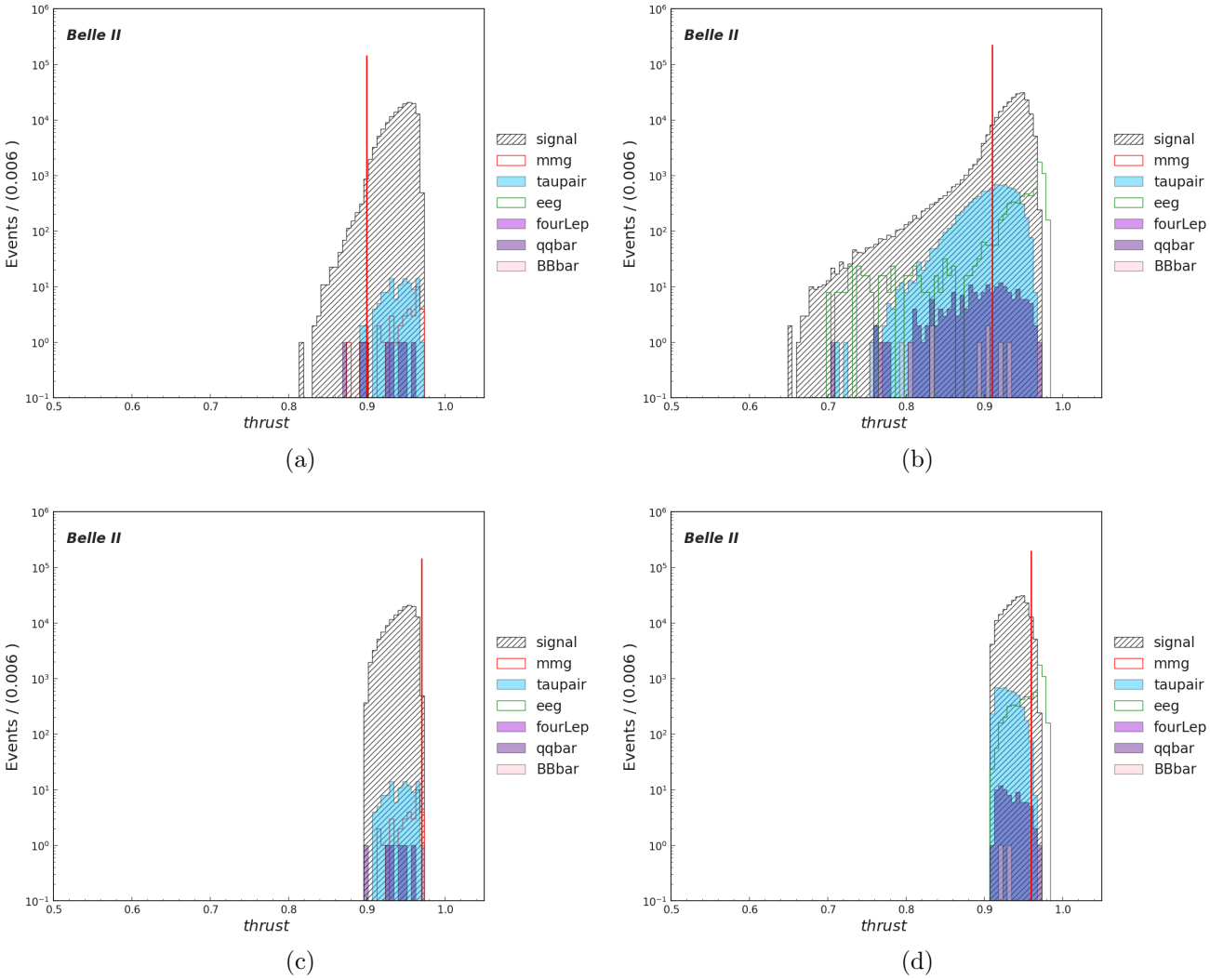
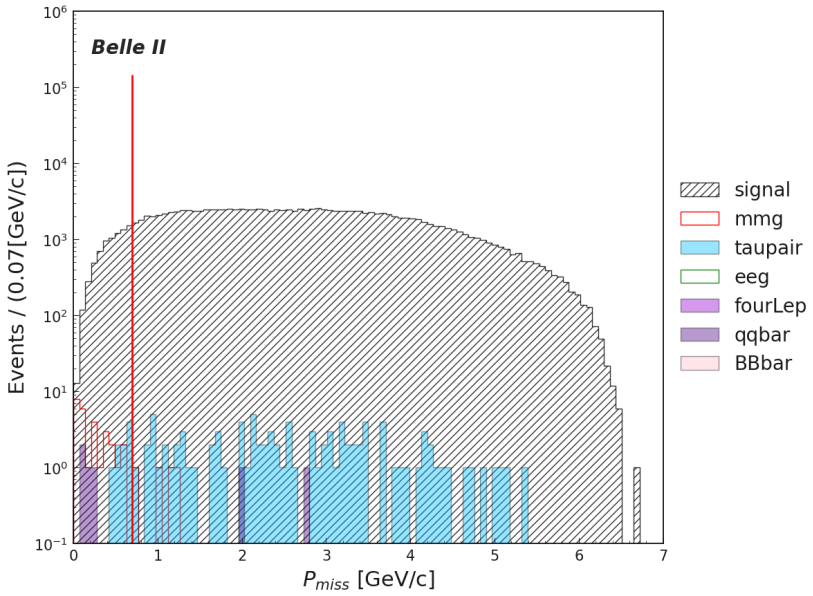


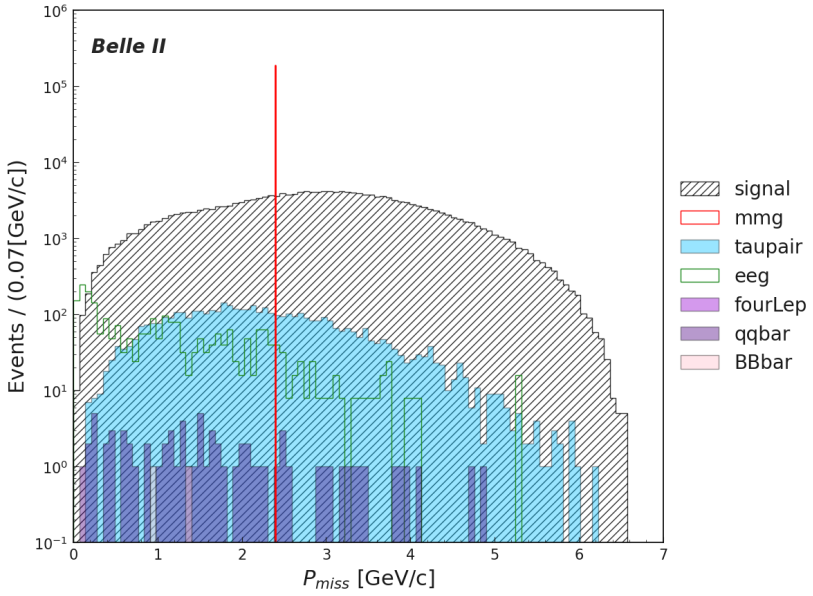
Figure 5.10: Distributions of the *Thrust* variable with the requirements of the previous variables. Lower (top) and upper (bottom) cuts are shown for (a), (c) the muon and (b), (d) electron channels. The requirements on these variables are indicated with solid red lines, we keep events between these lines.

9. P_{miss}

A lower cut in the missing momentum is required, since we expect to have missing particles (neutrinos) on the tag side. We required 0.7 GeV/c $< P_{miss}$ (2.4 GeV/c $<$), for the muon (electron) channel. This requirement helps to reduce the remaining low multiplicity events, as can be seen in Figure 5.11.



(a)



(b)

Figure 5.11: Distribution of the P_{miss} variable, with the requirements of the previous variables for (a) the muon and (b) the electron channels. The requirement on this variable is indicated with a solid red line, events to the left are rejected.

10. E_{vis}

For signal events, we expect to have missing energy since only the signal side is fully reconstructed. So this variable helps to reduce the remaining background. As can be seen in Figure 5.12, we require $5.6 < E_{vis} < 11.7$ GeV and $6.2 < E_{vis} < 8.9$ GeV, for the muon and electron channels, respectively.

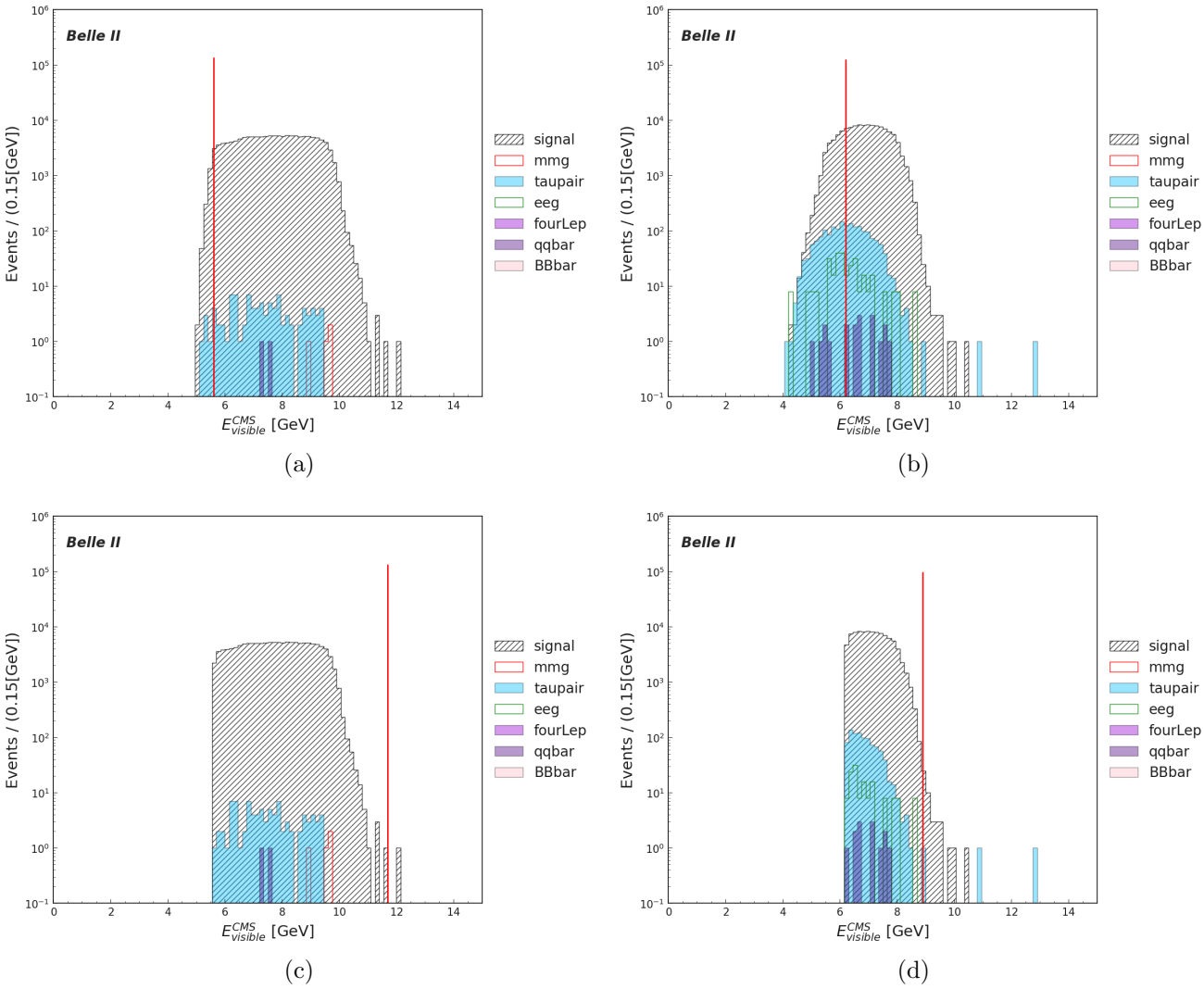
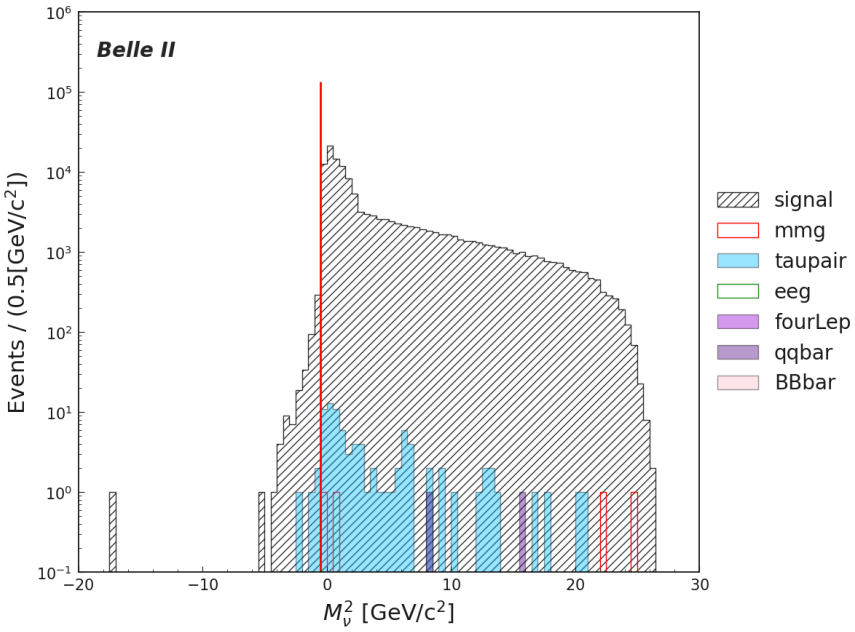


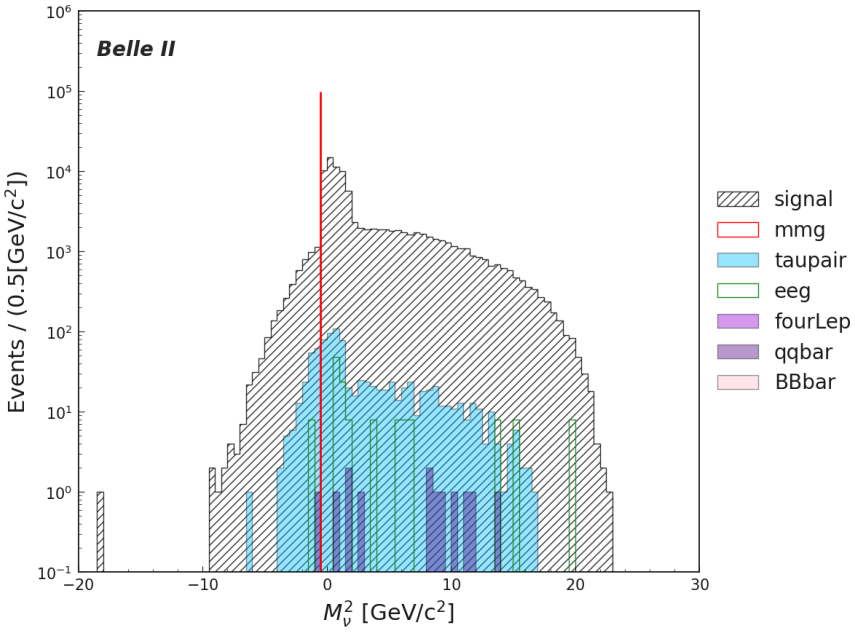
Figure 5.12: Distributions of the E^{CMS}_{vis} variable with the requirements of the previous variables. Lower (top) and upper (bottom) cuts are shown for (a), (c) the muon and (b), (d) electron channels. The requirements on these variables are indicated with solid red lines, we keep events between these lines.

