$\tau \to \ell\gamma$ at Belle II

Braden Moore

Under the supervision of Dr. Phillip Urquijo

School of Physics

The University of Melbourne

A thesis submitted for the degree of

Master of Science

28 October 2016

Contents

Acknowledgements

Abstract

1 | Lepton Flavour Violation and the Standard Model

Lepton flavour violation (LFV) is an exciting field of research at the frontier of particle physics. Searches for LFV probe a wide variety of new physics (NP) scenarios. **DISCUSS WHAT LFV IS AND WHERE LEPTON FLAVOUR FITS INTO THE SM** Charged LFV of the form $\tau \to \ell\gamma$. Of the tau processes, these modes are predicted to be the most sensitive to NP. We choose to investigate tau LFV rather than, say, muon LFV, for two main reasons. Firstly, the tau processes have predicted branching fractions of $\sim 5 - 6$ orders of magnitude greater than the analogous muon processes, due to the differences in mass [**?**]. The decay $\tau \to \mu\gamma$ has a predicted branching fraction ~ 6 orders of magnitude greater than the analogous $\mu \to e\gamma$.

Secondly, if this NP introduces Higgs-like particles, we would observe the NP more strongly in the tau sector, since taus couple more strongly to Higgs than do muons.

(a)   (b)
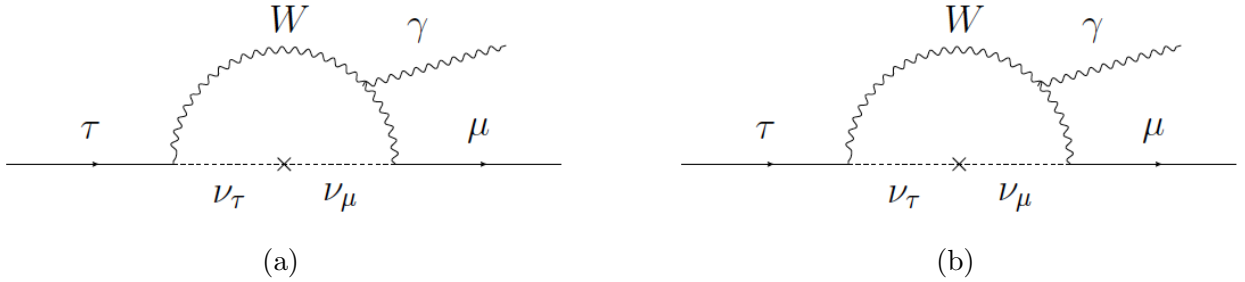
Figure 1.1: Feynman diagrams for (a) $\tau \to \mu\gamma$ and (b) $\tau \to e\gamma$ proceeding through SM processes.

LFV necessarily requires generation mixing between leptons to occur. Though this is prohibited in the Standard Model, the discovery of neutrino oscillations proves that flavour mixing does occur in our universe; that is, flavour is not conserved. We seek to discover whether this LFV can be observed in other areas of flavour physics.

In the Standard Model + massive neutrinos, the only source of LFV is from the operators responsible for neutrino mass. However, the relevant Feynman diagrams (see Figure) are "loop suppressed" and proportional to the GIM factor, given as $\left(\frac{m_\nu}{M_W}\right)^4$; as neutrino mass is very small

($\mathcal{O}(0.3\,\mathrm{eV})$) we expect the LFV effects to be negligible! With these operators the branching ratio for, say, $\tau \to \mu\gamma$ is $\sim 10^{-40}$ [2]. With such little SM background, observation of an LFV process of the type $\tau \to \ell\gamma$ would be an unambiguous signature of NP.

$$\mathcal{B}(\tau \to \ell\gamma) = \frac{3\alpha}{32\pi}|U_{\tau i}^{\star}U_{\mu i}\frac{\Delta_{3i}^2}{m_W^2}|^2 \leq 10^{-53} \sim 10^{-49} \tag{1.1}$$

1.1 Other LFV

In many NP models, LFV is not limited to just $\tau \to \ell\gamma$ decays. There have been searches for other LFV modes, such as $\mu \to e\gamma$, and $\tau \to 3\ell$. Current limits on the branching fractions are given in Figure 1.2 below [?].

LFV process	Present bound	Future sensitivity
$\mu \to e\gamma$	5.7×10^{-13}	$\approx 6 \times 10^{-14}$
$\mu \to 3e$	1.0×10^{-12}	$\approx 10^{-16}$
$\tau \to e\gamma$	3.3×10^{-9}	$\sim 10^{-8} - 10^{-9}$
$\tau \to \mu\gamma$	4.4×10^{-9}	$\sim 10^{-8} - 10^{-9}$
$\tau \to 3e$	2.7×10^{-8}	$\sim 10^{-9} - 10^{-10}$
$\tau \to 3\mu$	2.1×10^{-8}	$\sim 10^{-9} - 10^{-10}$

Table 1.1: Current experimental limits on various LFV processes

Moving into future sensitivities accessible from experiments such as Belle II, we see that the upper limits of branching fractions for $\tau \to \ell\gamma$ decays could be improved by $1-2$ orders of magnitude!

1.2 Hints of LFV beyond the Standard Model

Motivations behind the search for LFV come from both theoretical and experimental results. The primary experimental motivations is the existence of neutrino mixing, though anomalous results such as the $h \to \tau\mu$ excess observed at CMS in 2015 also hint at LFV beyond the Standard Model. On the theoretical side, LFV is predicted in a variety of NP models. In fact, many models which introduce mechanisms to generate neutrino mass also inadvertently allow LFV in other sectors of non-negligible order! We shall discuss these motivations below.

1.2.1 Neutrino mixing

The discovery that flavour mixing can occur in the neutrino sector [?] [?] proves that neutrinos have mass. Both the concept of massive neutrinos, and by extension the mechanisms which generate neutrino mass, are not predicted or explained by the SM. This tells us that the lepton sector is not fully understood.

There are many NP models which introduce mechanisms to give neutrinos mass. These include SUSY, seesaw models, and many others. In introducing these mechanisms, many of these models inadvertently introduce LFV! As a particle example, a Type-II seesaw model posits a scalar triplet of Higgs-like particles [2]. This triplet comprises a doubly-charged Higgs, a singly-charged Higgs, and a neutral Higgs. As show in Figure X below, lepton-flavour violating processes could proceed via leptons coupling to these scalars!
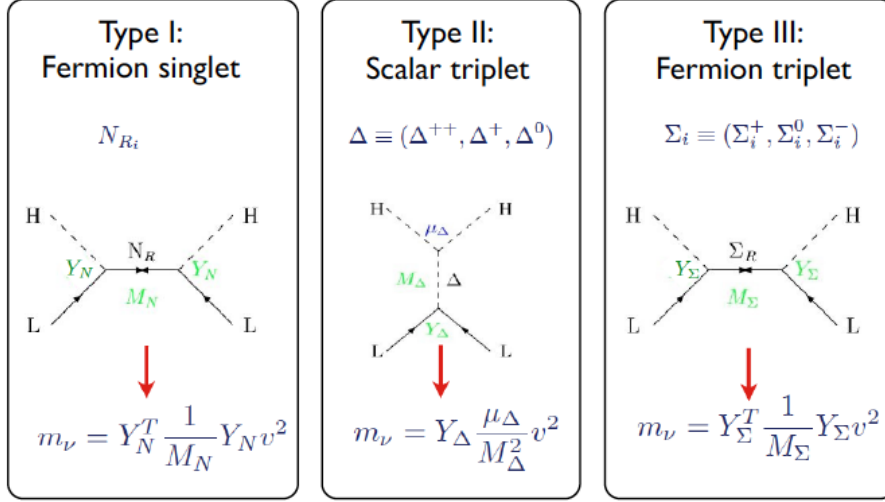
Figure 1.2: New particles introduced in seesaw models (Passemar, 2015)
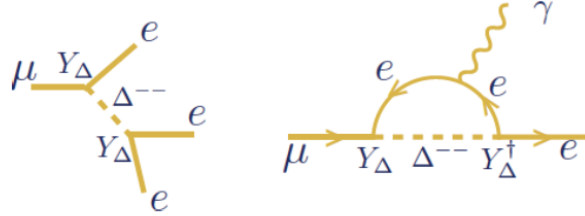
Figure 1.3: Scalars introduced in Type-II seesaw models mediating LFV decays (Passemar, 2015)

1.2.2 $h \to \tau\mu$ excess

Hints of LFV can come in the form of experimental results which are not consistent with the SM. One such "anomaly is the $h \to \tau\mu$ excess. In 2015, CMS found a 2.4σ excess in the branching fraction of $h \to \tau\mu$ [?]. This process is lepton flavour violating, so in the SM its branching fraction is expected to be consistent with zero. However it was determined

$$\mathcal{B}(h \to \tau\mu) = (0.84^{+0.39}_{-0.37})\% \tag{1.2}$$

Also in 2015 was a similar search performed by ATLAS [**?**], in which an excess of 1.2σ was found in the $h \to \tau\mu$ decay.

$$B(h \to \tau\mu) = (0.77 \pm 0.66)\% \tag{1.3}$$

Though this 1.2σ result is less indicative of NP, it still provides hints as to where NP could occur. These results indicate possible new physics in the Higgs sector! Several models, including Two-Higgs Doublet Models (2HDM), introduce new Higgs-like particles; these particles can couple with leptons to allow lepton flavour violating processes [**?**]. In fact, LFV can occur naturally in any model with more than one Higgs doublet.
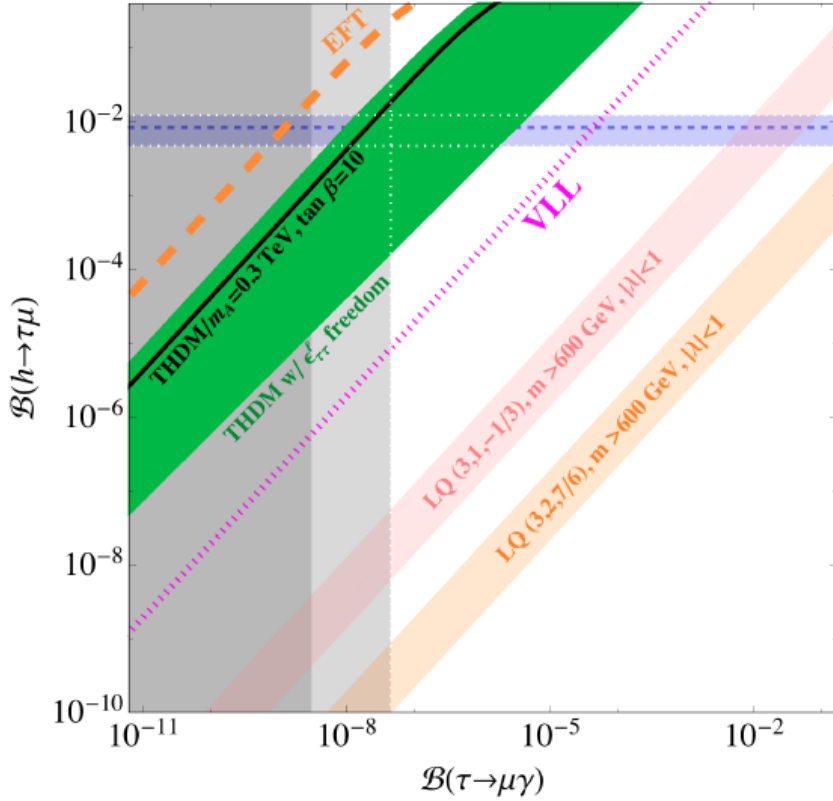
Figure 1.4: Correlation between $\mathcal{B}(h \to \tau\mu)$ and $\mathcal{B}(\tau \to \mu\gamma)$ in various NP scenarios (Dorsner et al., 2015)