11. M_{ν}^2

The missing mass squared in the tag side is expected to have different distributions for signal and background events, this will depend on the kinematic of the event, including the number of neutrinos that are produced. We require -0.5 GeV/c^2 $< M_{\nu}^2$ for both muon and electron channels, as can be seen in Figure 5.13



(a)



(b)

Figure 5.13: Distribution of the M_{ν}^2 variable, with the requirements of the previous variables for (a) the muon and (b) the electron channels. The requirement on this variable is indicated with a solid red line, events to the left are rejected.

Remaining events after cut-based optimization.

The remaining events after the cut-based optimization, for the "train" and "test" samples are shown in Table 5.5, and for the electron channel in Table 5.6.

MC14 samples	μ channel train	μ channel test
$\tau^{\pm} \to \mu^{\pm}\gamma$	$\epsilon = 6.65\%$	$\epsilon = 6.65\%$
$\tau^{+}\tau^{-}$	85 ± 9	99 ± 10
$\mu^{+}\mu^{-}(\gamma)$	4 ± 2	5 ± 2
$q\bar{q}$	2 ± 1	6 ± 2
$e^{+}e^{-}(\gamma)$	0	0
$\ell^{+}\ell^{-}\ell^{+}\ell^{-}$	0	0
$B\bar{B}$	0	0
Total BKG	91 ± 9	110 ± 10

Table 5.5: Final number of events for the muon channel. The signal efficiency is also shown. Background samples correspond to 400 fb^{-1}.

MC14 samples	e channel train	e channel test
$\tau^{\pm} \to e^{\pm}\gamma$	$\epsilon = 4.59\%$	$\epsilon = 4.62\%$
$\tau^{+}\tau^{-}$	771 ± 28	779 ± 28
$\mu^{+}\mu^{-}(\gamma)$	0	0
$q\bar{q}$	12 ± 3	10 ± 3
$e^{+}e^{-}(\gamma)$	136 ± 33	136 ± 33
$\ell^{+}\ell^{-}\ell^{+}\ell^{-}$	0	0
$B\bar{B}$	0	0
Total BKG	919 ± 43	925 ± 43

Table 5.6: Final number of events for the electron channel. The signal efficiency is also shown. Background samples correspond to 400 fb^{-1}.

5.3.3 Dual Annealing optimization

We tested another method of background reduction by implementing a Dual Annealing optimization. The implementation was made using the free and open-source Python library "SciPy". A Dual Annealing optimization is a stochastic approach that combines Classical Simulated Annealing (CSA) and Fast Simulated Annealing (FSA) coupled to a strategy for applying a local search on accepted locations [53–56].

In this section, we will present the results of the Dual Annealing optimization, for the specific case when the function to be minimized is the negative of the FOM, presented in Equation 5.3. This optimization is a 16-Dimension (16-D) problem, taking into account the upper and lower requirements that we want to find, for all the discriminative variables that are presented in Section 5.3.1.

Dual Annealing optimization for the muon channel

This alternative optimization was only applied for the muon channel $\tau^{\pm} \to \mu^{\pm}\gamma$, since the amount of background is much smaller compared with the electron channel. As we will see, the Dual Annealing optimization reaches a better signal efficiency.

The combination of requirements (cuts), that the algorithm find for the maximum of the FOM in Equation 5.3, are:

1. $P_{\ell}^{CMS} < 4.04$ GeV/c
2. $P_{\pi}^{CMS} < 5.5$ GeV/c
3. 1.58 GeV $< E_{\gamma}$
4. $0.73 < \cos\theta_{\ell-\gamma}^{CMS} < 0.95$
5. $0.15 < \cos\theta_{miss-\pi}^{CMS} < 0.97$
6. $\cos\theta_{\ell-\tau}^{R} < 0.68$
7. $-0.97 < \cos\theta_{\ell-\pi}^{CMS} < 0.42$
8. $0.87 < Thrust < 0.97$
9. 0.66 GeV/c $< P_{miss}$
10. $5.57 < E_{vis} < 10.2$ GeV
11. -0.13 GeV/c$^2 < M_{\nu}^2$

Events after Dual Annealing optimization

After considering the best combination of cuts above, the final number of events for the "train" and "test" samples are shown in Table 5.7. Here, it can be seen that compared with Table 5.5, we achieve an increase in the signal efficiency of 6.65% to 8.52%.

MC14 samples	Ann train	Ann test
$\tau^{\pm} \to \mu^{\pm}\gamma$	$\epsilon = 8.5\%$	$\epsilon = 8.52\%$
$\tau^{+}\tau^{-}$	191 ± 14	221 ± 15
$\mu^{+}\mu^{-}(\gamma)$	6 ± 2	9 ± 3
$q\bar{q}$	13 ± 4	10 ± 3
$e^{+}e^{-}(\gamma)$	0	0
$\ell^{+}\ell^{-}\ell^{+}\ell^{-}$	0	0
$B\bar{B}$	0	1 ± 1
Total BKG	210 ± 15	241 ± 16

Table 5.7: Final number of events of the "train" and "test" samples for the muon channel. The signal efficiency is also shown. Background samples correspond to 400 fb^{-1}.

5.3.4 Summary of the selection criteria

The selection criteria were obtained using the "train" sample with $\mathcal{L}_{int} = 400$ fb^{-1} of integrated luminosity. The summary of the event selection in Section 5.2 is presented in Table 5.8. Regarding the optimization, the summary for the cut-based and the Dual Annealing optimization is presented in Table 5.9.

	Selection criteria
Reconstruction level	
	No. good track = 2
tracks	Net charge = 0
	$\|dz\| < 3$ cm $\|dr\| < 0.5$ cm
	0.2 GeV $< E_\gamma$
photons	$-0.8660 < cos\theta_\gamma < 0.9535$
	clusterNHits > 1.5
	abs(clusterTiming) < 200 ns
	should not pass $M_{\gamma\gamma}$ cut of neutral pions
neutral pions	same that photons but with
	0.1 GeV$< E_\gamma$
	$0.115 < M_{\gamma\gamma} < 0.152$ GeV/c^2
LFV skim	$-1.5 < \Delta E < 0.5$ GeV
	$1 < M_{inv} < 2$ GeV/c^2
other selections	$P_t > 0.1$ GeV/c
	$-0.866 < cos(\theta_{miss}) < 0.9535$
Particle identification	
	muon (electron)
Signal track	$0.95 <$ muon ID ($0.95 <$ electron ID_noSVD_noTOP)
Tag track	muon ID < 0.5 (electron ID_noSVD_noTOP< 0.5)
Preselection level	
	muon (electron)
M_{inv}	[1.64,1.82] ([1.51,1.83])

Table 5.8: Summary of the event selection criteria before the optimization.

Variable	Cut base muon	Cut base electron	Ann muon
P_ℓ (GeV/c)	[—, 3.4]	[—, 3.5]	[—,4.04]
P_π (GeV/c)	[—, 4.4]	[—, 2.8]	[—,5.5]
E_γ (GeV)	[1.6, —]	[1.4, —]	[1.58, —]
$cos_{\ell-\gamma}$	[0.75, 0.78]	[0.24, 0.78]	[0.73, 0.95]
$cos_{miss-\pi}$	[0.11, 0.98]	[-0.04, 0.96]	[0.15, 0.97]
$cos^{R}_{\tau-\ell}$	[—, 0.3]	[—, 0.4]	[—,0.68]
$cos_{\ell-\pi}$	[—, 0.02]	[—, 0.25]	[-0.97, 0.42]
$thrust$	[0.9, 0.97]	[0.91, 0.96]	[0.87, 0.97]
P_{miss} (GeV/c)	[0.7, —]	[2.4, —]	[0.66,—]
$E^{CMS}_{visible}$ (GeV)	[5.6, 11.7]	[6.2, 8.9]	[5.57,10.2]
M^2_{miss} (GeV/c^2)	[-0.5, —]	[-0.5, —]	[-0.13,-]

Table 5.9: Summary of the optimized selection criteria.

5.4 Signal region definition

In the search of $\tau^{\pm} \to \ell^{\pm} + \gamma$, the candidates should have an invariant mass (Equation 5.2) close to the τ mass, $M_{inv} \sim m_\tau$ and an energy close to the beam energy, that is,

$$\Delta E = E^{CMS}_{\ell+\gamma} - E^{CMS}_{beam} \ , \tag{5.6}$$

should be close to zero, $\Delta E \simeq 0$.

These two variables are used to identify the signal. However, in order to select the final signal region efficiently, we perform a PCA-whitening (or sphering) to our data. This method guarantees two new uncorrelated variables, whit a covariance matrix equal to the identity. More details can be found in Appendix D.

The transformation is given by:

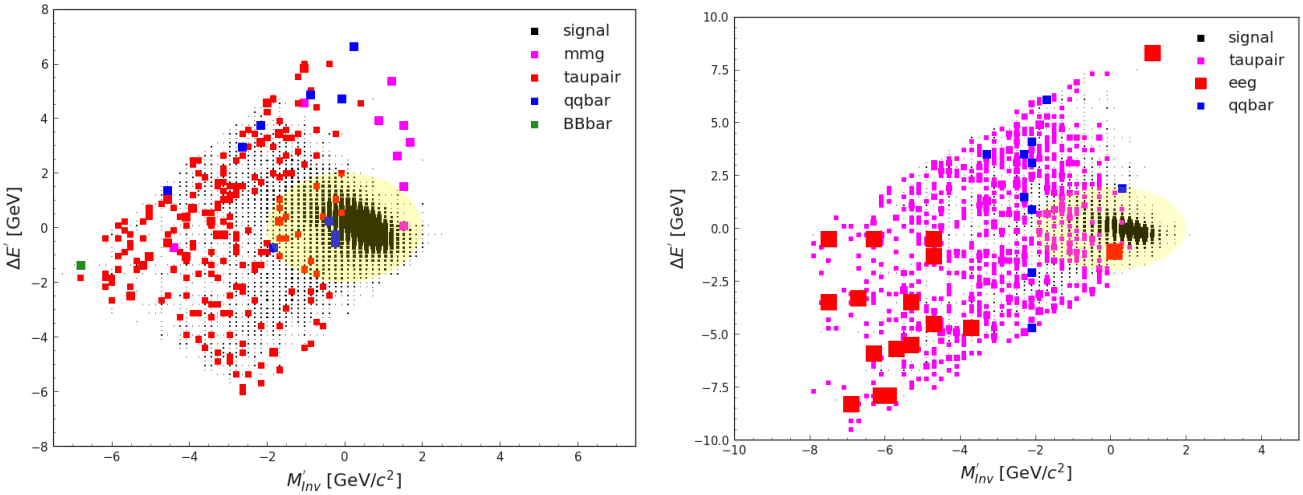
$$\begin{pmatrix} M^{'}_{inv} \\ \Delta E^{'} \end{pmatrix} = D^{-1/2} \cdot \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} M_{inv} \\ \Delta E \end{pmatrix} \ , \tag{5.7}$$

where D is the diagonal matrix with eigenvalues in the diagonal. For the muon channel, the rotation angle is $\theta = 78.3\ ^{\circ}$.

We define the signal region in terms of the variance (square of the standard deviation), choosing a circular region of radius of 2 ($r = \sigma^2 = 2$)

$$(M^{'}_{inv})^2 + (\Delta E^{'})^2 < (r = 2)^2 \ . \tag{5.8}$$

Figure 5.14 shows the signal region on the $M^{'}_{inv}$ and $\Delta E^{'}$ 2-D plane after the background reduction methods, for the electron and muon channels. In Table 5.10 we report the number of events inside the signal regions, calculated on the "test" sample.



(a) Dual Annealing muon channel.

(b) Cut-based electron channel.

Figure 5.14: Distributions of M'_{inv} vs $\Delta E'$. The right side distribution shows the remaining events for the muon channel after the Dual Annealing optimization. The remaining events for the electron channel are shown on the left side distribution after the cut-based optimization. The yellow circle (which may appear distorted by the plot dimensions) represents the signal region defined by Equation 5.8.

MC14 samples	Cut-based electron channel	Dual Annealing muon channel
$\tau^{\pm} \to \ell^{\pm}\gamma$	$\epsilon = 4.0\%$	$\epsilon = 7.56\%$
$\tau^{+}\tau^{-}$	56 ± 7	17 ± 4
$\mu^{+}\mu^{-}(\gamma)$	0	1 ± 1
$q\bar{q}$	0	3 ± 1
$e^{+}e^{-}(\gamma)$	8 ± 8	0
$\ell^{+}\ell^{-}\ell^{+}\ell^{-}$	0	0
$B\bar{B}$	0	0
Total BKG	64 ± 11	21 ± 5

Table 5.10: Final events of the "test sample", inside the signal region for the electron and muon channel.

5.5 Upper Limit strategy

In this section, we present a first sensitivity analysis, assuming zero signal observed events, to measure the upper limit on

$$\mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \gamma) \tag{5.9}$$

at 95% of C.L. using pseudo-data (MC toys) samples.

The statistical analysis is based on a frequentist approach using the asymptotic CLs technique [44] implemented using the RooStats and HistFactory statistical tools of the ROOT data analysis framework from CERN [45].

5.5.1 Statistical model

To measure the branching ratio of $\tau^{\pm} \to \ell^{\pm} + \gamma$, we model the pseudo data taking into account all the possible physics processes. For x and y discriminating variables, the model can be written as

$$\mathrm{F(x,y)} = N_{signal} \cdot S(x,y) + N_{background} \cdot B(x,y) \ , \tag{5.10}$$

where N_{signal} and $N_{background}$ are the respective yields for the signal events $\tau^{\pm} \to \ell^{\pm} + \gamma$ and SM processes (background). The corresponding probability density functions are $S(x,y)$ and $B(x,y)$.

The model satisfies $\int F(x,y) = N_{signal} + N_{background} = N_{total}$, the total number of observed events.

We can rewrite the expression above, using the following relation:

$$N_{signal} = \mathrm{Br}(\tau \to \ell\gamma) \cdot 2 \cdot \epsilon_{signal} \cdot \sigma(e^+e^- \to \tau^+\tau^-) \cdot \mathrm{L_{int}} \ . \tag{5.11}$$

Then,

$$\mathrm{F(x,y)} = \mathrm{Br}(\tau \to \ell\gamma) \cdot 2 \cdot \epsilon_{signal} \cdot \sigma(e^+e^- \to \tau^+\tau^-) \cdot \mathrm{L_{int}} \cdot S(x,y) + N_{background} \cdot B(x,y) \tag{5.12}$$

In this way, we can extract the value of the branching ratio directly by fitting the model to the pseudo data, with N_{signal} and $N_{background}$ allowed to float in the fit.

The following values of the involved parameters are used for the muon and the electron channels:

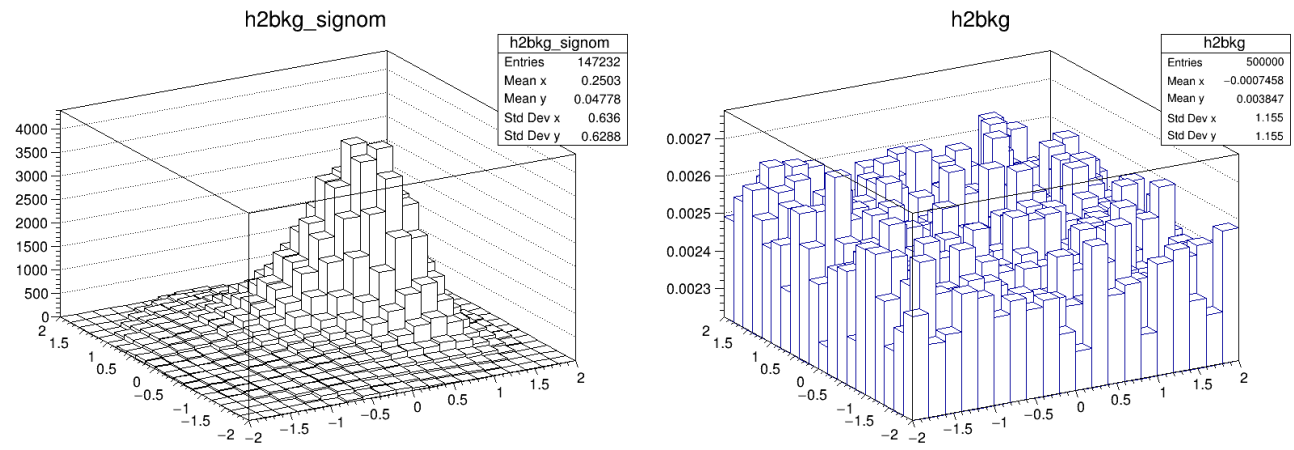
- $L = 400.0$ fb^{-1} ;
- $\sigma(e^+e^- \to \tau^+\tau^-) = 0.919$ nb ;
- $\varepsilon_{\mu\gamma} = 7.31$ ($\varepsilon_{e\gamma} = 3.81$) ;
- $N_{background} = 18(57)$.

The uncertainties on these parameters are considered in Section 5.6.

The signal efficiency and the number of expected background events are slightly different from Table 5.10, due to extra requirements that we are considering, these are $\#\gamma_{\mathrm{signal-side}} = 1$ (Section 5.2.6), the trigger (Secction 5.2.4), and the Lepton ID corrections of Section 5.5.2.

In this analysis, the signal probability density function in Equation 5.12 is a 2-D histogram template obtained from the signal MC sample. Regarding the background probability density function, we assume it is a 2D uniform distribution. Then, we generate random uniform values (500k) to create a histogram template.

An example of the template distributions for the muon channel are shown in Figure 5.15. In this example, the signal template was obtained after all the requirements of the Dual Annealing optimization.



(a) (b)

Figure 5.15: 2D normalized templates of M'_{inv} vs $\Delta E'$. For the muon channel, shown are (a) the signal template obtained from signal MC simulation after the Dual Annealing optimization, and (b) the background template generated from a uniform distribution. The distribution binned size is 0.2.

5.5.2 Corrections

To ensure the correct calculation of the sensitivity of the measurement to the upper limit. The following corrections are applied to MC:

- The electron and muon ID efficiencies. As the performance of the lepton ID is different in MC simulation and data, correction factors are needed to take the discrepancies in lepton identification efficiency into account. These factors are taken from the tables provided by the Belle II Lepton ID Performance Group [46].

- The $\pi \to e$ and $\pi \to \mu$ fake rates. For the difference performance in data and MC, we apply the corrections due to the probability that a charged pion ($\pi^{\pm}$) is misidentified as a charged lepton ($\ell^{\pm}$). These factors are taken from the tables provided by the Belle II Lepton ID Performance Group [46].

The correction factors were applied to the binned MC distributions, as weights in each bin. The changes in the signal efficiency are taken into account in the calculations of this section.

However, there are corrections that we are not considering but will be applied in the future as mentioned in Chapter 7.

5.5.3 Upper limit on pseudo-data

To calculate the sensitivity to measure the upper limit on $\mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \gamma)$, without taking into account uncertainties, we generate a set of 1000 independent pseudo datasets (MC toys) corresponding to the background-only hypothesis. Then, we calculate the upper limit on each pseudo-data using the model described in the previous Section 5.5.1, where the involved parameter are fixed to their nominal values.

We present in Table 5.11 the mean of the upper limit calculation for the asymptotic implementation of the CLs method, for both the muon and electron channels.

Channel	Expected UL $\times 10^{-8}$ at 90 C.L.
muon	9.29
electron	23.52

Table 5.11: Upper limit calculation at 90 % C.L. without taking into account systematics uncertainties.

5.6 Systematic uncertanties

In order to accurately compute the upper limit on the branching ratio $\mathcal{B}(\tau \to \ell\gamma)$ using the statistical model presented in Equation 5.12, it is necessary to take into account both statistical and systematic uncertainties associated to the relevant parameters.

Integrated luminosity L_{int}

We consider as systematic uncertainty due to the knowledge of $\mathcal{L}_{int}$, the estimated percent uncertainty of 0.6% presented in [57]. The measurement of $\mathcal{L}_{int}$ was made using Bhabha and digamma events in Phase 2 (2018) data.

Cross section $\sigma(e^+e^- \to \tau^+\tau^-)$

The uncertainty of the tau pair cross section was taken from the calculation in [58], being $\sigma(e^+e^- \to \tau^+\tau^-) = 0.919 \pm 0.0033$ nb, corresponding to 0.033%.

Number of background events $N_{background}$

The number of background events is extracted from MC. Here, we consider the uncertainty propagation of the different processes involved in the calculation of the final events.

For example, for the muon channel, we have

$$N_{BKG} = N_{\mu\mu} + N_{\tau\tau} + N_{q\bar{q}} \,, \tag{5.13}$$

where $N_i = L_{int} \cdot \epsilon_i \cdot \sigma_i$ is the number of events of the i-sample, then, we can write

$$N_i = L_{int} \cdot (\epsilon_{\mu\mu} \cdot \sigma_{\mu\mu} + \epsilon_{\tau\tau} \cdot \sigma_{\tau\tau} + \epsilon_{q\bar{q}} \cdot \sigma_{q\bar{q}}) \tag{5.14}$$

where the $\epsilon_{\mu\mu} = 2.2 \times 10^{-9}$, $\epsilon_{\tau\tau} = 3.8 \times 10^{-8}$, $\epsilon_{q\bar{q}} = 4.7 \times 10^{-9}$, and the cross sections are given in Table 3.1. The uncertainty of each parameter is considered.

Signal efficiency

-Lepton Identification

The signal identification involves requirements on the global likelihood lepton ID, as was mentioned in Section 5.2.3. Since the performance of the lepton ID is different in MC simulation

and data, correction factors must be applied to the MC. They consist of lepton identification efficiency and $\pi^{\pm} \to \ell^{\pm}$ fake rate.

The lepton ID efficiency and fake rate corrections factors are provided by the Belle II Lepton ID performance group and have associated statistical and systematic uncertainties binned in momentum p_{lab} and polar angle θ_{lab} of the lepton.

We require the events to have a global LID weight correction, that is, a positive global LID weight. This condition modifies the number of signal events, hence, the signal efficiency.

The statistical and systematic uncertainties of this correction are combined on quadrature:

$$\sigma_{\text{lid}} = \sqrt{(\sigma_{\text{lid}}^{\text{stat}})^2 + (\sigma_{\text{lid}}^{\text{sys}})^2}. \tag{5.15}$$

For the muon channel, $\sigma_{\text{lid}} = 0.00228$, or 3.08%.

-Trigger efficiency

For data and MC simulation, we use a combination of the *lmlX* triggers presented in Section 5.2.4 to select the events of interest. In MC, the events are selected using the trigger simulation TSIM and this selection affects the final efficiency.

At the moment, we are not considering an associated systematic uncertainty due to the trigger selection.

We performed a calculation of the trigger efficiency in side bands data (see Section 5.7) and MC under the same conditions, and we found consistency within $\sim 2\sigma$:

$$\sigma_{MC}^{trigger} = 96.24 \pm 0.26\%$$

$$\sigma_{Data}^{trigger} = 95.50 \pm 0.25\%$$

-Uncertainty due to MC statistics

After considering the lepton identification, trigger efficiencies, and the requirement of $\gamma_{signal} = 1$, the total signal efficiency is 7.31% and 3.81 %, for the muon and electron channels, respectively.

We calculate the uncertainty of the total efficiency using the binomial error formula:

$$\sigma_{\epsilon}^{stat} = \left(\frac{1}{N}\right) \cdot \sqrt{k\left(1 - \frac{k}{N}\right)} \tag{5.16}$$

where k is the expected number of events passing the cuts and N is the generated sample size [59].

For the muon channel, this uncertainty is $\sigma_{\epsilon}^{stat} = 0.000184$, corresponding to 0.25%. For the electron channel, $\sigma_{\epsilon}^{stat} = 0.0001386$, that is 0.35%. We take these numbers as systematic uncertainties due to the generated MC statistics.

5.6.1 Summary of the uncertainties

Parameters	Percentage muon channel	Percentage electron channel
$N_{background}$	23.58%	13.25%
L_{int}	0.6%	0.6%
$\sigma(e^{\pm}e^{\pm} \to \tau^{+}\tau^{-})$	0.33%	0.33%
ϵ_{signal}	3.08%	2.48%
Total percentage	3.17%	2.58%

Table 5.12: Summary of the systematic uncertainties of the involved parameters on the upper limit calculation.

5.6.2 Upper limit on pseudo-data considering systematics uncertainties

The associated uncertainty of both, the product of $\epsilon_{signal} \cdot \sigma(e^{\pm}e^{\pm} \to \tau^{+}\tau^{-}) \cdot L_{int}$ and the expected number of background events $N_{background}$, are introduced in the model of Equation 5.12 as new nuisance parameters that follows Gaussian constraints.

Channel	Expected UL $\times 10^{-8}$ at 90 % C.L.
muon	9.88
electron	24.55

Table 5.13: Upper limit calculation at 90 % C.L. taking into account systematics uncertainties.

5.7 Data-MC validation in sidebands

In this section, we checked the Data-MC agreement following a blind procedure, at the moment focusing only on the $\tau^{\pm} \to \mu^{\pm}\gamma$ channel. To define a sidebands (control) region, it is useful to calculate the signal resolution of the invariant mass variable, as shown in Figure 5.16 after the Dual Annealing optimization in Section 5.3.3. We compare outside a 10σ region, where σ is found by a fit of asymmetric Gaussians to be $\sigma_{left} = 34.29$ MeV/c^2 and $\sigma_{right} = 12.24$ MeV/c^2, whit a mean of 1.7778 GeV/c^2. Thus, the sideband region is defined as $M_{inv} < 1.54$ GeV/c^2 and 1.9 GeV/c^2 $< M_{inv}$.