The present experimental result for $\mathcal{B}(h \to \tau\mu)$ is shown in horizontal blue band; current and future projections for $\mathcal{B}(\tau \to \mu\gamma)$ experimental sensitivity are represented with vertical light and dark gray bands. We note that certain 2HDM models predict a branching fraction for $\mathcal{B}(\tau \to \mu\gamma)$, consistent with the CMS results, at sensitivities which could be observed by Belle II. It is important to note that the Higgs sector could contribute to LFV in NP scenarios, and that both theory and experimental limits on other LFV processes such as $h \to \tau\mu$ all interweave

with limits on $\tau \to \ell\gamma$ branching fractions to provide information on NP, even just through reducing the available phase space for certain models [**?**]. '

Figure 1: Diagrams contributing to the flavor violating decay $\tau \to \mu\gamma$, mediated by a Higgs boson with flavor violating Yukawa couplings.

Figure 1.5: Diagrams contributing to $\tau \to \mu\gamma$ decay, mediated by a Higgs boson with flavour violating Yukawa coupling (Harnick et al., 2012)

As seen in Figure above, new Higgs particles can mediate LFV processes and allow for measurable amounts of LFV beyond the Standard Model [**?**].
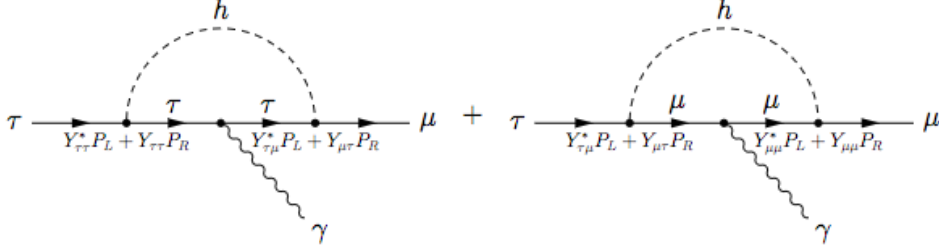
1.2.3 Models predicting $\tau \to \ell\gamma$

As mentioned previously, LFV in the τ sector is introduced in many NP scenarios as a consequence of generating neutrino mass (and hence facilitating neutrino mixing). Branching fractions of the modes $\tau \to \ell\gamma$ are highly calculable - there is little theoretical uncertainty.

model	$Br(\tau \to \mu\gamma)$
mSUGRA + seesaw	10^{-7}
SUSY + SO(10)	10^{-8}
SM + seesaw	10^{-9}
Non-Universal Z'	10^{-9}
SUSY + Higgs	10^{-10}

Figure 1.6: Upper limits of branching fractions from $\tau \to \mu\gamma$, predicted by models of new physics beyond the SM (various sources) [1]

Figure above lists a few NP models with their predictions of $\mathcal{B}(\tau \to \mu\gamma)$. We see that the phase space of some of these models has already been ruled out with our current experimental limits on LFV branching fractions.

1.3   Past searches for $\tau \to \ell\gamma$

The most recent searches for $\tau \to \ell\gamma$ were undertaken at Belle (2007) and Babar (2010), for both $\ell = \mu, e$ modes. These detectors are e^+e^- colliders a signal of the form $e^+e^- \to \tau^+\tau^-$, with one tau (signal-side) decaying $\tau \to \ell\gamma$ and the other tau (tag-side) decaying generically, with the requirement that the tag-side track is not ℓ.

1.3.1   Belle searches

(a)   (b)
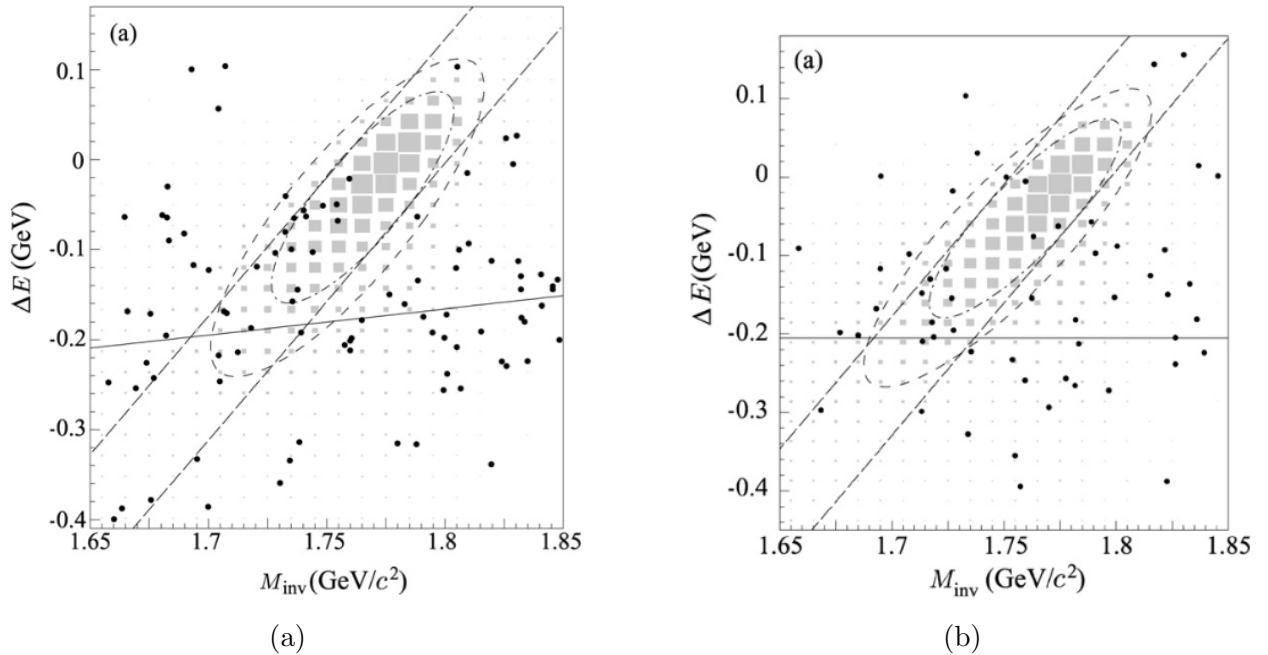
Figure 1.7: M_{inv}-ΔE distribution in the search for (a) $\tau \to \mu\gamma$ and (b) $\tau \to e\gamma$ at Belle over a $515\,\mathrm{fb}^{-1}$ data set. M_{inv} is the reconstructed mass of the τ from these processes, ΔE is the energy difference between this reconstructed τ and half center-of-mass frame beam energy. Dots are the data and shaded boxes indicate the signal MC. The dashed ellipse shows the 3σ region and the dot-dashed ellipse is the 2σ signal region.

The Belle detector records events from an asymmetric e^+e^- collider with electron (positron) energy of $8\,\mathrm{GeV}$ ($3.5\,\mathrm{GeV}$). A detailed discussion of the detector can be found at Ref. [?]. In 2007, the Belle Collaboration performed a search over $535\,\mathrm{fb}^{-1}$ of e^+e^- data and set constraints [?] on $\tau \to \ell\gamma$ branching fractions as

$$\mathcal{B}(\tau \to \mu\gamma) < 4.5 \times 10^{-8},$$
$$\mathcal{B}(\tau \to e\gamma) < 1.2 \times 10^{-7}.$$
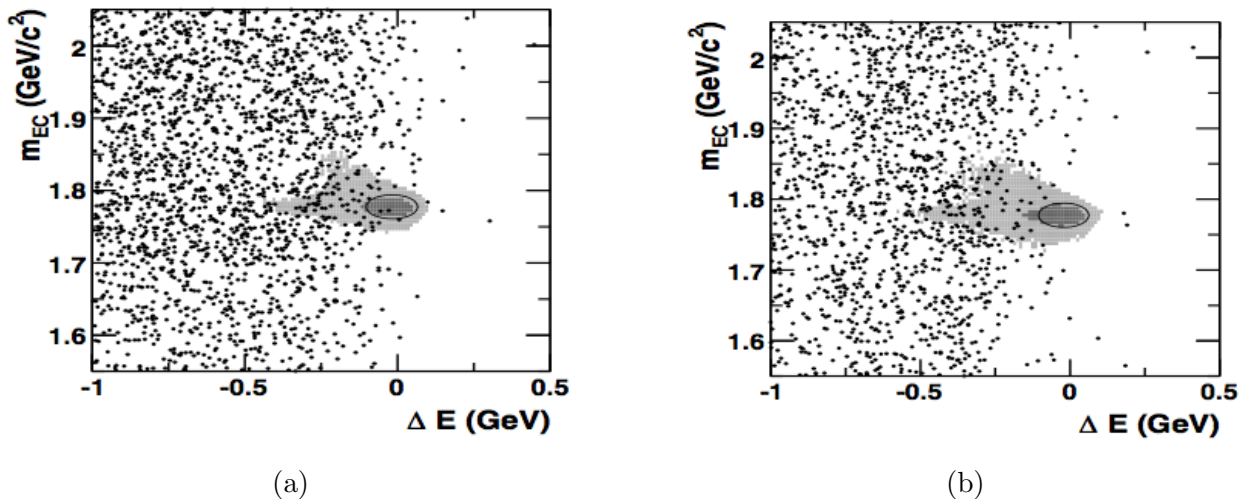
Figure 1.8: M_{bc}-ΔE distribution in the search for (a) $\tau \to \mu\gamma$ and (b) $\tau \to e\gamma$ at Babar over a $515\,\mathrm{fb}^{-1}$ data set. M_{bc} is the beam-constrained mass of the τ from these processes, ΔE is the energy difference between this reconstructed τ and half center-of-mass frame beam energy. Data are shown as dots and contours containing 90% (50%) of signal MC events are shown as light- (dark-) shaded regions.