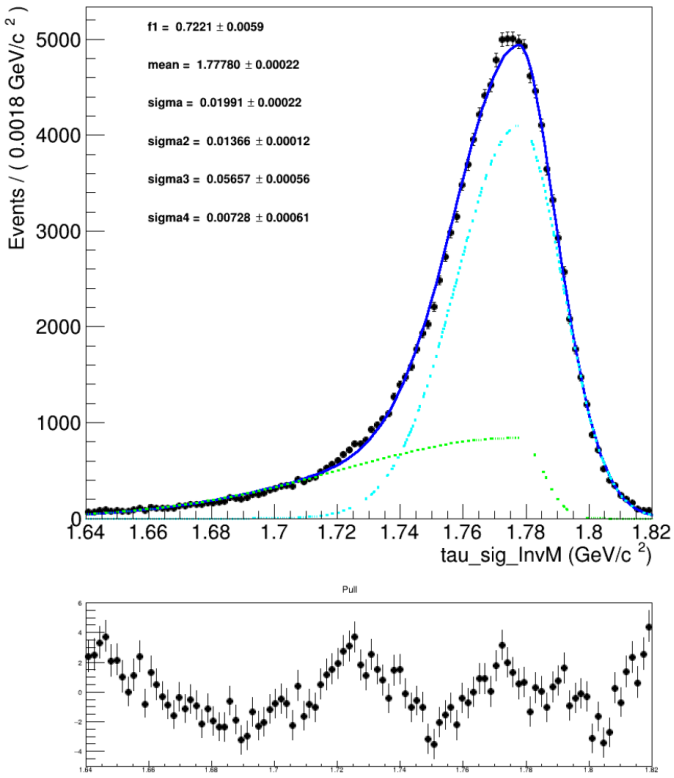


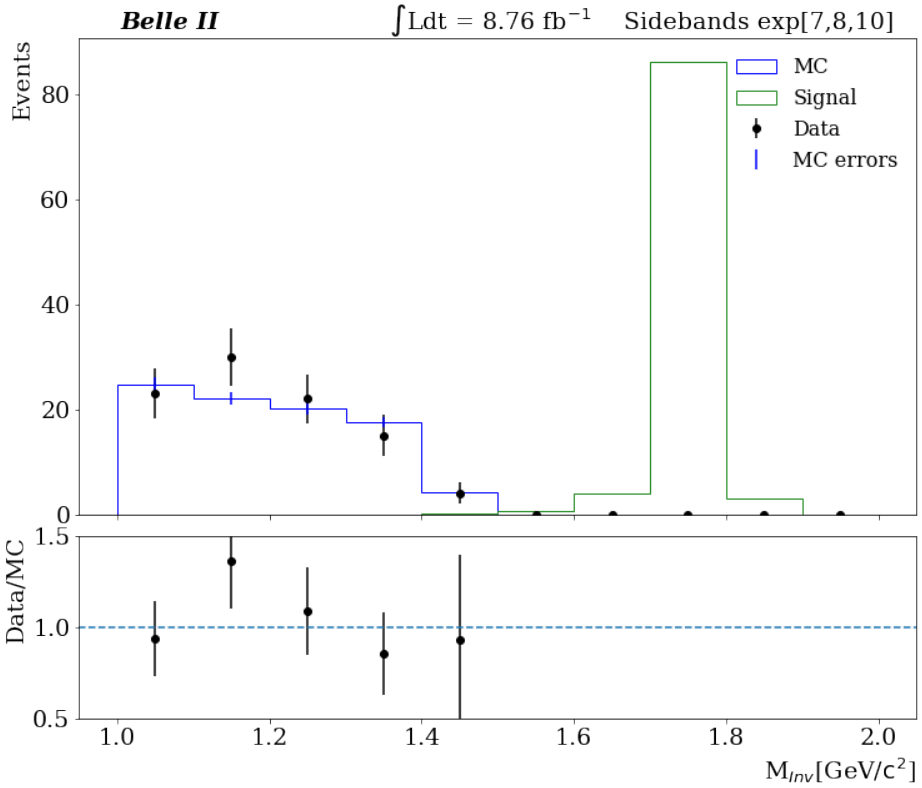
Figure 5.16: Signal fit of M_{inv} for the $\tau^{\pm} \to \mu^{\pm}\gamma$ channel after the Dual Annealing optimization.

In the plots presented in this section, the MC error band includes the following uncertainties summed in quadrature:

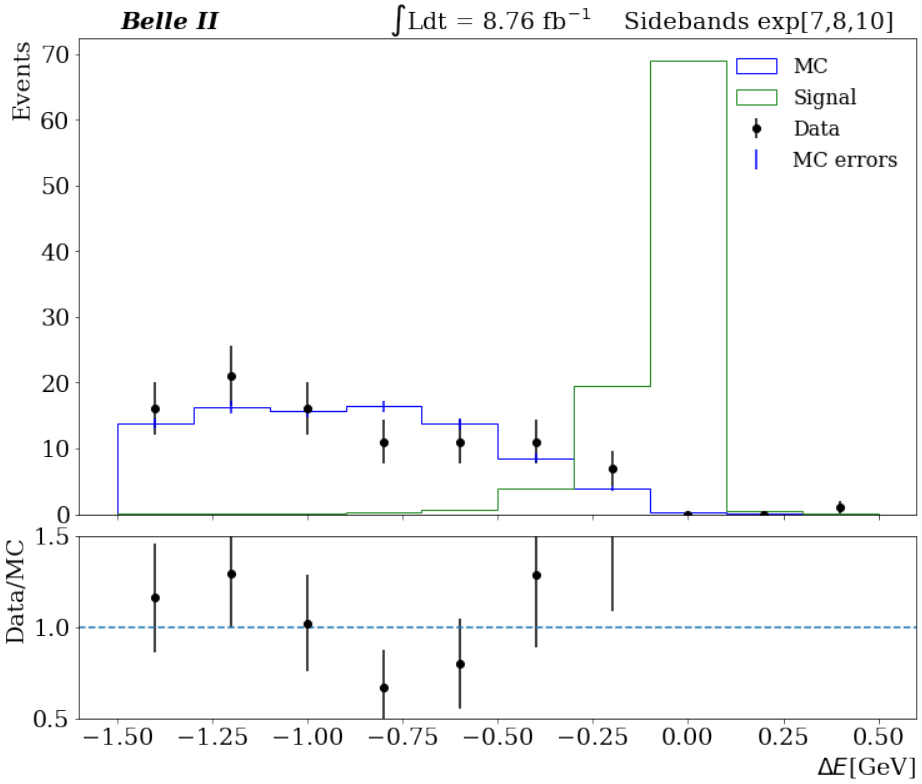
- MC statistical uncertainty.
- Lepton ID efficiency correction uncertainties.
- $\pi \to \ell$ fake-rate correction uncertainties.

The data satisfy the trigger selection mentioned in Section 5.2.4, as well as, all the requirements of the MC samples. The uncertainty in data is statistical. The energy and momentum scale uncertainties in data will be added in these comparisons in the future.

In Figure 5.17, the distributions of M_{inv} and ΔE for sideband events are shown. The other variables of interest are shown in Figure 5.18 and 5.19. In each figure, the signal MC simulation is plotted as a reference. We find good agreement between data and MC in sidebands.

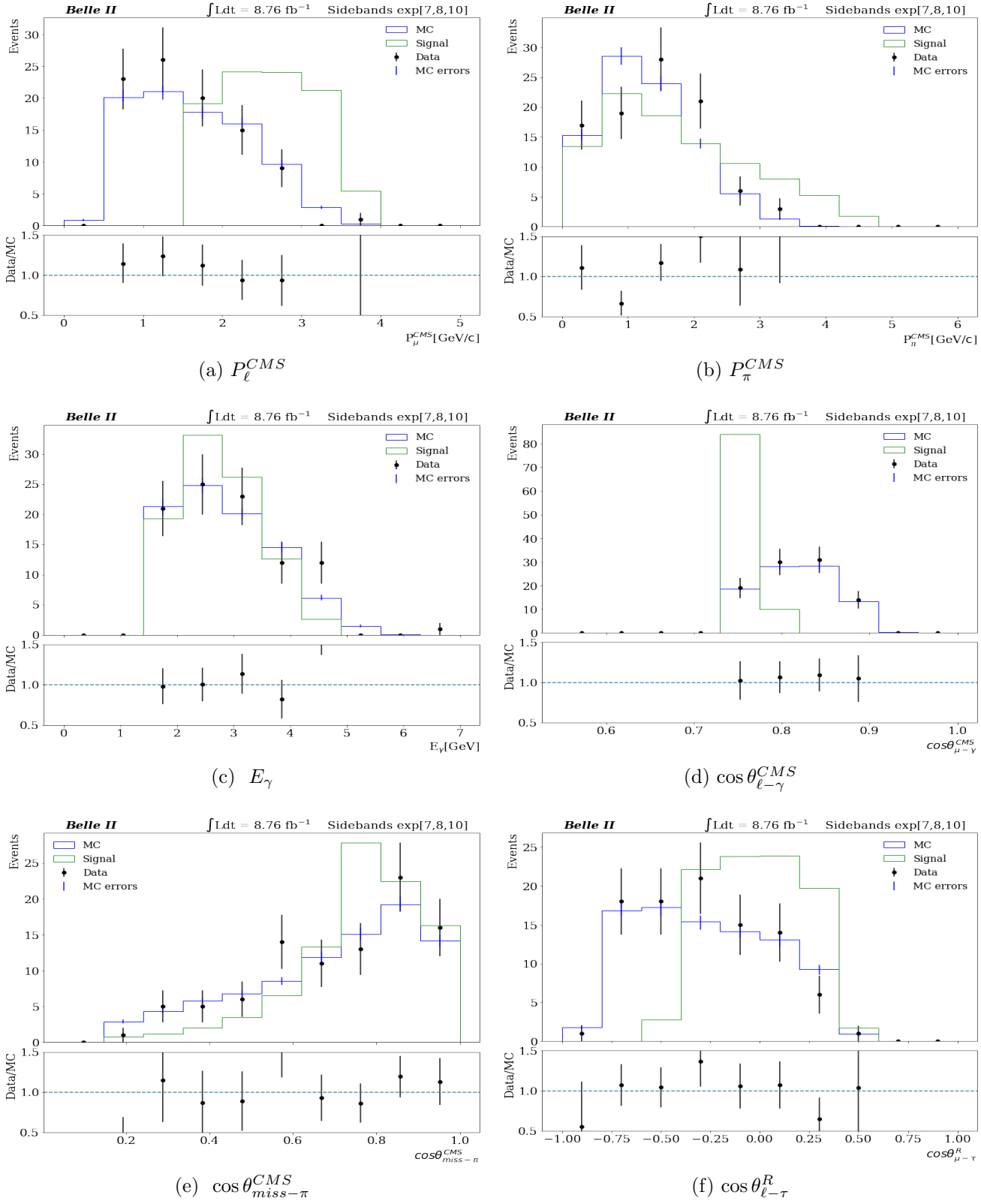


(a) M_{inv}



(b) ΔE

Figure 5.17: Data vs MC distributions in sidebands of (a) the invariant mass M_{inv} of the $\mu\gamma$ system and (b) the energy difference ΔE. The signal MC simulation is plotted in green as a reference.

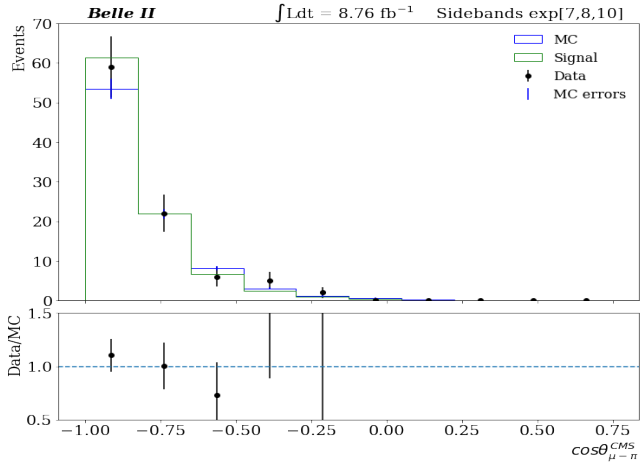


(a) P_ℓ^{CMS} (b) P_π^{CMS}

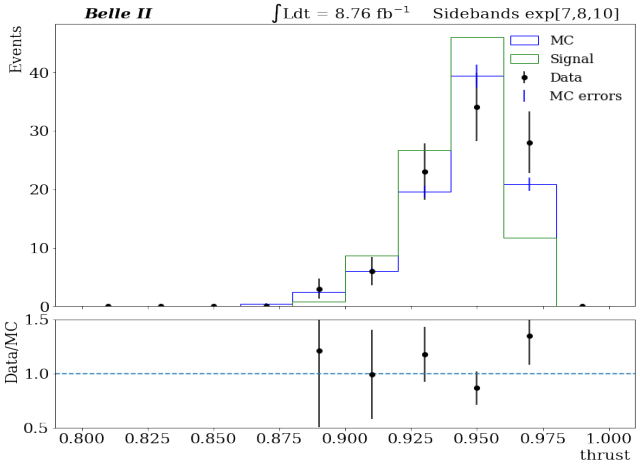
(c) E_γ (d) $\cos\theta_{\ell-\gamma}^{CMS}$

(e) $\cos\theta_{miss-\pi}^{CMS}$ (f) $\cos\theta_{\ell-\tau}^{R}$

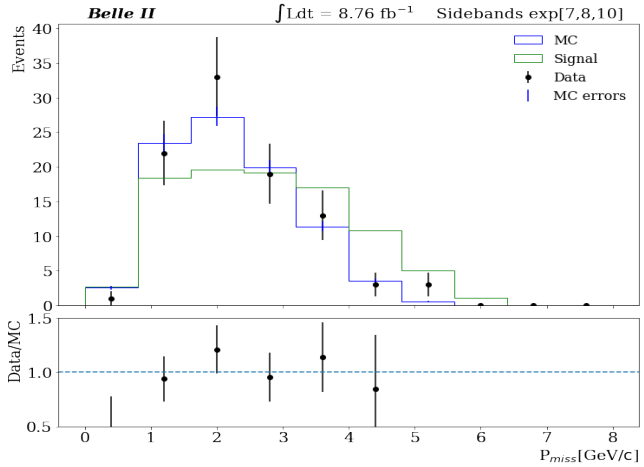
Figure 5.18: Data vs MC distributions in sidebands of the different discriminative variables described in Section 5.3.1. The signal MC simulation is plotted in green as a reference.



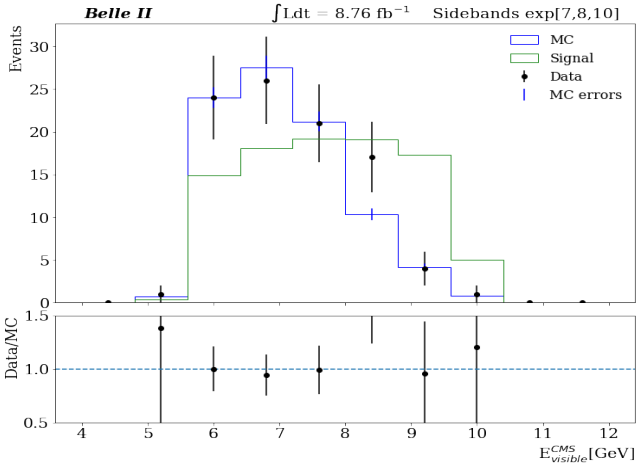
(a) $\cos\theta^{CMS}_{\ell-\pi}$



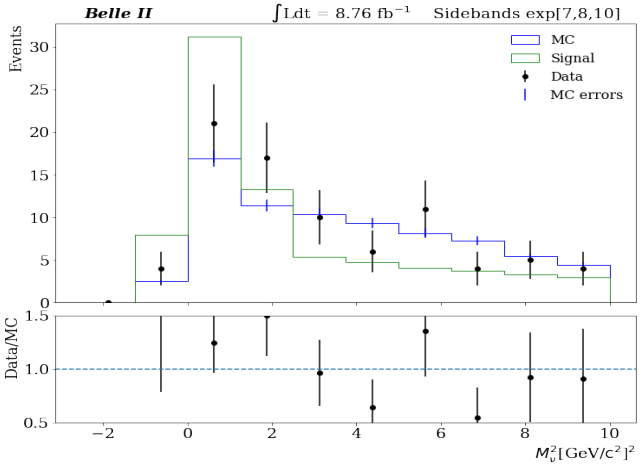
(b) *Thrust*



(c) P_{miss}



(d) E_{vis}



(e) M^2_{ν}

Figure 5.19: Data vs MC distributions in sidebands of the different discriminative variables described in Section 5.3.1. The signal MC simulation is plotted in green as a reference.

5.8 Results

The search of the $\tau^{\pm} \to \ell^{\pm} + \gamma$ decay is still an ongoing analysis, and at this moment we present the expected upper limits calculation on the branching ratio using pseudo data.

We determine the upper limits using the CLs method and following the details given in Section 5.5. The systematic uncertainties discussed in Section 5.6 are considered.

The expected upper limit results at 90 % C.L., considering an integrated luminosity of 400 fb^{-1} are:

$$\mathcal{B}(\tau^{\pm} \to \mu^{\pm} + \gamma) < 9.88 \times 10^{-8} \tag{5.17}$$

$$\mathcal{B}(\tau^{\pm} \to e^{\pm} + \gamma) < 24.55 \times 10^{-8} \tag{5.18}$$

Our result for the muon channel is twice the current Belle limit, shown in Table 2.2, which uses 988 fb^{-1} of integrated luminosity.

However, other sources of systematic uncertainties need to be considered. The Data-MC comparison needs to be updated with all relevant corrections and uncertainties. On the other hand, the electron channel requires several improvements. These tasks will be part of the prospects, which will be discussed further in Chapter 7.

Chapter 6

Conclusions

Throughout this research work, we have searched for new physics using the lepton flavor violating decays $\tau^{\pm} \to \ell^{\pm} + \alpha$ and $\tau^{\pm} \to \ell^{\pm} + \gamma$.

To search for the $\tau^{\pm} \to \ell^{\pm} + \alpha$ decay we used an integrated luminosity of 62.8 fb^{-1} collected with the Belle II detector. The search strategy was based on the ARGUS technique, looking for an excess of the $\tau^{\pm} \to \ell^{\pm} + \alpha$ signal on top of the SM processes in the lepton-energy spectrum variable computed in the τ pseudo-rest frame, $x_{prf} = 2E_{\ell}/m_{\tau}$. We achieved a purity of 91.95% and a signal efficiency between $9.1 - 17.4\%$ for the muon channel. For electron, we achieved a purity of 95.85% and a signal efficiency between $9.4 - 13.9\%$. It is worth mentioning that the validation of the MC data helped us to include pertinent corrections and improvements before making the final measurement. We "burned" 10% of the data and sidebands, and further selections on photons and the missing momentum polar angle were applied. Upon opening the entire box, no excess was found in the Data/MC comparison.

We calculated the upper limit on the ratio $\mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \alpha)/\mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \nu\bar{\nu})$, for $\ell = \mu, e$, using the asymptotic CLs, where binned extended likelihood 1D fits were performed. As a novel approach, we included systematics uncertainties in the model as shape-correlated nuisance parameters. The dominant sources of systematics uncertainties were due to lepton ID efficiency and fake rate, as well as the trigger efficiency. On average, the systematics uncertainties degraded our limit sensitivity by approximately 35% in both channels. We report 90 and 95% C.L. upper limits on the branching-fraction ratio $\mathcal{B}(\tau^{\pm} \to e^{\pm} + \alpha)/\mathcal{B}(\tau^{\pm} \to e^{\pm} + \nu\bar{\nu})$ in the range $(1.1 - 9.7) \times 10^{-3}$ and on $\mathcal{B}(\tau^{\pm} \to \mu^{\pm} + \alpha)/\mathcal{B}(\tau^{\pm} \to \mu^{\pm} + \nu\bar{\nu})$ in the range $(0.7 - 12.2) \times 10^{-1}$ for α masses between 0 and 1.6 GeV/c^2. The 95% C.L. upper limits represent a 2.2 to 14 improvement depending on the α mass compared with the previous bounds presented by the ARGUS collaboration.

It is possible to update these limits with the current Belle II integrated luminosity of 424 ± 3 fb^{-1}. Furthermore, new methods are under development that can increase the limits up to one order of magnitude, and there is also the possibility to perform this search using the 1×1 prong topology [60].

On the other hand, the search of $\tau^{\pm} \to \ell^{\pm} + \gamma$ is planned to be carried out with all the data collected so far by the Belle II detector, which is about 400 fb^{-1} at the $\Upsilon(4S)$ resonance.

For the $\tau^{\pm} \to \mu^{\pm} + \gamma$ channel, we were able to increase our signal efficiency by using a Dual Annealing optimization. The signal efficiencies in the signal region after all the selections are 7.31% and 3.81% for the muon and electron channels, respectively.

We calculated the expected upper limit on the branching ratio $\mathcal{B}(\tau^{\pm} \to \ell^{\pm} + \gamma)$ using the asymptotic CLs technique, where binned extended likelihood 2D fits were performed. Systematic uncertainties were considered for the parameters involved in the branching ratio calcula-

tion, including the signal efficiency ϵ_{signal}, the integrated luminosity L_{int}, and the cross section $\sigma(e^+e^- \to \tau^+\tau^-)$. The percent uncertainties on the product $\epsilon_{signal} \cdot \sigma(e^+e^- \to \tau^+\tau^-) \cdot L_{int}$ are 3.17% and 2.58% for the muon and electron channels, respectively. Additionally, we estimate the systematic uncertainty on the expected number of background events $N_{background}$ based on MC inputs. The associated uncertainties of both sources are included in the model as nuisance parameters that follow a Gaussian constraint.

The expected upper limits corresponding to 400 fb^{-1} are $\mathcal{B}(\tau^\pm \to \mu^\pm + \gamma) < 9.88 \times 10^{-8}$ and $\mathcal{B}(\tau^\pm \to e^\pm + \gamma) < 24.55 \times 10^{-8}$ at 90% C.L. Our results represent the most updated version of the search for $\tau^\pm \to \ell^\pm + \gamma$ within Belle II. For the muon channel, the result is twice the current Belle limit, which uses 988 fb $^{-1}$ (2.5 times our luminosity). Regarding the electron channel, better background suppression of $e^+e^-\gamma$ (Bhabha) events is needed but without affecting the signal efficiency. Therefore, before proceeding with the calculation on data, several improvements need to be considered.

Chapter 7

Prospects

Looking ahead, there are several promising directions in the search of the golden channel $\tau^{\pm} \to \ell^{\pm} + \gamma$, building on the contributions of this work. In this chapter, we will mention some:

- Systematics uncertainties need to be considered due to trigger and tracking efficiencies, the momentum scale, and the beam energy. Although some of these were negligible in the $\tau^{\pm} \to \ell^{\pm} + \alpha$ analysis, they need to be reassessed for $\tau^{\pm} \to \ell^{\pm} + \gamma$.
- Moreover, systematic uncertainties associated with the signal photon, such as the photon energy correction and photon resolution calibration, need to be taken into account.
- We need to update the Data/MC comparison for both the muon and electron channels with all the available data collected so far by the Belle II detector.

For the electron channel:

- We need to apply the Bremsstrahlung correction, which can be important to increase the efficiency and suppress further backgrounds.
- Specifically, new methods need to be explored to suppress the $e^+e^-(\gamma)$ background.

Based on the latest results from Belle:

- There are new promising variables that we can explore, such as the cosine of the angle between the tau and the track on the tag side, and the asymmetric energy between the lepton and the photon on the signal side.
- It seems convenient to change from M_{inv} to M_{bc} (beam constrained mass).
- Separating the sample in different 1-prong tags (μ, e, π and ρ) could also improve the background suppression.

Finally:

- A parallel effort is being carried out to analyze the 3×1 prong topology. Preliminary results in the muon channel show that the signal efficiency in this topology is about 2.5%, which could be reflected as a 20% improvement in the combined limit. Therefore, it is worth to continue these investigations in the future.

Bibliography

[1] Standard model of elementary particles. https://upload.wikimedia.org/wikipedia/commons/0/00/Standard_Model_of_Elementary_Particles.svg, 2021. Accessed: May 10, 2023.

[2] The Belle II Collaboration. Belle II experiment sensitivity to the LFV decay $\tau \to \ell+\alpha$. https://docs.belle2.org/record/2043/files/BELLE2-NOTE-PL-2020-018.pdf, Sep 2021.

[3] L. Calibbi, D. Redigolo, R. Ziegler, and J. Zupan. Looking forward to lepton-flavor-violating ALPs. *J. High Energy Phys*, 09:173, 2021.

[4] The BABAR Collaboration. Searches for lepton flavor violation in the decays $\tau^{\pm} \to e^{\pm}\gamma$ and $\tau^{\pm} \to \mu^{\pm}\gamma$. *Phys. Rev. Lett. 104, 021802 (2010)*, arXiv:0908.2381.

[5] The Belle collaboration. Search for lepton-flavor-violating tau-lepton decays to $\ell\gamma$ at Belle. *Journal of High Energy Physics*, 2021(10):19, Oct 2021.

[6] Belle II Collaboration. Super KEKB and Belle II. https://www.belle2.org/project/super_kekb_and_belle_ii/, 2023.

[7] W. Altmannshofer et al. The Belle II Physics Book. *PTEP*, 2019(12):123C01, 2019. [erratum:PTEP2020,no.2,029201(2020)].

[8] Valerio Bertacchi et al. Track Finding at Belle II. *Computer Physics Communications*, 259:107610, Feb 2021.

[9] T. Abe et al. Belle II Technical Design Report. *arXiv:1011.0352*, 11 2010.

[10] Swagato Banerjee, Vincenzo Cirigliano, Mogens Dam, Abhay Deshpande, Luca Fiorini, Kaori Fuyuto, Ciprian Gal, Tomáš Husek, Emanuele Mereghetti, Kevin Monsálvez-Pozo, Haiping Peng, Francesco Polci, Jorge Portolés, Armine Rostomyan, Michel Hernández Villanueva, Bin Yan, Jinlong Zhang, and Xiaorong Zhou. Snowmass 2021 white paper: Charged lepton flavor violation in the tau sector. *arXiv:2203.14919*, 2022.

[11] E. De La Cruz Burelo, A. De Yta Hernandez, M. Garcia Hernandez, I. Heredia de la Cruz, P. Rados, A. Rostomyan, F. Tenchini, and T. Kraetzschmar. Search for lepton flavor violating decay $\tau \to \ell + \alpha$. https://docs.belle2.org/record/1326/files/BELLE2-NOTE-PH-2019-009-v9.pdf, Mar 2019.

[12] I. et al. Adachi. Search for lepton-flavor-violating τ decays to a lepton and an invisible boson at Belle II. *Phys. Rev. Lett.*, 130:181803, May 2023.

[13] R. L. Workman et al. Review of Particle Physics. *PTEP*, 2022:083C01, 2022.

[14] G. Hernández-Tomé, G. López Castro, and P. Roig. Flavor violating leptonic decays of τ and μ leptons in the standard model with massive neutrinos. *The European Physical Journal C*, 79(1):84, 2019.

[15] S. T. Petcov. The Processes $\mu \to e+\gamma, \mu \to e+\overline{e}, \nu' \to \nu+\gamma$ in the Weinberg-Salam Model with Neutrino Mixing. *Sov. J. Nucl. Phys.*, 25:340, 1977. [Erratum: Sov.J.Nucl.Phys. 25, 698 (1977), Erratum: Yad.Fiz. 25, 1336 (1977)].

[16] Latika Aggarwal and others. Snowmass white paper: Belle II physics reach and plans for the next decade and beyond. *arXiv:2207.06307*, 2022.

[17] Z. G. Berezhiani and M. Y. Khlopov. Cosmology of Spontaneously Broken Gauge Family Symmetry. *Z. Phys. C*, 49:73–78, 1991.

[18] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm. Axionlike Particles, Lepton-Flavor Violation, and a New Explanation of a_μ and a_e. *Phys. Rev. Lett.*, 124(21):211803, 2020.

[19] C. Cornella, P. Paradisi, and O. Sumensari. Hunting for ALPs with Lepton Flavor Violation. *J. High Energy Phys*, 01:158, 2020.

[20] Andrea Brignole and Anna Rossi. Lepton flavour violating decays of supersymmetric higgs bosons. *Phys.Lett. B566 (2003) 217-225 hep-ph/0304081 DFPD-03-TH-11.*

[21] Takeshi Fukuyama and Tatsuru Kikuchi. Lepton flavor violating processes and muon $g-2$ in minimal supersymmetric $so(10)$ model. *Phys. Rev. D68, 033012 (2003)*, arXiv:hep-ph/0304190.

[22] Alejandro De Yta Hernández. Search for the decay channel $\tau^\pm \to \ell^\pm\alpha$ at the Belle II experiment. Cinvestav-IPN. Ph.D. Theses.

[23] Thomas Kraetzschmar. Search for an Invisible Boson in Tau Decays with Early Belle II Data and Development of New Analysis Methods for Tau Physics. `https://mediatum.ub.tum.de/doc/1684577/1684577.pdf`. Ph.D. Theses.

[24] R. M. Baltrusaitis et al. τ Leptonic Branching Ratios and a Search for Goldstone Decay. *Phys. Rev. Lett.*, 55:1842, 1985.

[25] H. Albrecht et al. A Search for lepton flavor violating decays $\tau \to e\alpha$, $\tau \to \mu\alpha$. *Z. Phys. C*, 68:25–28, 1995.

[26] Alejandro Ibarra, Marcela Marín, and Pablo Roig. Flavor violating muon decay into an electron and a light gauge boson. *Physics Letters B*, 827:136933, 2022.

[27] Michael Dine, Willy Fischler, and Mark Srednicki. A simple solution to the strong cp problem with a harmless axion. *Physics Letters B*, 104(3):199–202, 1981.

[28] C.D. Froggatt and H.B. Nielsen. Hierarchy of quark masses, cabibbo angles and cp violation. *Nuclear Physics B*, 147(3):277–298, 1979.

[29] J. Schechter and J. W. F. Valle. Neutrino decay and spontaneous violation of lepton number. *Phys. Rev. D*, 25:774–783, Feb 1982.

[30] Y. Chikashige, R.N. Mohapatra, and R.D. Peccei. Are there real goldstone bosons associated with broken lepton number? *Physics Letters B*, 98(4):265–268, 1981.

[31] Benjamin W. Lee and Robert E. Shrock. Natural suppression of symmetry violation in gauge theories: Muon- and electron-lepton-number nonconservation. *Phys. Rev. D*, 16:1444–1473, Sep 1977.