1.3.2 Babar searches

Similar to Belle, the Babar detector records events from an asymmetric e^+e^- collider, with electron (positron) energy of $9\,\mathrm{GeV}$ ($3.1\,\mathrm{GeV}$). A detailed discussion of the detector can be found at Ref. [?]. The most recent search for $\tau \to \ell\gamma$ was performed in 2010 by the Babar Collaboration [?], over a $515.5\,\mathrm{fb}^{-1}$ dataset, setting constraints on $\tau \to \ell\gamma$ branching fractions as

$$\mathcal{B}(\tau \to \mu\gamma) < 4.4 \times 10^{-8},$$
$$\mathcal{B}(\tau \to e\gamma) < 3.3 \times 10^{-8}.$$

1.4 The role of this analysis (rephrase plz)

Collection of data at Belle II is scheduled to begin in 2017. A physics program has been designed for the experiment, detailing specific decays to investigate and areas of particle physics to probe. It is important before data collection begins to perform sensitivity studies on these decay processes, to find areas in which the analysis performs better or worse than expected. In some cases we may find that background rejection is greater than previously modelled due to, for example, good performance in particle identification by the detector. It is also important if areas for improvement are found, such as beam background degrading analysis performance. Finding areas for improvement during the development stage informs future development, in-

cluding tracking and reconstruction software and hardware-based beam background mitigation techniques, and leads to more optimal conditions for physics events to be investigated.

This analysis is the first τ physics study for Belle II. Some key points studied are muon and electron track reconstruction and the effects of beam background on XXXXX (???). Information on these areas of event analysis is not only useful for $\tau \to \ell\gamma$ searches at Belle II — the physics program for the experiment includes a range of τ decay processes (see Table XXX) of which many require muon and electron track reconstruction. Low-multiplicity events require XXXXXX, so XXXXXX.

LFV process	Present bound	Future sensitivity
$\tau \to e\gamma$	3.3×10^{-9}	$\sim 10^{-8} - 10^{-9}$
$\tau \to \mu\gamma$	4.4×10^{-9}	$\sim 10^{-8} - 10^{-9}$
$\tau^{\pm} \to e^{\pm}e^{\mp}e^{\pm}$	2.7×10^{-8}	$\sim 10^{-9} - 10^{-10}$
$\tau^{\pm} \to \mu^{\pm}\mu^{\mp}\mu^{\pm}$	2.1×10^{-8}	$\sim 10^{-9} - 10^{-10}$

Table 1.2: Current experimental limits on various LFV processes

Results from this analysis will be published in the Belle II physics book, which is currently being compiled by members across all working groups in the collaboration. The book aims to (????); this study will provide the basis for future τ LFV studies at Belle II. This study also validates the projected increase in sensitivity for $\tau \to \ell\gamma$ branching fractions from Belle to Belle II.

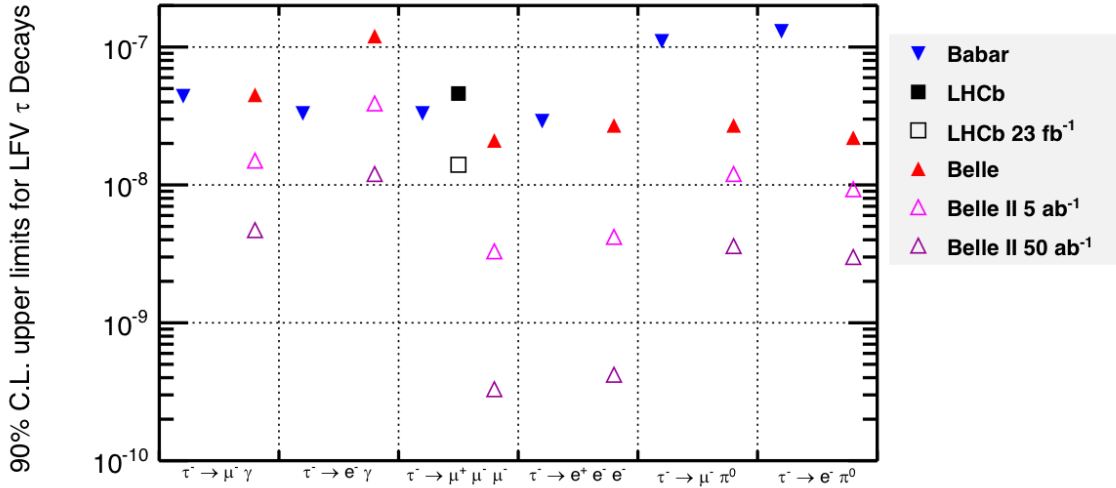Figure 1.9: Current and future sensitivities on τ LFV branching fractions (*Urquijo*, 2016). **NOTE: remove the unrelated LFVs**

2 | The Belle and Belle II detectors

Located in Tsukuba, Japan, the KEKB particle accelerator was used for the Belle experiment from 1999 to 2010 and collided high energy electrons and positrons of $8\,\text{GeV}$ and $3.5\,\text{GeV}$ respectively. This experiment made important discoveries on the flavour structure of elementary particles, notably in the quark section where CP symmetry violation was discovered. Efforts by the Belle Collaboration culminated in the 2008 Nobel Prize awarded to Makoto Kobayashi and Toshihide Maskawa for their theory of CP violation. Over its lifetime, the Belle detector collected a total time-integrated luminosity of $1\,\text{ab}^{-1}$, corresponding to ???? tau-pair events. An upgrade to the KEKB collider known as SuperKEKB is expected to be complete by 2014 (**LOL**) as part of the Belle II experiement.
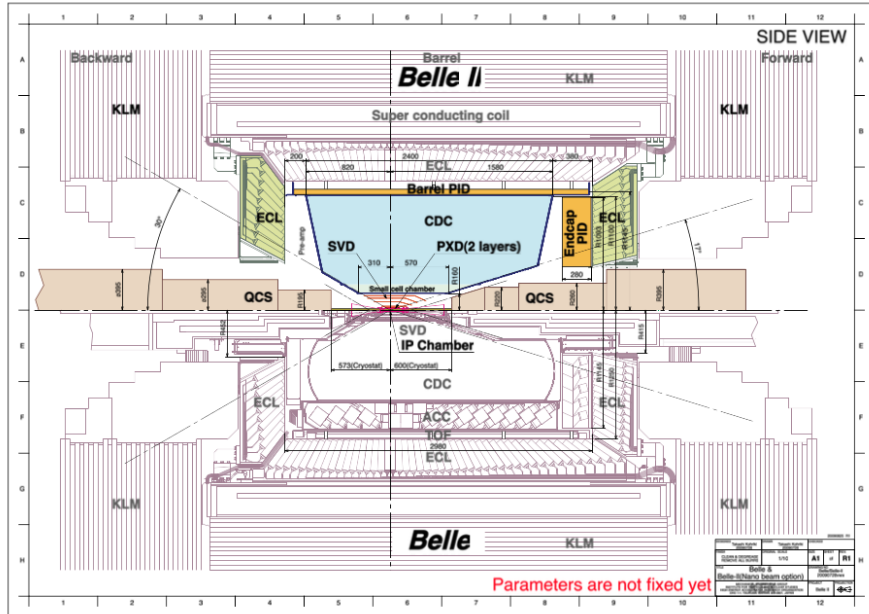
Figure 2.1: Super-KEKB detector design, side-view

Over the course of the experiment, it is projected that Belle II will collect a total time-integrated luminosity of $50\,\text{ab}^{-1}$. This is a much larger data sample than the $1\,\text{ab}^{-1}$ collected at Belle. Electron-positron beam energies differ from those used at Belle, with high-energy-ring (HER) electron beam energy of $7\,\text{GeV}$ and low-energy-ring (LER) positron beam energy of $4\,\text{GeV}$. Design luminosity in the detector is increased by 40 over Belle; this increases background

hit rate (???) by a factor of 20, while physics event rate is expected to be 50 times higher than at Belle.

Key components of the detector are the vertex detector (VXD), the central drift chamber (CDC), the electromagnetic calorimeter (ECL), the time-of-flight/Cerenkov aerogel chamber (TOP), and the K-long/muon detector (KLM). Many of these will be upgraded in Belle II to provide better performance; in some cases the full reconstruction software is rewritten even when the hardware is not improved.

The old silicon strip detector immediately outside the beam pipe will be replaced with a twolayer pixel detector. The endcap and inner layers of the K L and detector are to be replaced with scintillators

2.1 Vertex Detector

Determing the decay vertices of particles — that is, the point inside the detector which the decay occurs — was performed at Belle by the Silicon Vertex Detector (SVD2), comprising four layers of double-sided silicon detectors (DSSDs). By the end of its lifetime SVD2 the innermost layer had an average occupancy of 10%, defined as the fraction of channels hit in each triggered event. This level of occupancy leads to worsened track resolution and reconstruction "due to merging of clusters with lead to incorrect track association".

With beam luminosity at Belle II projected to increase by factor 20 over Belle, the greater rate of events due to beam-related backgrounds would worsen SVD2 occupancy and degrade detector performance. As such the vertexing detector received a redesign moving to Belle II, with a focus on reducing occupancy. This is achieved by the introduction of a two-layer Pixel Detector (PXD) replacing the the DSSDs closest to beam interaction point, and a four-layer Silicon Vertex Detector (SVD) beyond the PXD.

Vertexing is important BECAUSE XXXXX.

2.1.1 Pixel Detector

The PXD at Belle II has a far larger number of channels than DSSDs - 3.072 million pixels in the inner layer, and 4.608 million pixels in the outer layer. This affords a much smaller occupancy even with higher event rate. Pixel layers are located at $14\,\text{mm}$ and $22\,\text{mm}$ from the beampipe.

2.1.2 Silicon Vertex Detector

"The remaining silicon strip detector is to be extended to have a larger radius than in Belle."