[32] Monika Blanke, Andrzej J. Buras, Björn Duling, Anton Poschenrieder, and Cecilia Tarantino. Charged lepton flavour violation and $(g-2)$ μ in the littlest higgs model with t-parity: a clear distinction from supersymmetry. *JHEP 0705, 013 (2007)*, arXiv:hep-ph/0702136.

[33] G. Cvetič, C. Dib, C. S. Kim, and J. D. Kim. On lepton flavor violation in tau decays. *Phys. Rev. D66, 034008 (2002)*, arXiv:hep-ph/0202212.

[34] Chong-xing Yue, Yan-ming Zhang, and Lan-jun Liu. Nonuniversal gauge bosons z-prime and lepton flavor violation tau decays. *Phys. Lett. B547 252-256 (2002).*

[35] Antonio Masiero, Sudhir K. Vempati, and Oscar Vives. See-saw and lepton flavour violation in susyso(10). *Nuclear Physics B*, 649(1):189–204, 2003.

[36] SuperB Collaboration. Superb: A high-luminosity asymmetric e+ e- super flavor factory. conceptual design report. *arXiv:0709.0451*, 2007.

[37] Belle II Analysis Software Framework (basf2). `https://doi.org/10.5281/zenodo.5574115`.

[38] Swagato Banerjee, Bolek Pietrzyk, J. Michael Roney, and Zbigniew Was. Tau and muon pair production cross sections in electron-positron annihilations at $\sqrt{s} = 10.58$ GeV. *Phys. Rev. D*, 77:054012, Mar 2008.

[39] Alejandro De Yta Hernández. Tau physics results and prospects at Belle II. *PoS*, EPS-HEP2021:527, 2022.

[40] P. Rados. Tracking performance confluence page. `https://confluence.desy.de/display/BI/Tracking+and+Vertexing+Performance`. Accessed: 2021-08-19.

[41] S. Longo. Neutrals performance confluence page. `https://confluence.desy.de/display/BI/Neutrals+Performance`. Accessed: 2021-08-19.

[42] M. Milesi. Lid corrections for eps 2021. `https://confluence.desy.de/display/BI/Recommendations+for+LeptonID+-+EPS+2021`. Accessed: 2021-08-19.

[43] Belle II Collaboration. τ lepton mass measurement at Belle II. *arXiv:2008.04665*, 2021.

[44] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011. [Erratum: Eur.Phys.J.C 73, 2501 (2013)].

[45] L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf. The RooStats Project. *PoS*, ACAT2010:057, 2010.

[46] Belle II collaboration. Recommendations for leptonid–moriond 2021. `https://confluence.desy.de/display/BI/Recommendations+for+LeptonID+-+Moriond+2021`, June 2021. Last accessed on 31.10.2021.

[47] Belle II Collaboration. Conference readiness. https://confluence.desy.de/display/BI/Conference+readiness, June 2021. Last accessed on 31.10.2021.

[48] T. Ferber. Neutrals performance confluence page. https://confluence.desy.de/display/BI/Conference+readiness. Accessed: 2021-08-19.

[49] TH1 - root 6.26/06 reference guide. https://root.cern.ch/doc/master/classTH1.html#aeadcf087afe6ba203bcde124cfabbee4, June 2021. Last accessed on 31.08.2021.

[50] Zuzana Gruberova, Zdenek Dolezal, Petar Rados, and Ami Rostomyan. Study of the τ lepton decays at the Belle II experiment. *BELLE2-MTHESIS-2021-007*, June 2021.

[51] Belle II collaboration. Tracking and vertexing performance. https://confluence.desy.de/display/BI/Tracking+and+Vertexing+Performance, June 2021. Last accessed on 31.10.2021.

[52] Jerome H. Friedman. Data Analysis Techniques for High-Energy Particle Physics. In *3rd CERN School of Computing*, page 271, 10 1974.

[53] Y Xiang, D.Y Sun, W Fan, and X.G Gong. Generalized simulated annealing algorithm and its application to the thomson model. *Physics Letters A*, 233(3):216–220, 1997.

[54] Constantino Tsallis. Possible generalization of boltzmann-gibbs statistics. *Journal of Statistical Physics*, 52(1):479–487, 1988.

[55] Constantino Tsallis and Daniel A. Stariolo. Generalized simulated annealing. *Physica A: Statistical Mechanics and its Applications*, 233(1):395–406, 1996.

[56] Y. Xiang and X. G. Gong. Efficiency of generalized simulated annealing. *Phys. Rev. E*, 62:4473–4476, Sep 2000.

[57] BELLE II collaboration. Measurement of the integrated luminosity of the Phase 2 data of the Belle II experiment. *Chinese Physics C*, 44(2):021001, feb 2020.

[58] Swagato Banerjee, Bolek Pietrzyk, J. Michael Roney, and Zbigniew Was. Tau and muon pair production cross sections in electron-positron annihilations at $\sqrt{s} = 10.58$ GeV. *Phys. Rev. D*, 77:054012, Mar 2008.

[59] Marc Paterno. Calculating efficiencies and their uncertainties. https://lss.fnal.gov/archive/test-tm/2000/fermilab-tm-2286-cd.pdf, 12 2004.

[60] E. De La Cruz-Burelo, A. De Yta-Hernandez, and M. Hernandez-Villanueva. New method for beyond the standard model invisible particle searches in tau lepton decays. *Phys. Rev. D*, 102:115001, Dec 2020.

Appendix A

Efficiency and purity definitions

In order to estimate the reconstruction efficiency and purity of $ee \to \tau(\to 3\,\text{prong})\tau(\to e)$, where the $\tau \to e$ includes both, the SM $\tau \to e\nu\nu$ and NP $\tau \to e\alpha$ processes, the 3- and 1-prongs in Monte Carlo are defined using basf2 built-in variables and functions:

τ 3-prong:

- tau(Plus/Minus)MCProng = 3 ;
- all the tracks have a τ ancestor: hasAncestor(15) > 0;

$\tau \to e\nu\nu$ (SM):

- mcPDG = $\pm$ 11 ;
- the track has a τ parent: hasAncestor(15) = 1 ;
- tau(Plus/Minus)MCMode = 1;

$\tau \to e\alpha$ (NP):

- mcPDG = $\pm$ 11 ;
- genMotherPDG = $\pm$ 15;

Then the "signal" event should satisfy the combined requirements on the 3-prong and either SM or NP 1-prong side.

The efficiencies ϵ^{SM} and ϵ^{NP} of the SM and NP samples, respectively, are given by:

$$\epsilon^{SM} = \frac{N_{rec}^{SM}}{N_{gen}^{SM}}, \qquad \epsilon^{NP} = \frac{N_{rec}^{NP}}{N_{gen}^{NP}} \tag{A.1}$$

where $N_{rec}^{SM(NP)}$ is the number of events after certain selection requirements. N_{gen}^{SM} and N_{gen}^{NP} are the number of generated SM and NP events,

$$N_{gen}^{SM} = 2 \times N_{gen}^{total} \times \mathcal{B}(\tau \to hhh \geq 0 \text{ neutrals}) \times \mathcal{B}(\tau \to \text{e} \geq 0 \text{ neutrals}) , \tag{A.2}$$

$$N_{gen}^{NP} = N_{gen}^{total} \times \mathcal{B}(\tau \to hhh \geq 0 \text{ neutrals}) \times 1 , \tag{A.3}$$

with the N_{gen}^{total} being the number of total generated events, $\mathcal{B}(\tau \to hhh \geq 0 \text{ neutrals}) = 15.2\%$ and $\mathcal{B}(\tau \to e \geq 0 \text{ neutrals}) = 17.82\%$.

The purity of the sample, SM or NP, is defined as,

$$\mathcal{P} = \frac{N_{signal}}{N_{rec}} , \tag{A.4}$$

where N^{signal} is the true number of signal events defined at the beginning of this section, while N_{rec} is the sum of events passing the selection criteria in signal plus background samples.

Appendix B

Trigger bits and trigger efficiency

The description of the triggers (Level 1 logic) used in this work is the following :

- **lml0** : ≥ 3 clusters with at least one having $E^* > 300$ MeV, $1 < \theta_{ID} < 17$ (corresponding to $12.4° < \theta < 154.7°$, full ECL) and not an ECL Bhabha.
- **lml1** : exactly 1 cluster with $E^* > 2$ GeV and $4 < \theta_{ID} < 14$ ($32.2° < \theta < 124.6°$)
- **lml2** : ≥ 1 cluster with $E^* > 2$ GeV, $\theta_{ID} = 2$, 3, 15, or 16 ($18.5° < \theta < 32.2°$ or $124.6° < \theta < 139.3°$) and not an ECL Bhabha.
- **lml4** : ≥ 1 cluster with $E^* > 2$ GeV, $\theta_{ID} = 1$ or 17 ($12.4° < \theta < 154.7°$) and not an ECL Bhabha.
- **lml6** : exactly 1 cluster with $E^* > 1$ GeV, $4 < \theta_{ID} < 15$ ($32.2° < \theta < 128.7°$, full ECL barrel) and no other cluster with $E >$ 300 MeV anywhere.
- **lml7** : exactly 1 cluster with $E^* > 1$ GeV, $\theta_{ID} = 2$, 3 or 16 ($18.5° < \theta < 31.9°$ or $128.7° < \theta < 139.3°$) and no other cluster with $E >$ 300 MeV anywhere.
- **lml8** : cluster pair with $170° < \Delta\phi < 190°$, both clusters with $E^* > 250$ MeV and no 2 GeV cluster in the event.
- **lml9** : cluster pair with $170° < \Delta\phi < 190°$, one cluster with $E^* < 250$ MeV with the other having $E^* > 250$ MeV, and no 2 GeV cluster in the event.
- **lml10** : cluster pair with $160° < \Delta\phi < 200°$, $160° < \sum\theta < 200°$ and no 2 GeV cluster in the event.
- **lml12** : ≥ 3 clusters with at least one having $E^* > 500$ MeV, $2 < \theta_{ID} < 16$ (corresponding to $18.5° < \theta < 139.3°$, full ECL) and not an ECL Bhabha.

B.0.1 Absolute trigger efficiency

The trigger correction was calculated as a data absolute trigger efficiency, obtained from a data relative trigger efficiency with a Monte Carlo correction described below.

The efficiency of the *lmlX* trigger bits (see Section 4.2.4) used in this analysis is measured **relative** to the orthogonal CDC triggers for data and MC :

$$\epsilon^{rel}_{Data,MC} = \frac{\text{(ffo or fff or fso) and lmlX}}{\text{ffo or fff or fso}} \ , \tag{B.1}$$

where we considered the logical OR of three CDC triggers:

- ffo : ≥ 2 full tracks, track pair with $\Delta\phi > 90^\circ$ and not an ECL Bhabha.
- fff : ≥ 3 full tracks.
- fso : ≥ 1 full tracks, ≥ 1 short tracks, track pair with $\Delta\phi > 90^\circ$ and not an ECL Bhabha.

The $\tau^\pm \to \ell^\pm + \alpha$ analysis uses the "MC13a" production samples as describe in Chapter 4. In this production the trigger simulation (TSIM) is unreliable. However, with the "MC14a" production samples, the trigger efficiency of the *lmlX* trigger bits is well reproduced by TSIM. Therefore, we use the "MC14a" samples to derive a small correction to the **relative** *lmlX* efficiency in data (ϵ^{rel}_{Data}) so as to estimate the **absolute** efficiency in data (ϵ^{abs}_{Data}).

The absolute efficiency in the "MC14a" samples is calculated as:

$$\epsilon^{abs}_{MC} = \frac{\text{lmlX}}{\text{no trigger}} \tag{B.2}$$

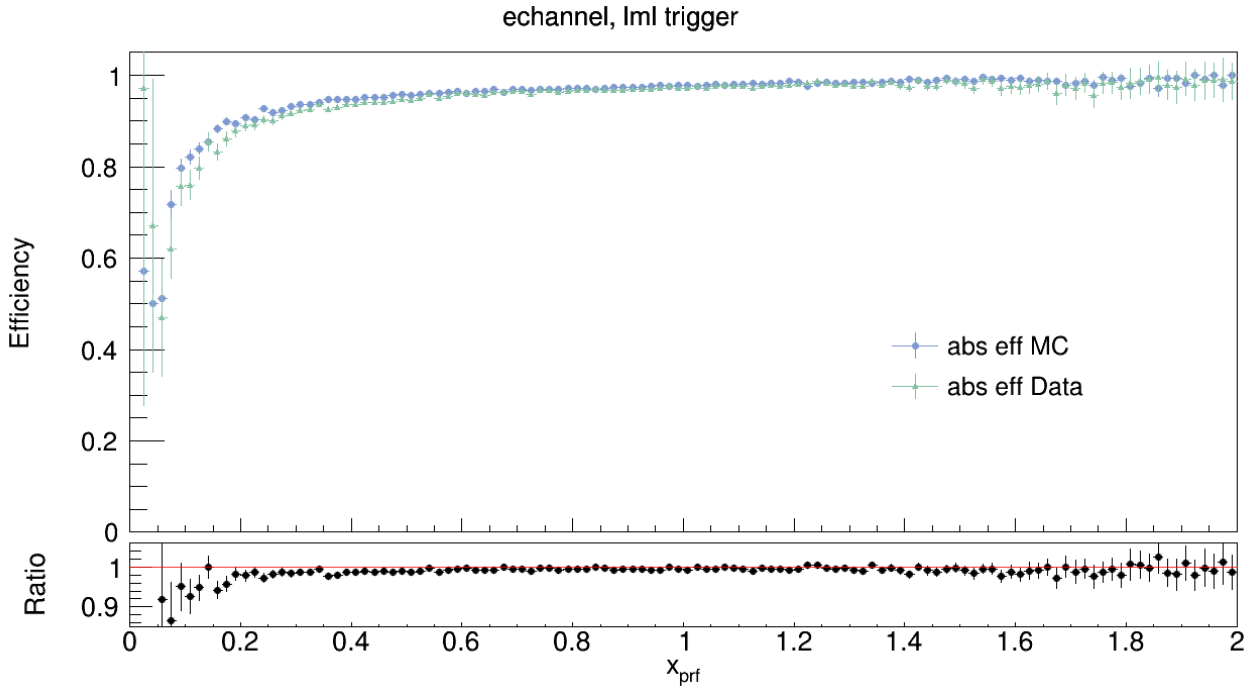
The bin difference between the relative efficiency and absolute efficiency in the "MC14a" samples is given by:

$$\delta_{MC} = \epsilon^{rel}_{MC} - \epsilon^{abs}_{MC} \tag{B.3}$$

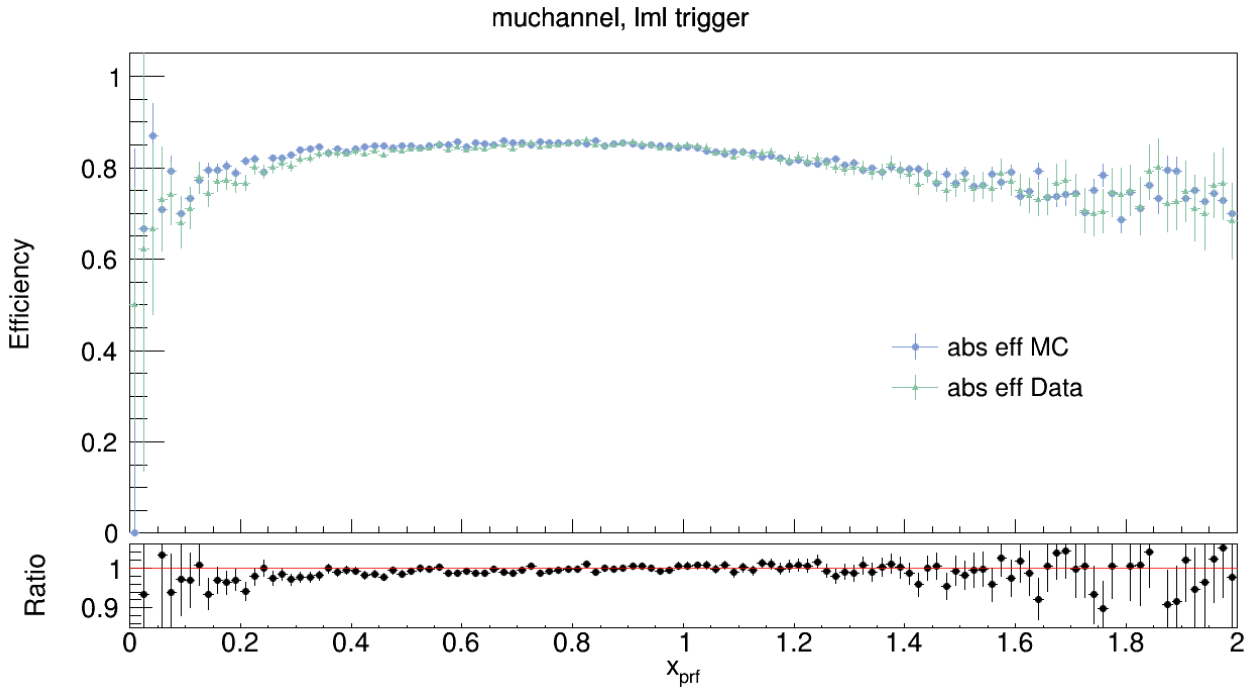
Finally, the absolute trigger efficiency in data is estimated as:

$$\epsilon^{abs}_{Data} = \epsilon^{rel}_{Data} - \delta_{MC} \tag{B.4}$$

Figure B.1 shows the absolute trigger efficiency curves for the muon and electron channels.



(a) e channel.



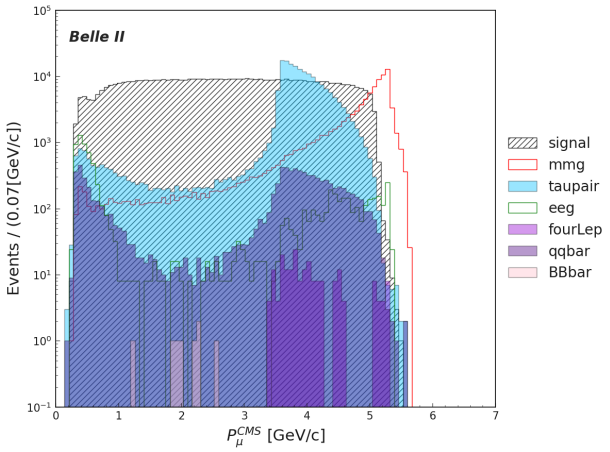
(b) μ channel.

Figure B.1: Absolute trigger efficiency (ϵ^{abs}_{Data}) curves for the (a) electron and (b) muon channels [11].

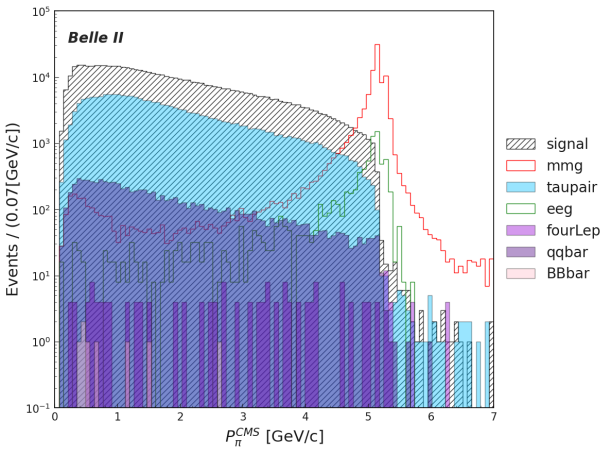
Appendix C

Distributions of the variables used for the optimization of $\tau \to \ell\gamma$

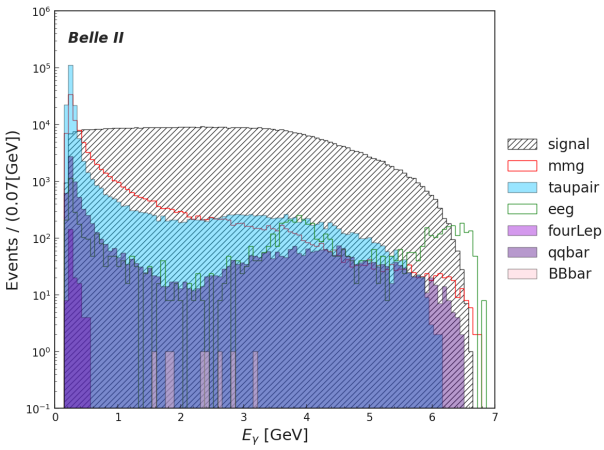
In Figure C.1, C.2, is shown the distribution of these variables for the muon channel, and analogously, in Figure C.4, C.5, and C.6 the distributions for electron channel. These distributions are presented after all the requirements in Section 5.2.



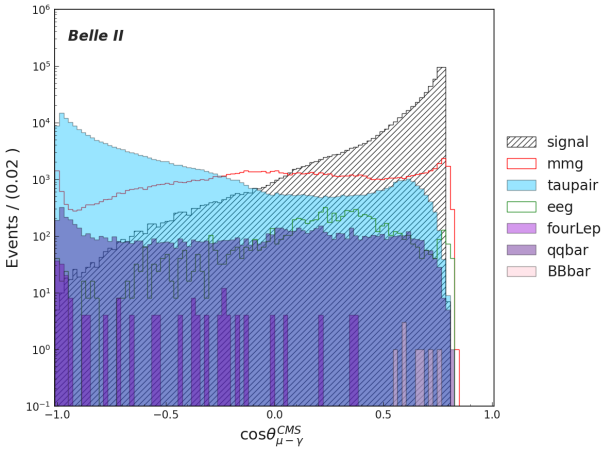
(a) P_ℓ^{CMS}



(b) P_π^{CMS}

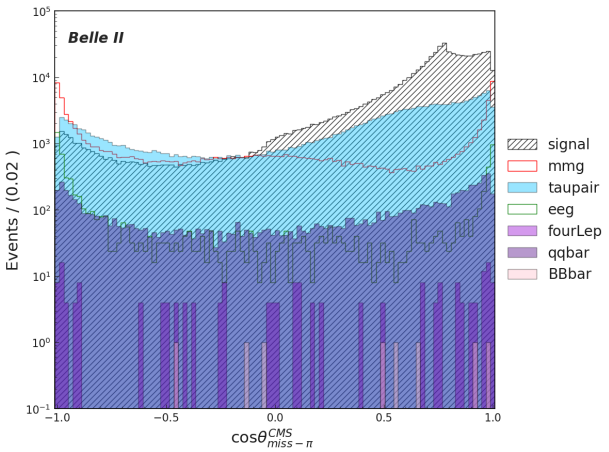


(c) E_γ

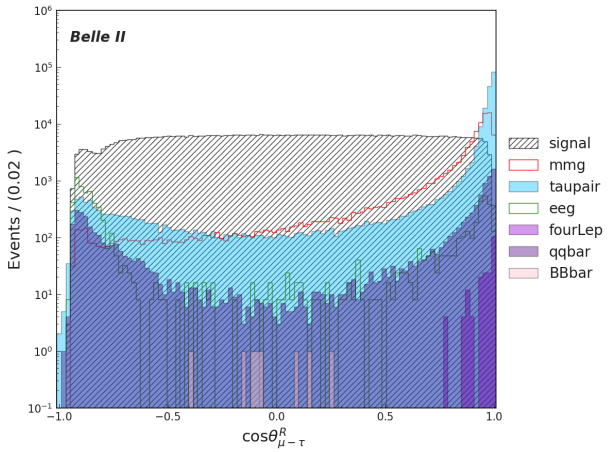


(d) $\cos\theta_{\ell-\gamma}^{CMS}$

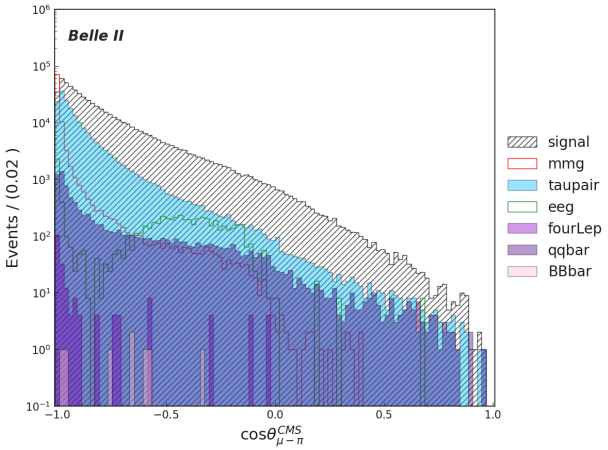
Figure C.1: Distributions for the muon channel $\tau^\pm \to \mu^\pm + \gamma$. Figures (a) and (b) show the event distributions for the momentum of the signal and tag tracks in the CMS frame, respectively. The distribution of the photon energy (signal candidate) is shown in (b). Figure (c) shows the distribution of the cosine of the opening angle between the signal track and the photon candidate in the CMS. This value is expected to be greater than zero for the signal.



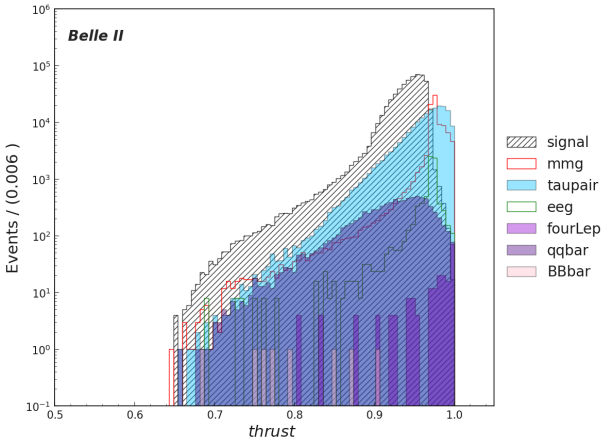
(a) $\cos\theta^{CMS}_{miss-\pi}$



(b) $\cos\theta^{R}_{\ell-\tau}$

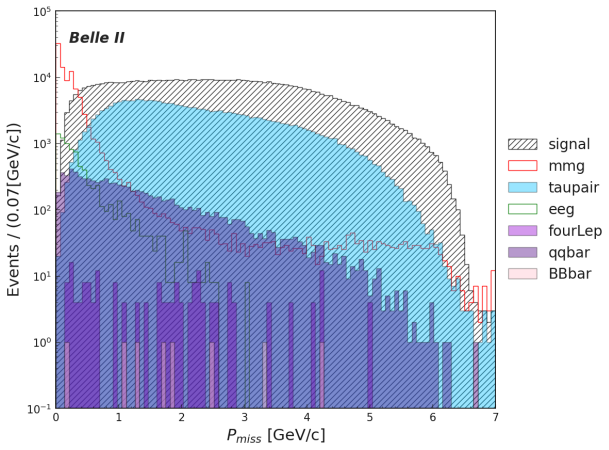


(c) $\cos\theta^{CMS}_{\ell-\pi}$

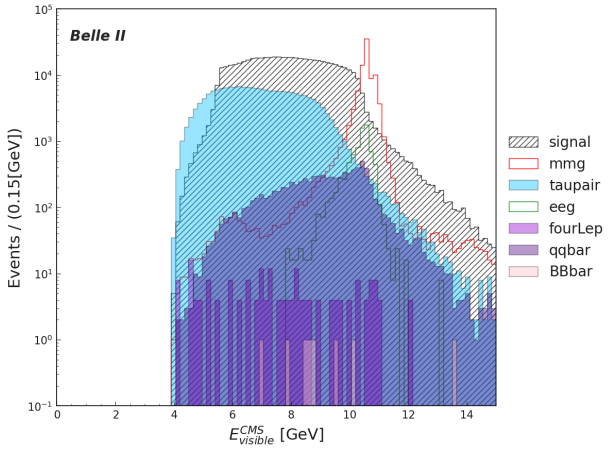


(d) Thrust

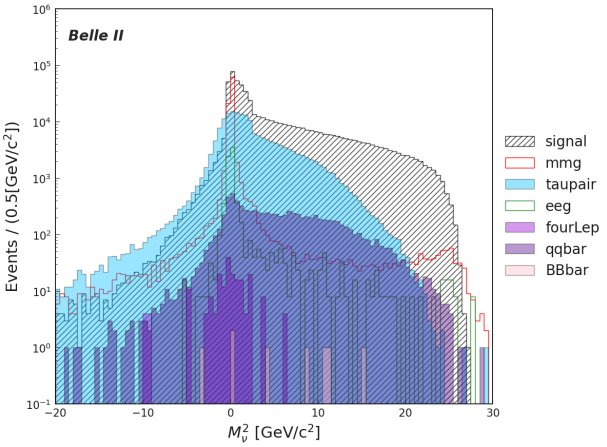
Figure C.2: Distributions for the muon channel $\tau^{\pm} \to \mu^{\pm} + \gamma$. Figure (a) shows the distribution of the cosine of the angle between the missing particles and the track on the tag side in the CMS. This value is expected to be greater than zero because only missing particles are expected on the tag side. Figure (b) is the distribution of the cosine of the opening angle between the signal lepton ℓ and its mother τ in the τ rest frame. Figure (c) is the distribution of the cosine of the opening angle between the signal and tag tracks in the CMS. Figure (d) is the distribution of the thrust value.