Four ladder layers: 38, 80, 115, 140mm. Excellent timing resolution (~ 2 - 3ns) complements PXD

The main characteristics of the Belle II SVD are the following:

- It covers the full Belle II angular acceptance of $17^{\deg} < \theta < 150^{\deg}$. The inner radius is 38 mm and outer radius is 140 mm; these are determined by the radii of the PXD and CDC.

- It improves the quality of the reconstruction of charged tracks compared to Belles SVD2.

- It provides data to extrapolate the tracks reconstructed in the CDC to the PXD with high efficiency; this is the strategy we employ to determine a vertex.

The SVD is used together with the PXD and CDC to measure decay vertices, as well as to extrapolate charged particle tracks.

2.2   Central Drift Chamber

"The central drift chamber, the primary tracking device, will have a larger volume and smaller cell sizes than in Belle."

Smaller drift cells than Belle; extends to a large radius.

$\sim 51\,500$ sense wires inside 1.5T magnetic field. Key roles: 1. Reconstruct charged tracks with precision momentum measurements. 2. Particle identification using measurements of dE/dx 3. Trigger for charged particles.

Tracking of charged particle trajectory is accomplished through combination of several detector components. The CDC is the primary detector for this task.

2.3   Electromagnetic Calorimeter

"The endcap scintillator crystals (CsI(T1)) in the electromagnetic calorimeter will be replaced with faster, more radiation tolerant pure CsI crystals, and new electronics will be used."

Possible replacement of endcap scintillator crystral with a faster and radiation tolerant version.

New electronics of ECL considerably reduce noise pile-up. ' I can talk about the importance of cluster timing

`https://docs.belle2.org/record/344/files/BELLE2-NOTE-TE-2016-006.pdf`

2.4   Time-of-flight/Cerenkov aerogel chamber

"Particle identification is to be performed by entirely new devices using Cerenkov imaging with faster readouts than in Belle."

TOP: Particle identification in barrel region. Sixteen modules, each containing: Two 2.7m long quartz bars. A spherical mirror. An expansion prism. An array of photodetectors.

Aerogel: Particle identification in forward endcap. Components: Aerogel radiator -¿ produces Cherenkov photons. Expansion volume. 2D array of photon detectors (420 Hybrid Avalanche Photo Detectors). Readout system for photon detectors.

Focusing constructed to seperate K and pi photons across most of their momentum range. USEFUL FOR K - PI SEPARATION.

2.5 K_L/muon detector

"The end-cap and inner layers of the K L and detector are to be replaced with scintillators."

Used in muon identification. Can discriminate charged pions from muons by penetration depth.

2.6   Particle identification (PID)

These PID values are generated through combination of many components of the detector, and provide the probability of a track being a particular particle. Some tracks may be double counted or misreconstructed. To account for this, several tighter PID cuts are implemented later.

2.7   Beam backgrounds

SEE TECHNICAL DESIGN REPORT

Beam background is an important issue in B-factories, and especially so for the upgraded energies of SuperKEKB. Key sources of this beam background are synchrotron radiation (SR), beam-gas scattering, and Touschek radiation.

Scattered particles and radiation photons generated by these processes collide with the beam pipe and generate showers of photons, leptons and hadrons. Additionally, other particles can be generated at the interaction point of the collision between opposite beams; these are mostly electron-positron pairs. These particles then interact with the detector, and are called beam background. This background can make searches for interesting physics events difficult, due to the large number of clusters and tracks introduced.
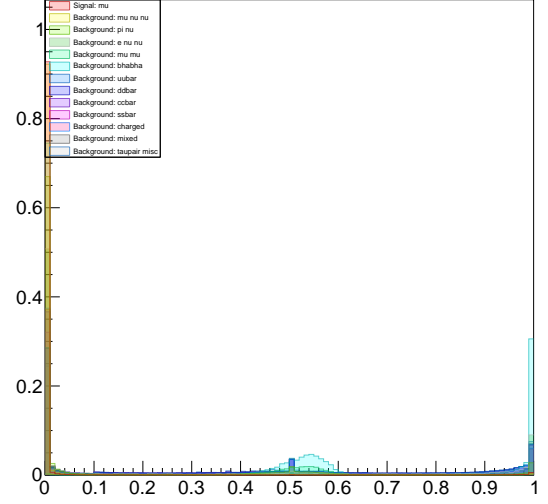
2.7.1   Synchrotron radiation

As beam particles are bent by magnets while travelling through the accelerator, they emit synchrotron radiation.
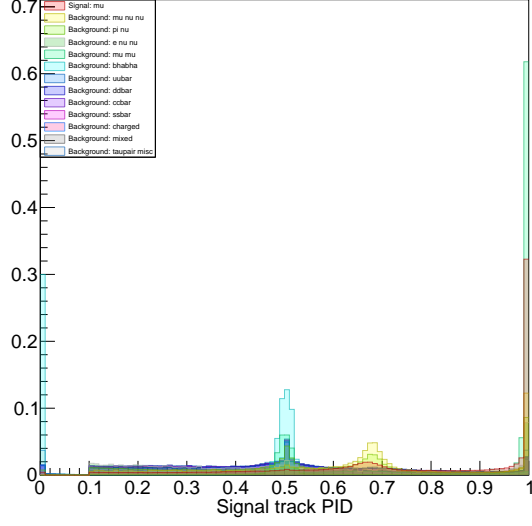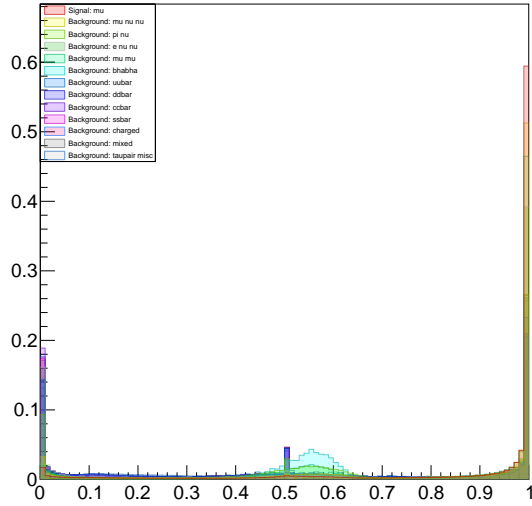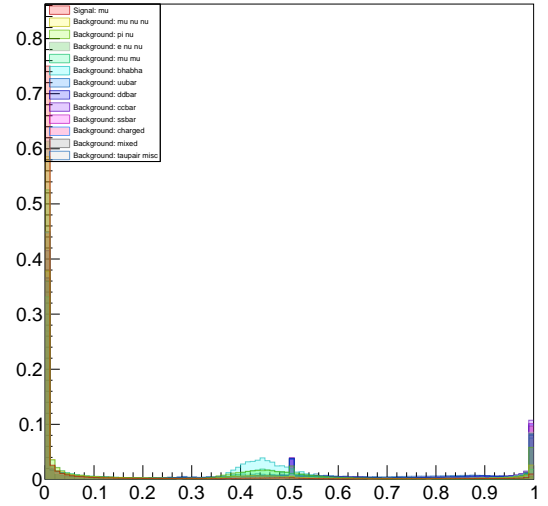
Figure 2.2: Particle likelihood probabilities for signal track.
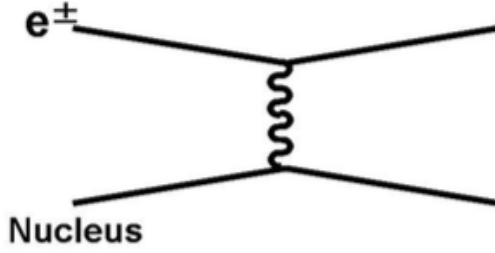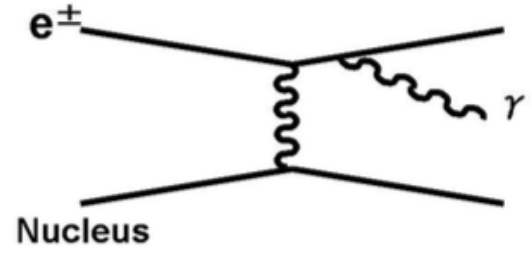
2.7.2 Beam-gas scattering

The region inside the beam pipe is not a perfect vacuum; the designed gas pressure for SuperKEKB is 10^{-7} Pa. Most of this gas is composed of neutral H_2 and CO_2. Beam particles can collide with the gas molecules, resulting in elastic scattering where energy is unchanged but direction is changed (Coulomb scattering), or inelastic scattering whereby a photon is emitted from the scattered particle (bremsstrahlung).

(a) Coulomb scattering   (b) Bremsstrahlung

2.7.2.1 Coulomb scattering

Beam electrons and positrons can elastically scatter off beam gas particles, changing direction such that the scattered particle does not reach the interaction point.

2.7.2.2 Bremsstrahlung

Beam-gas scattering can also occur as bremsstrahlung, where electrons (or positrons) recoil off gas nuclei and emit photons (see figure XXXX b) above). The photon carries away for fraction of the scattered particle's energy.

2.7.3 Touschek scattering

Particle beams do not exist as continuous "lines" of electrons and positrons marching through the accelerator in single file. Instead, we have tightly packed beam bunches, containing $10^{10} - 10^{11}$ particles each. In Super-KEKB the bunches have a designed cross-section perpendicular to the beam trajectory of XXXXX, up from XXXXX in KEKB (LENGTH??). These bunches are aligned almost perfectly parallel so that beam particles do not collide with the beam pipe during their many cycles around the accelerator.