(a) P_{miss}

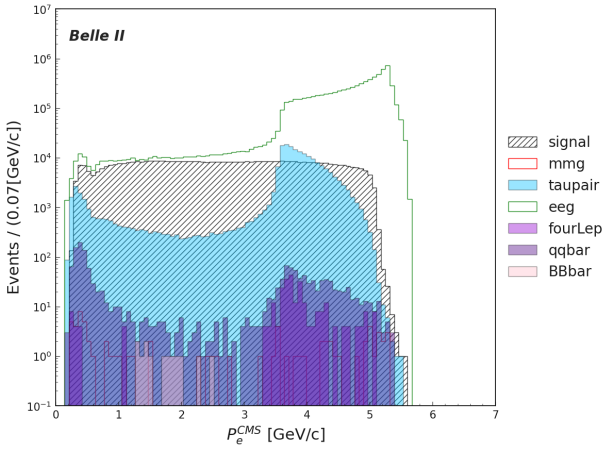


(b) E_{vis}

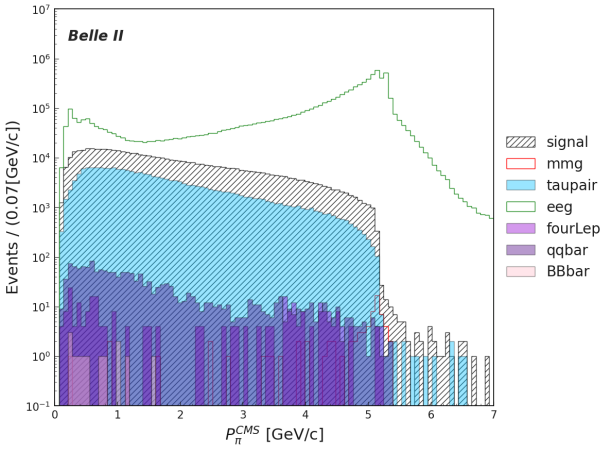


(c) M^2_ν

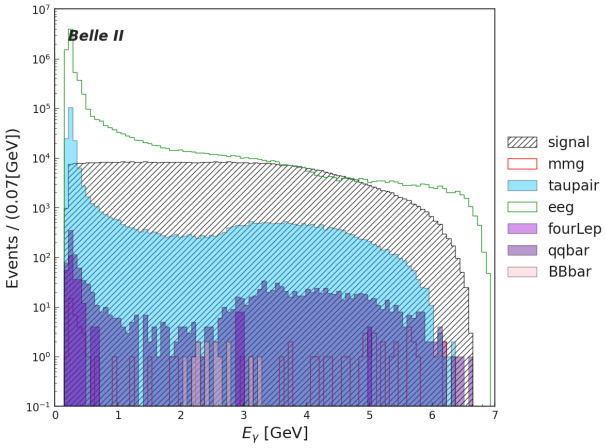
Figure C.3: Distributions for the muon channel $\tau^\pm \to \mu^\pm + \gamma$. Figure (a) shows the distribution for the momentum corresponding to the missing energy. On the other hand, in Figure (b) the distribution of the visible energy is shown. Figure (c) is the distribution of the missing mass squared on the tag side.



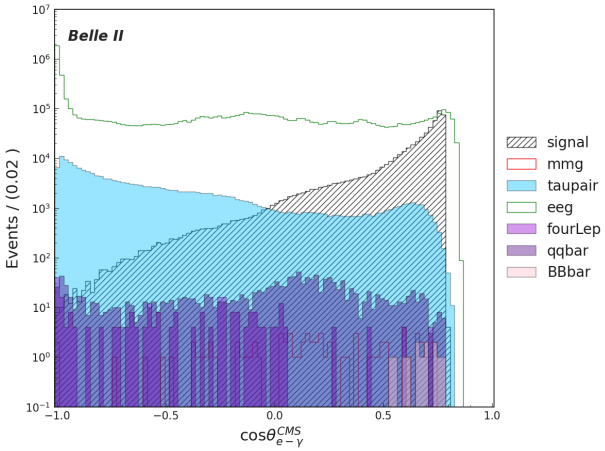
(a) P_ℓ^{CMS}



(b) P_π^{CMS}

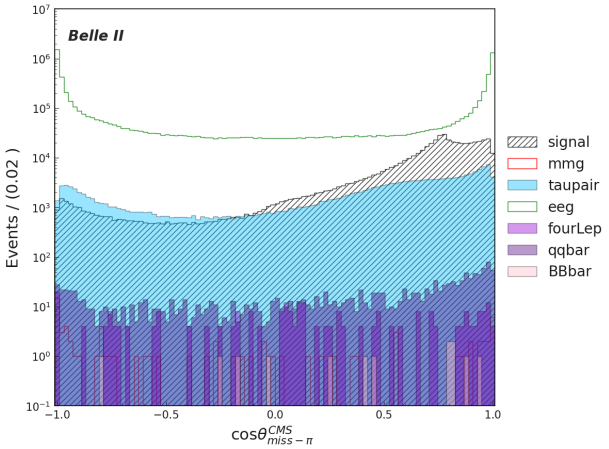


(c) E_γ

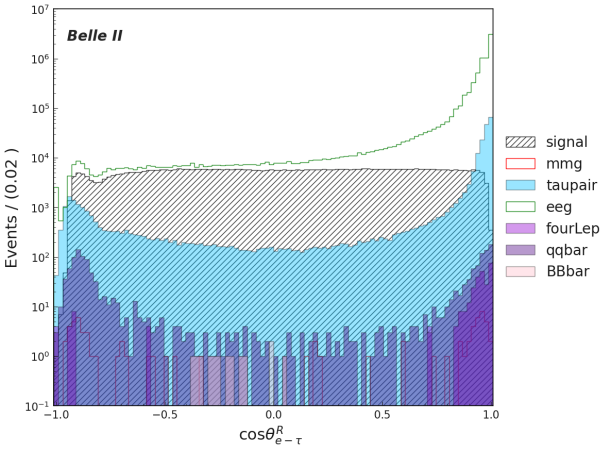


(d) $\cos\theta_{\ell-\gamma}^{CMS}$

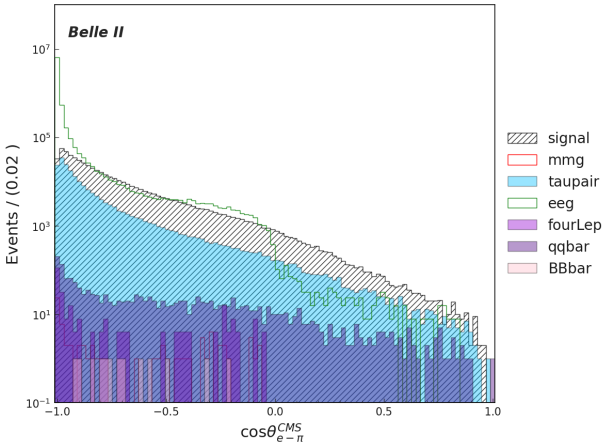
Figure C.4: Distributions for the electron channel $\tau^\pm \to e^\pm + \gamma$. Figures (a) and (b) show the event distribution for the momentum of the signal and tag tracks in the CMS frame, respectively. The distribution of the photon energy (signal candidate) is shown in (b). Figure (c) shows the distribution of the cosine of the opening angle between the signal track and the photon candidate in the CMS. This value is expected to be greater than zero for the signal.



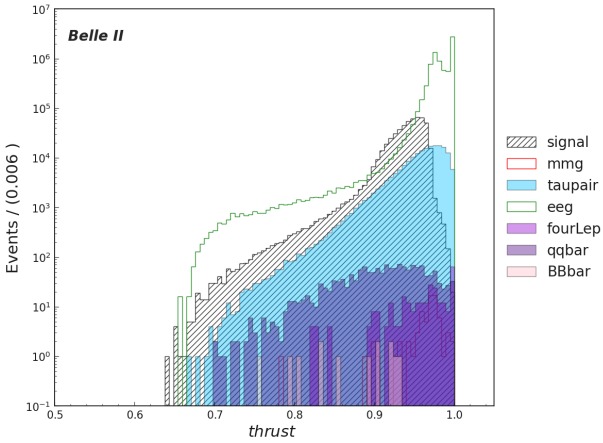
(a) $\cos\theta^{CMS}_{miss-\pi}$



(b) $\cos\theta^{R}_{\ell-\tau}$

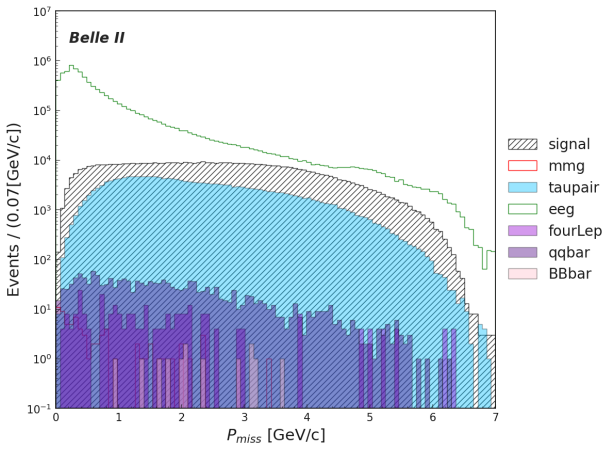


(c) $\cos\theta^{CMS}_{\ell-\pi}$

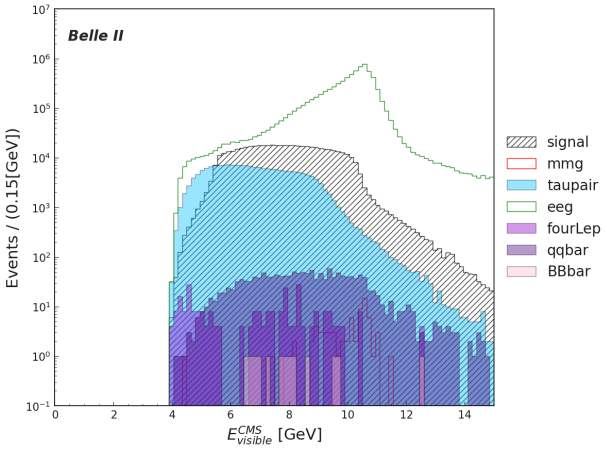


(d) *Thrust*

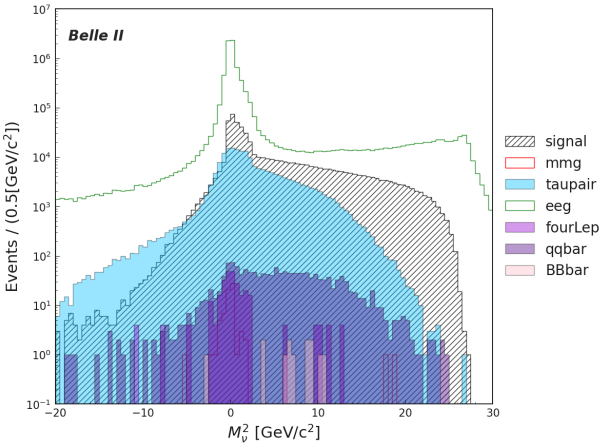
Figure C.5: Distributions for the electron channel $\tau^{\pm} \to e^{\pm} + \gamma$. Figure (a) shows the distribution for the cosine of the angle between the missing particles and the track on the tag side in the CMS. This value is expected to be greater than zero because only missing particles are expected on the tag side. Figure (b) is the distribution of the cosine of the opening angle between the signal lepton ℓ and its mother τ in the τ rest frame. Figure (c) is the distribution of the cosine of the opening angle between the signal and tag tracks in the CMS. Figure (d) is the distribution of the thrust value.



(a) P_{miss}



(b) E_{vis}



(c) M^2_ν

Figure C.6: Distributions for the electron channel $\tau^\pm \to e^\pm + \gamma$. Figure (a) shows the distribution for the momentum corresponding to the missing energy. On the other hand, in Figure (b) the distribution of the visible energy is shown. Figure (c) is the distribution of the missing mass squared on the tag side.

Appendix D

PCA-whitening

We want to perform a linear transformation of our data in order to remove the correlation and to have a new covariance matrix equal to the identity, that is each column has 1-variance.

Let X the data set with m samples, X= $x_1, x_2, ..., x_m$, and with zero mean, then, the correlation matrix C is given by:

$$C = \frac{1}{m} X \cdot X^T \tag{D.1}$$

The covariance matrix C is real and symmetric, and so has an eigen decomposition, i.e., we can rewrite it as:

$$C = E \cdot D \cdot E^T \,, \tag{D.2}$$

where E is the matrix of eigenvectors (as columns) and D is the diagonal matrix with eigenvalues in the diagonal. E^T is also called "rotation" matrix since gives a rotation needed to de-correlated the data.

Then, the PCA-whitening transformation is given by:

$$W_{PCA} = D^{-1/2} \cdot E^T \,, \tag{D.3}$$

Therefore, this transformation de-correlates the data with E^T, then, makes the variances equal to 1 scaling by $D^{-1/2}$.

Let $X^{'}$ the new transformed data, below is the proof that the correlation matrix is the identity under this transformation:

$$\begin{aligned}
C' &= \tfrac{1}{m} X' \cdot X'^T \\
&= \tfrac{1}{m} (W_{PCA} \cdot X) \cdot (W_{PCA} \cdot X)^T \\
&= \tfrac{1}{m} (D^{-1/2} \cdot E^T \cdot X) \cdot (D^{-1/2} \cdot E^T \cdot X)^T \\
&= \tfrac{1}{m} D^{-1/2} \cdot E^T \cdot X \cdot X^T \cdot E \cdot D^{-1/2} \\
&= D^{-1/2} \cdot E^T \cdot C \cdot E \cdot D^{-1/2} \\
&= D^{-1/2} \cdot E^T \cdot E \cdot D \cdot E^T \cdot E \cdot D^{-1/2} \\
&= D^{-1/2} \cdot D \cdot D^{-1/2} \\
&= I
\end{aligned}$$

Using $A \cdot A^T = I$, $(A \cdot B)^T = B^T \cdot A^T$ and $D^T = D$.