Within these bunches the particles oscillate in a direction perpendicular to the beam trajectory, so that in addition to interacting with beam-gas, particles within a beam bunch collide with each other resulting in an transfer of energy and momentum. Trajectories of these particles may be changed by this interaction, so that they do not reach the interaction point.
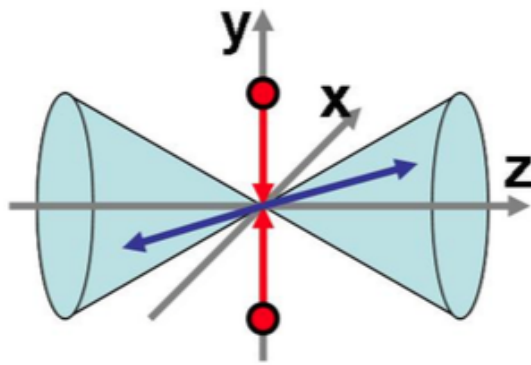
Figure 2.4: An illustration of Touschek scattering in the bunch frame

3 | Monte Carlo production and background events

Electron-positron collisions at Belle II will produce a wide range of different physics events. Within these, we hope to find the signal modes $\tau \to \mu\gamma$ and $\tau \to e\gamma$. However, a gallimaufry of non-signal mode events are also produced; these events which are referred to as background (or sometimes physics background, distinct from beam-background) must be suppressed in order to properly investigate signal processes. We discriminate signal from background through applying selection criteria such as energy, momentum, angular relations and event shape variables. This criteria is optimised by examining Monte Carlo (MC) simulated signal and background events. MC events are produced using physics event generators and the response of the detector is simulated.
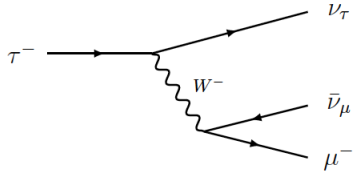
3.1 Physics backgrounds

It is unfeasible to study MC in-depth for every possible physics background; we focus on high event rate processes and events with final states similar to our signal. In this analysis, we investigate all generically decaying tau-pair processes, mu-pair events ($e^+e^- \to \mu^+\mu^-(\gamma)$), Bhabha scattering ($e^+e^- \to e^+e^-(\gamma)$), $e^+e^- \to q\bar{q}$ events (where $q = u, d, c, s$) usually referred to as continuum background, and generic B$\bar{\text{B}}$ processes ($e^+e^- \to B^+B^-$ or $e^+e^- \to B^0\bar{B}^0$). Of the tau-pair processes, we isolate the modes $\tau \to \mu\nu\nu$, $\tau \to e\nu\nu$, and $\tau \to \pi\nu$ (for charged pions $\pi^\pm$) for investigation, as these are expected to be dominant backgrounds [?]. Feynman diagrams of the backgrounds are shown in Figure 3.1a - 3.1i. Note that $e^+e^- \to B\bar{B}$ proceeds though production of the meson $\Upsilon(4S)$.

Fiducial cross-sections for these backgrounds are given in Table 3.1. These values are for collisions at center-of-mass beam energies of $10.573\,\text{GeV}$ and with Belle II detector geometry.

3.2 Event generation

All MC was generated in the Belle II Analysis Framework (basf2) with geometry and energies from SuperKEKB. HOW DOES GENERATION HAPPEN?

16
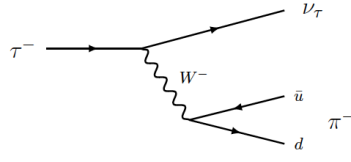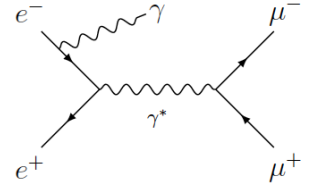
(a) $\tau \to \mu\nu\nu$ (b) $\tau \to \pi\nu$ (c) $\tau \to e\nu\nu$
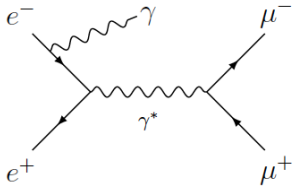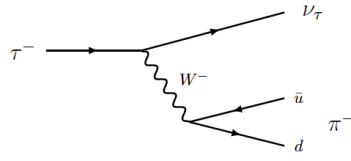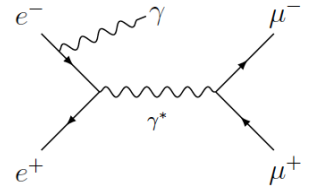
(d) $e^+e^- \to \mu^+\mu^-(\gamma)$ (e) $e^+e^- \to e^+e^-(\gamma)$ (f) $e^+e^- \to q\bar{q}$
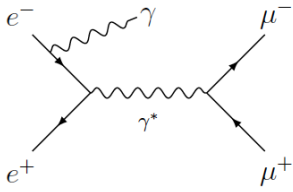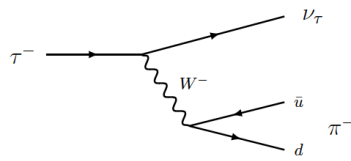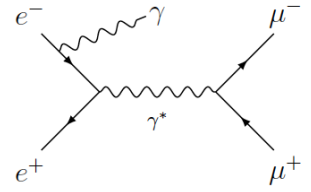
(g) $e^+e^- \to \Upsilon(4S)$ (h) $\Upsilon(4S) \to B^+B^-$ (i) $\Upsilon(4S) \to B^0\bar{B}^0$

Figure 3.1: Feynman diagrams for the dominant physics backgrounds for $\tau \to \mu\gamma$.

Background process	Cross-section [nb]	Generator
$\tau \to \mu\nu\nu$	0.160	KKMC, TAUOLA
$\tau \to \pi\nu$	0.100	KKMC, TAUOLA
$\tau \to e\nu\nu$	0.164	KKMC, TAUOLA
$\tau \to$ generic	0.919	KKMC, TAUOLA
$e^+e^- \to \mu^+\mu^-(\gamma)$	1.148	KKMC
$e^+e^- \to e^+e^-(\gamma)$	300	BABAYAGA.NLO
$e^+e^- \to u\bar{u}$	1.61	KKMC
$e^+e^- \to d\bar{d}$	0.40	KKMC
$e^+e^- \to s\bar{s}$	0.38	KKMC
$e^+e^- \to c\bar{c}$	1.30	KKMC
$e^+e^- \to B^+B^-$	0.525	EvtGen 1.3, PYTHIA 8.2
$e^+e^- \to B^0\bar{B}^0$	0.525	EvtGen 1.3, PYTHIA 8.2

Table 3.1: Fiducial cross-sections for backgrounds at Belle II.

Geometry information of the detector includes very specific location, thickness and material values; magnetic field strength through the detector is also known. The detector components are simulated by the generator, so that timing information and energy deposition is recorded for these components as MC particles interact with them. Following the simulation of a physics event, information from the sub-detectors is used to reconstruct tracks (correponding to charged particles) and clusters (referring to cells of the ECL with which a particle has interacted with, corresponding to photons).

3.2.1 Signal generation

Signal MC was produced using the KKMC generator [?]. Decays proceeded as $e^+e^- \to \tau^+\tau^-$, with one tau decaying to the signal mode $\tau \to \ell\gamma$ (the signal-side), and the other (the tag-side) decaying to all experimentally measured SM decay modes of the tau, scaled by their branching ratio; decaying according to the SM is referred to generic decay. The modes $\tau^+ \to \ell^+\gamma$, τ^- decaying generically, and $\tau^- \to \ell^-\gamma$, τ^+ decaying generically were both generated so that differences were not overlooked. To investigate both the electron and muon modes, two distinct final states were produced. The number of produced events was $3\,100\,000$ for the muon final state, and $2\,550\,000$ for the electron final state.

The KKMC generator....

3.2.2 Background generation

Samples for a range of background events were produced by the Belle II collaboration. **DO I NEED TO TALK ABOUT EACH OF THE GENERATORS? HOW MUCH DETAIL?**

3.3 Event scaling

Belle II is projected to collect a time-integrated luminosity of $50\,\mathrm{ab}^{-1}$ over the length of the experiment. Instead of running over the equivalent amount of background events expected in such a sample size we can run over a smaller number of events then scale results. This is done for computing reasons, as the number of background events recorded for these luminosities quickly becomes prohibitively large. **REPHRASE**.

We choose to scale up to a luminosity of $1\,\mathrm{ab}^{-1}$; this is the total time-integrated luminosity of the complete Belle dataset, and so is useful as a point of comparison to previous searches. The scale factor for each event type is calculated by

$$n_{1\,\mathrm{ab}^{-1}} = n_{\mathrm{generated}} \times \text{scale factor}, \tag{3.1}$$

where $n_{1\,\mathrm{ab}^{-1}} = \mathcal{L}\sigma$, and $n_{\mathrm{generated}}$ is the number of events generated. Scaled event numbers are presented in Table 3.2 below.

Event type	$n_{\mathrm{generated}}$	scale factor	$n_{1\,\mathrm{ab}^{-1}}$
$\tau \to \mu\gamma$	3 200 000	2.58×10^{-5}	83
$\tau \to e\gamma$	2 550 000	8.65×10^{-5}	221
$\tau \to \mu\nu\nu$	127 998 320	1.250	159 997 900
$\tau \to \pi\nu$	267 245 200	1.250	334 056 500
$\tau \to e\nu\nu$	131 086 160	1.250	163 857 700
$\tau \to \text{generic}$	208 870 320	1.250	261 087 900
$e^+e^- \to \mu^+\mu^-(\gamma)$	148 600 000	7.725	1 148 000 000
$e^+e^- \to e^+e^-(\gamma)$	15 630 000	19 193.858	300 000 000 000
$e^+e^- \to u\bar{u}$	1 268 991 935	1.269	1 610 000 000
$e^+e^- \to d\bar{d}$	317 048 262	1.262	400 000 000
$e^+e^- \to c\bar{c}$	1 039 855 756	1.250	1 300 000 000
$e^+e^- \to s\bar{s}$	289 900 586	1.310	380 000 000
$e^+e^- \to B^+B^-$	451 320 000	1.219	550 000 000
$e^+e^- \to B^0\bar{B}^0$	427 680 000	1.286	550 000 000

Table 3.2: Scaled event numbers

Unless explicitly stated, scaled event numbers will be used throughout this analysis as to provide accurate points of comparison between events.

3.4 Version differences

The software framework on which event generation was performed is undergoing continuous development; a majority of signal and background MC was produced on the same release version to ensure accuracy between MC types. This version was made available on XXXXX, and was one of the most up-to-date releases at the time of analysis. However, the backgrounds

$\mu^+\mu^-(\gamma)$ and $e^+e^-(\gamma)$ (mu-pair and bhabha, respectively) did not have any events generated using this release, and instead used an older release dated XXXXX.

A full investigation into the differences in generated MC between releases has not been performed in this analysis, as it is assumed they are negligible in most relevant cases. Changes relate mostly to generation of beam backgrounds (??); however one major change is in the reporting of timing data from the ECL. From the release dated XXXX onwards, these associated times will be reported in nanoseconds, rather than uncalibrated clock ticks (REF). In comparing samples of background MC generated in both the older and newer releases, Figure 3.2a and 3.2b below, as well as private correspondance with the developer responsible for these changes, it was found that conversion to the newer scale could be achieved by adding 80 units to the cluster timing values.
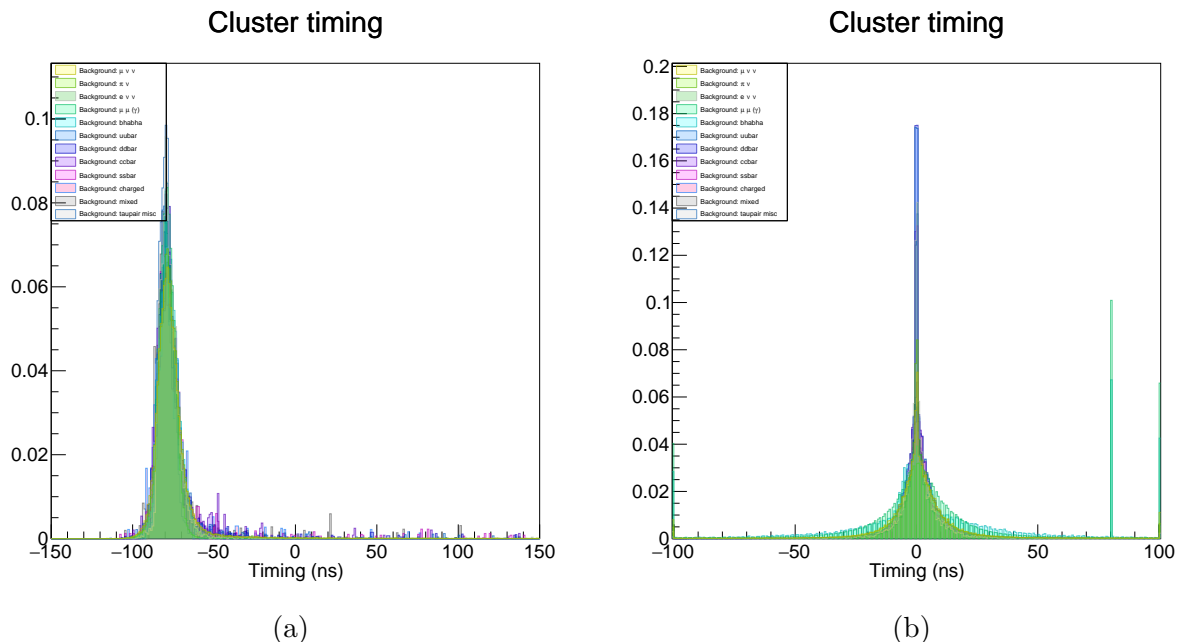
Figure 3.2: Comparisons of cluster timing with background MC (a) prior to ECL timing change and (b) after ECL timing change. Note the shift in peak by 80 units between versions.

3.5 Implementation of beam backgrounds

Signal and background MC was generated with simulated beam background. The most important sources have been generated by a dedicated accelerator group software within the Belle II collaboration. This software simulates particles travelling the detector and records the position and momentum of particles which leave the nominal beam trajectory and collide with the beam pipe or collimator (???). Detector response simulation is then processed for these particles to produce background samples of a given type. Separate backgrounds were made for electron

(HER) and positron (LER) beams, due to differences in energies and currents. Implemented background types are listed in Table 3.3.

type	source	rate [MHz]
radiative Bhabha	HER	1320
radiative Bhabha	LER	1294
radiative Bhabha (wide angle)	HER	40
radiative Bhabha (wide angle)	LER	85
Touschek scattering	HER	31
Touschek scattering	LER	83
beam-gas interactions	HER	1
beam-gas interactions	LER	156

Table 3.3: Beam background types simulated at Belle II (REF - Belle II Simulation (Belle II note)). Rate is calculated

For comparison, some samples of MC were generated without beam background. Comparisons between some measureables are shown in Figures 3.3a and 3.3b, 3.4a and 3.4b below.

(a) $E_{\text{total}}^{\text{CM}}$ without beam background

(b) $E_{\text{total}}^{\text{CM}}$ with beam background
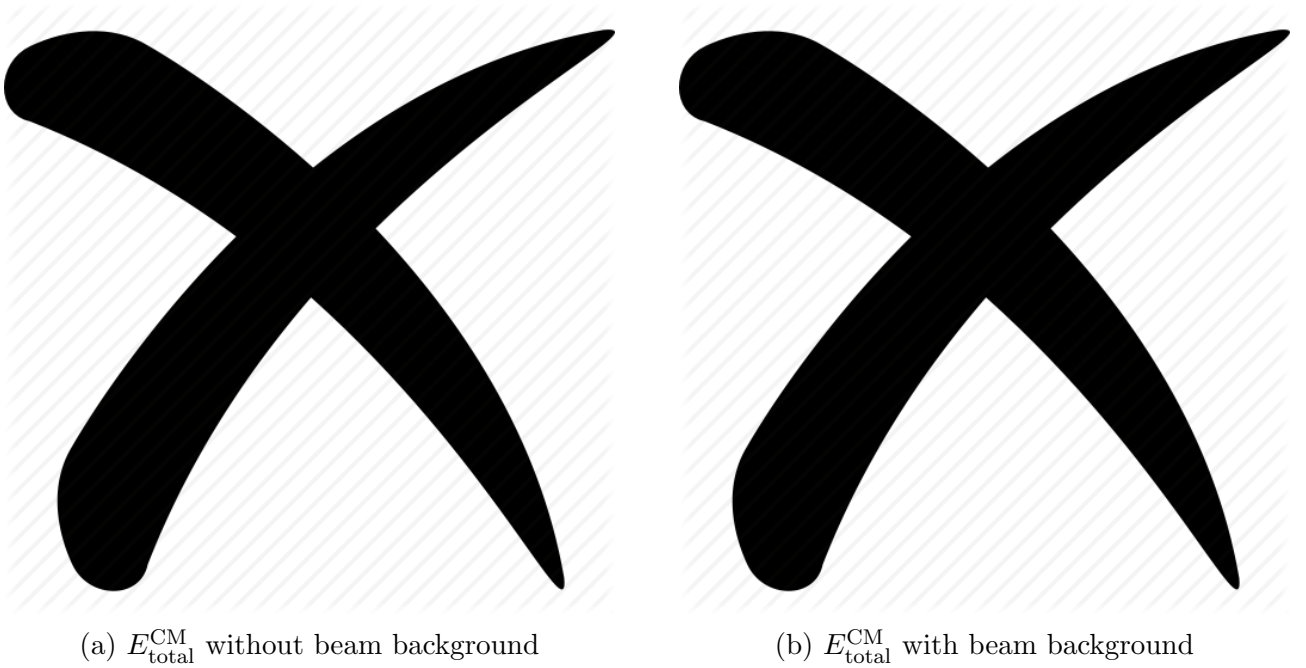
Figure 3.3: Comparison of total center-of-mass energies for signal and background distributions with and without beam background.

There are many differences between MC with and without beam background; only a few key points will be discussed here. Most obvious is the number of tracks recorded. Taking, for instance, our signal mode, we would expect for most events 2 or 4 tracks - one signal-side track corresponding to μ/e, and one or three tracks on the tag-side (one- or three-pronged) coming from standard model τ decays, which are dominantly one- or three-pronged. Beam background particles in the detector produce a number of new tracks (as well as clusters) which are tracked by sub-detector components, then later reconstructed.
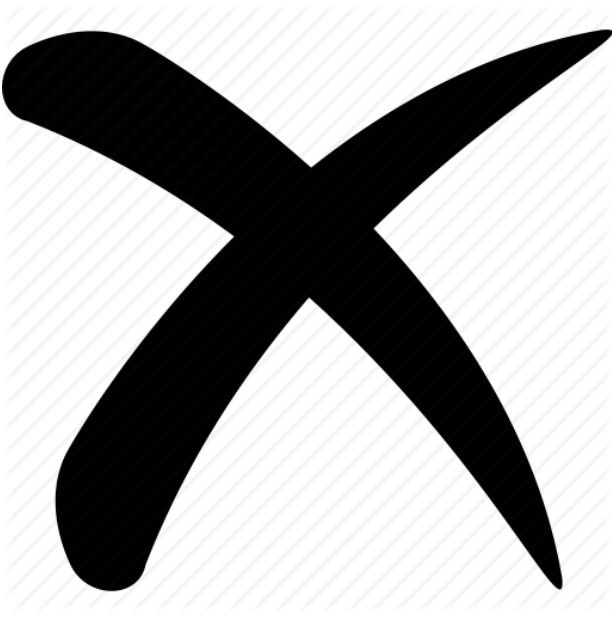
(a) Total center-of-mass energies w/o beam background

(b) Total center-of-mass energies w/ beam background

Kinematic variables such as energy and momentum are also obviously affected by beam background. We observe an increase in low energy and low momentum tracks; in these cases low energy beam background particles have been misidentified as signal- or tag-side particles originating from tau-pair processes.

4 | Reconstruction

Reconstruction of events was performed through `basf2`. The generated MC consists of information about charged tracks travelling through the detector geometry and interacting with sub-detectors, as well as cluster information relating to photons from the ECL, among other information.

4.1   Process

Belle II Note: Belle II Reconstruction Software and Performance

To reconstruct events, we fill lists of particles by categorising tracks as specific charged particles - muons, electrons or pions, and by associating clusters with photons. These categorisations are performed with loose criteria to remove obvious background events. Tracks with PID values greater than 0.1 (or 0.5?) are considered muons, and similar for electrons and pions; the photon list is populated by clusters passing a "goodness" test.

DISCUSS SIGNAL SIDE/TAG SIDE (+signal track, signal photon).

We define the signal track to be reconstructed track associated with the final state electron or muon from $\tau \to \ell\gamma$; similarly the signal photon describes the cluster associated with the final state photon. The signal-side particles do not necessarily refer to the physical charged particle or photon coming from the signal modes under analysis; misidentified particles (such as a pion being identified as a muon during reconstruction) or particles from different decay processes (such as a muon from the tau process $\tau \to \mu\nu\nu$).

We reconstruct the signal side tau as $\tau \to \mu\gamma$ for the muon mode, or $\tau \to e\gamma$ for the electron mode. The tag side tau is reconstructed by requiring at least one muon, electron or charged pion. In addition to PID cuts, we also apply loose cuts to our ΔE and M_{inv}. We define our signal region variables as

$$\Delta E = E_{\text{signal } \tau}^{\text{CM}} - E_{\text{beam}}/2, \tag{4.1}$$

$$M_{\text{inv}} = \text{invariant mass of reconstructed signal side tau}, \tag{4.2}$$

where $E_{\text{signal } \tau}^{\text{CM}}$ is the center-of-mass energy of the reconstructed signal-side tau, and E_{beam} is the total center-of-mass energy of the electron-positron beam system. Nominally we expect

signal events to have $\Delta E \sim 0$ and $M_{\text{inv}} \sim m_\tau$; signal region selection is further discussed in Section XXXX.

Once background has been sufficiently reduced through use of selection criteria, signal candidates will be chosen from events in the ΔE vs. M_{inv} phase space. As such we cut (???) only very loosely on these criteria wherever possible; in reconstruction we require $-0.4\,\text{GeV} < \Delta E < 0.2\,\text{GeV}$ and $1.6\,\text{GeV} < M_{\text{inv}} < 1.9\,\text{GeV}$ for the signal side tau.

The number of events over which reconstruction was performed, as well as the percentage of events after reconstruction/events before reconstruction (which we shall call *reconstruction efficiency* ϵ_{recon}), is shown in Table 4.1. **(I WANT TO REWORD THIS)**.

4.2   Reconstruction efficiencies

Event type	events in	events out	ϵ_{recon}
$\tau \to \mu\gamma$	3 200 000	5 435 660	169.86 %
$\tau \to e\gamma$	2 550 000	4 155 023	162.94 %
$\tau \to \mu\nu\nu$	127 998 320	542 504	0.42 %
$\tau \to \pi\nu$	267 245 200	861 406	0.32 %
$\tau \to e\nu\nu$	131 086 160	87 848	0.07 %
$\tau \to \text{generic}$	208 870 320	1 843 380	0.88 %
$e^+e^- \to \mu^+\mu^-(\gamma)$	148 600 000	5 159 295	3.47 %
$e^+e^- \to e^+e^-(\gamma)$	15 630 000	370 496	2.37 %
$e^+e^- \to u\bar{u}$	1 268 991 935	52 373 200	4.13 %
$e^+e^- \to d\bar{d}$	317 048 262	12 987 072	4.10 %
$e^+e^- \to c\bar{c}$	1 039 855 756	48 007 101	4.62 %
$e^+e^- \to s\bar{s}$	289 900 586	11 122 646	3.84 %
$e^+e^- \to B^+B^-$	451 320 000	46 498 047	10.30 %
$e^+e^- \to B^0\bar{B}^0$	427 680 000	46 912 228	10.97 %

Table 4.1: Reconstruction efficiency (unscaled events)

4.3   Track counting

Low-momentum tracks can "curl" around in the detector and be counted twice.

We expect our signal mode to be have two or four charged particle tracks in most cases - a muon on the signal-side, and one (or less often three) charged particles on the tag-side according to SM decays. However due to tracks introduced by beam backgrounds (and not-omniscient reconstruction) we cannot naively select events with only two or four tracks without removing prominent amounts of signal. Instead we require the number of reconstructed tracks to be less than 15 to suppress hadronic backgrounds.

4.4 Fake rates

5 | Event signatures

We discuss expected energies, kinematics and signal-side event topologies with comparison to MC events. Knowledge of event signatures is important in optimising selection criteria and hence the ratio of signal-to-background. This also serves as validation of the MC, as we compare physical expectation of certain variables to the simulated events.

5.1 Signal

5.1.1 Muon mode

The mode $\tau \to \mu\gamma$ has a simple signal side, with only one track and one photon, and no missing energy (in the form of neutrinos). If we assume that on average both τ particles generated in the e^+e^- collision share equally the energy generated in the interaction, we expect the mean signal-side center-of-mass frame energy to be $5.5\,\text{GeV}$. No physical preference is given to the energy distribution between the signal muon and signal photon, so their center-of-mass energies should be normally distributed around $2.75\,\text{GeV}$. Recalling the relation between energy, momentum and mass,

$$E^2 = p^2 + m^2, \tag{5.1}$$

since the muon energy is much greater than its mass of $105.66\,\text{MeV/c}^2$ this can simplify to

$$E^2 \approx p^2. \tag{5.2}$$

Note that p is the magnitude of the particle's 3-momentum.

We can predict and compare with MC several topological measureables, such as polar angle (i.e. angle from an axis parallel to the beam trajectory) as well as opening angles between particles; this is done for several key measureables. The average opening angle between the signal muon and signal photon in the center-of-mass frame can be calculated by boosting the expected back-to-back pair from the signal τ frame into the lab frame.

CALCULATIONS.

Similarly for SOME OTHER PAIRING,

CALCULATIONS.

As shown in Figures 5.3 - 5.4 below, the MC matches the predicted kinematics and topologies for the muon mode.

5.1.2 Electron mode

Much of the kinematics and topologies as described for the signal muon mode also applies to the signal electron mode. However, it is prudent to note the major difference between the two final state charged particles - specifically the difference in mass. Muons have mass $105.66\,\mathrm{MeV/c^2}$, over 200 times heavier than electrons, with mass of $0.511\,\mathrm{MeV/c^2}$. Due to their lighter mass, electrons lose far more energy due to bremsstrahlung (braking radiation), as given by

$$P = \text{radiated power} = \frac{e^2 a^2 \gamma^4}{6}, \tag{5.3}$$

where

$$e = \text{electron charge}, \tag{5.4}$$
$$a = \text{particle acceleration}, \tag{5.5}$$
$$\lambda = \text{relativistic factor}, \tag{5.6}$$
$$\tag{5.7}$$

and we have of course taken particle physics units $c = \pi = \epsilon_0 = 1$. Since $E = \gamma m c^2$, radiated power from bremsstrahlung goes as m^{-4}; electrons lose more energy via bremsstrahlung than muons by factor $(m_\mu/m_e)^4 \sim 200^4$.

This energy in the form of photons is deposited in the ECL clusters; a non-negligible amount of radiation from beam background is also deposited in these clusters, making accurate energy reconstruction of the signal electron difficult (??? is this right, or is it just that the electron has less energy? What is the "energy/momentum" variable.... E/P at creation?). Due to this energy loss we see a different energy and momentum signature for the signal track compared to the muon mode; peaks for both are located around XXXX GeV, but a large fraction of events have energies much lower.

Accurate reconstruction via `basf2` is more difficult for these lighter particles, with only a fraction of $\tau \to e\gamma$ events having a reconstructed τ invariant mass peak anywhere near the τ mass.

5.2 Backgrounds

5.2.1 Tau-pair processes

A key difference between tau-pair backgrounds and the signal modes investigated is the lack of signal-side photon, and the existence of signal-side neutrinos.
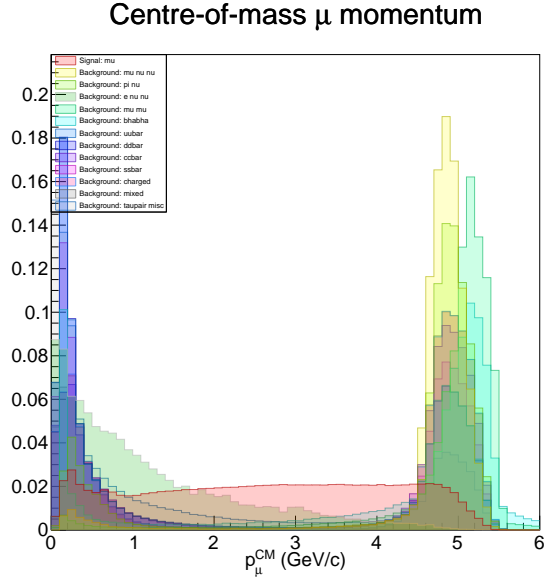
Figure 5.1: Center-of-mass frame momentum of signal muon.
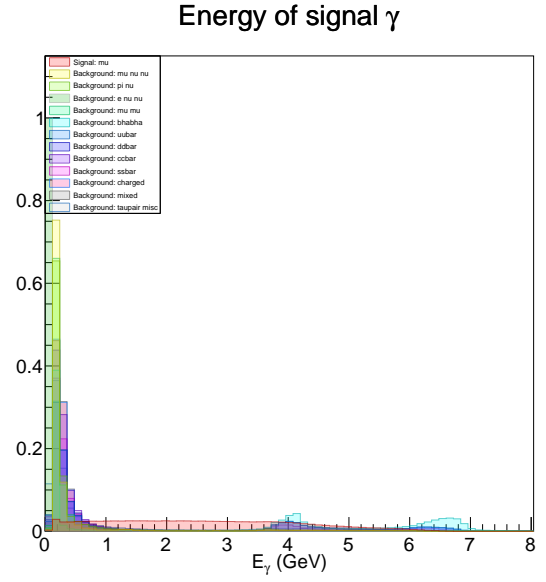

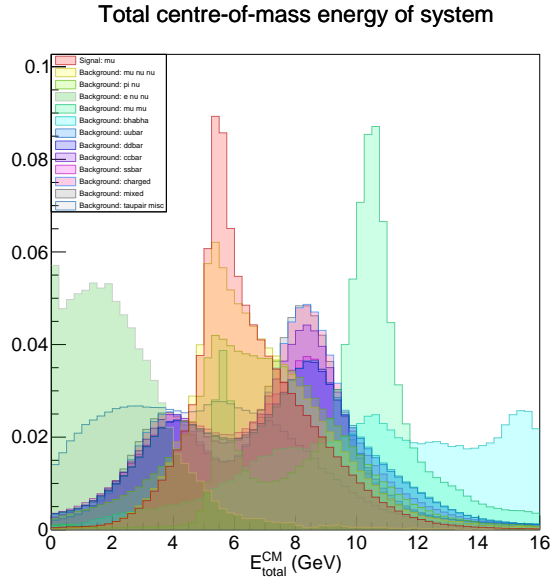Figure 5.2: Lab frame energy of signal photon.

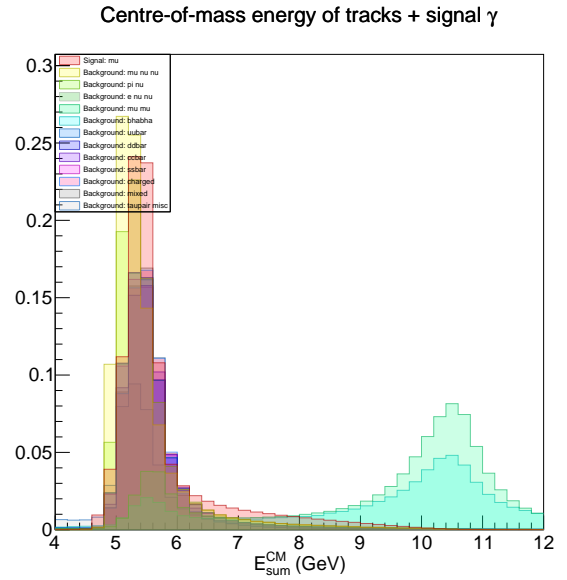Figure 5.3: Total center-of-mass energy of the system.

Figure 5.4: Center-of-mass frame energy of signal and tag tracks and signal photon.

We discuss the dominant tau-pair backgrounds $\tau \to e\nu\nu$, $\tau \to \mu\nu\nu$ and $\tau \to \pi\nu$ (where the pion is, of course, charged), with some minor discussion of the remaining modes. In tau-pair processes, the signal-side τ of energy $5.5\,\mathrm{GeV}$ decays dominantly into a single charged track and a neutrino (or neutrinos, in the pion modes). Light neutrinos carry away only a fraction of the energy from its mother particle; we expect the signal track to have a peak in energy of $\approx 5.5\,\mathrm{GeV}$. However, the electron will lose a significant fraction of its energy due to bremsstrahlung as it passes through the detector and hence have a lower average energy. Signal track center-of-mass momentum is shown in Figure XXXXX below.

None of the dominant tau decay processes proceed with the final state photon; in most cases the reconstructed signal photon comes from initial state radiation (???? discuss this) or beam backgrounds. Photons produced as beam background are often low energy, especially compared to the average photon energy from signal processes of $\sim 2.25\,\mathrm{GeV}$. Signal photon energy for tau-pair processes is shown in Figure XXXX; we note the peak energy of of $< 1\,\mathrm{GeV}$.

Since the signal photon does not originate from the signal-side τ, the signal tracks appears "boosted" by comparison. This "boost" leads to the signal track and signal photon travelling almost back-to-back in the center-of-mass frame, in contrast to the signal mode where these final state particles have only a small opening angle between them.

Since the signal track originates from the signal-side τ and is hence boosted in the center-of-mass frame by its mother particle,

Since this signal photon does not originate from the signal-side τ, but instead from beam background or occasionally the tag-side τ.

SOMETHING ABOUT E9E25.

5.2.2 Mu-pair processes

Muons are very cleanly reconstructed by the detector; they do not lose any significant amount of energy through bremsstrahlung radiation, and they penetrate deeply leaving a distinct signal. Unlike the tau-pair processes or the signal modes discussed above, mu-pair processes $e^+e^- \to \mu^+\mu^-(\gamma)$, both signal- and tag-side channels consist of only a single charged track each, sometimes with a final state photon. Hence, mu-pair processes are reconstructed very well. This is evidenced by the majority of events with only two reconstructed tracks (see Figure XXXX); this is is stark comparison to the spread of tracks from two up to fifteen in tau-pair events. Total reconstructed energy peaks around $10.5\,\mathrm{GeV}$.

By momentum conservation, the signal and tag tracks are generated back-to-back for $\mu\mu$ final states (XXX % of all mu-pair processes), or with an opening angle similar to the signal mode for $\mu\mu\gamma$ final states (XXXX % of all mu-pair processes). The signal photon is often reconstructed from low energy beam background photons, as with background tau-pair processes. A similar photon energy spectrum can be seen in Figure XXXX.
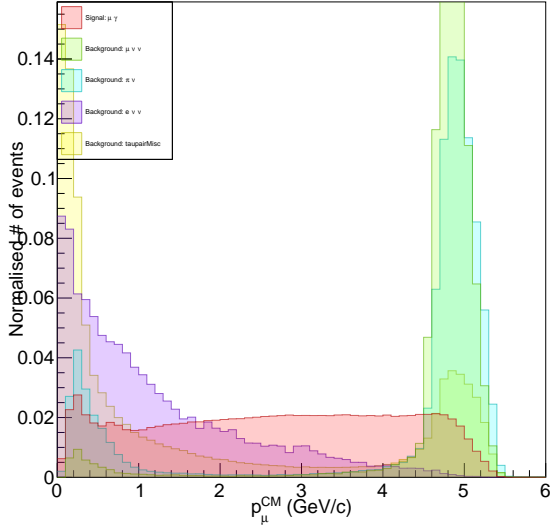
Figure 5.5: Taupair - muCM P.
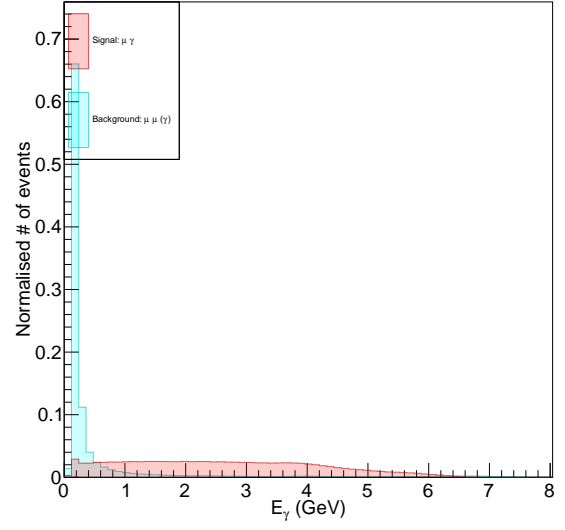
Figure 5.6: Mu-pair: gamma E.

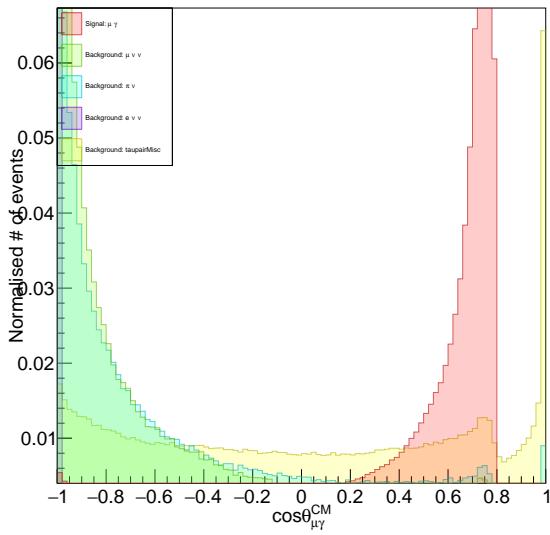Figure 5.7: Taupair - muGammaOpeningCos-ThetaCM.
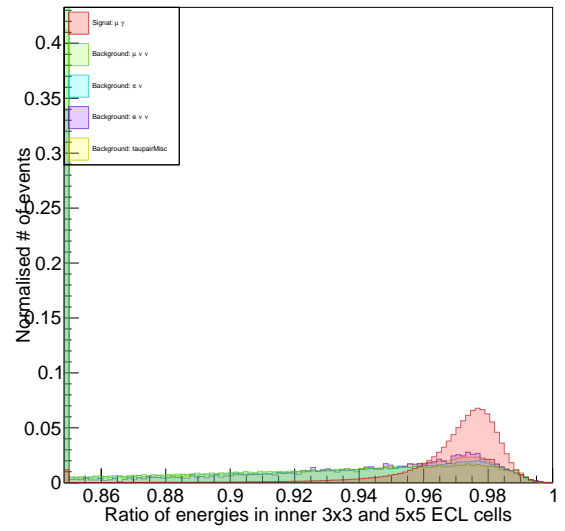

Figure 5.8: Taupair - E9E25.

5.2.3 Bhabha

Bhabha events $e^+e^- \to e^+e^-(\gamma)$ have similar event signatures to mu-pair processes $e^+e^- \to \mu^+\mu^-(\gamma)$ across a range of variables, for obvious reasons. However, as is case with electrons in the detector, both signal- and tag-track particles lose a significant portion of their energy to bremsstrahlung and are not reconstructed as cleanly as muons.

Smearing of energies and momenta occurs for bhabha event reconstruction, with measureables not peaking as strongly as for mu-pair events and instead spreading across a values. These features are shown in Figures 5.9 and 5.10 below, comparing distributions for mu-pair events to bhabha events. Much of the event topology is still common between these processes, however.
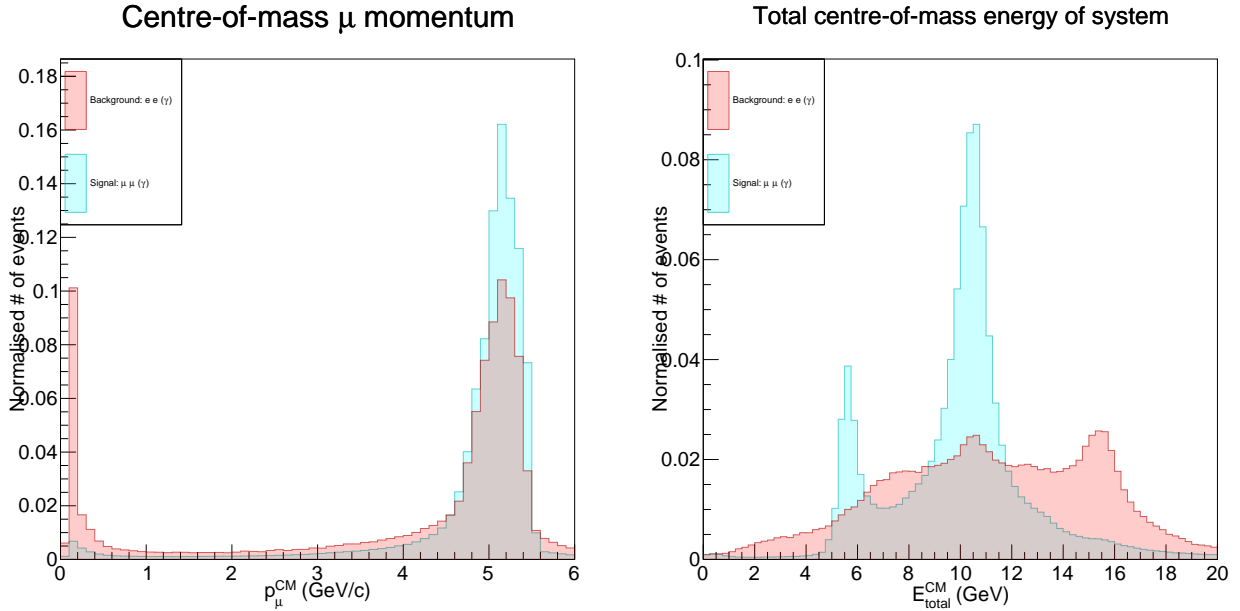
Figure 5.9: Comparison of center-of-mass frame signal track momentum for Bhabha and mu-pair events.

Figure 5.10: Comparison of total energy of the system in center-of-mass frame for Bhabha and mu-pair events.

5.2.4 Continuum and $B\bar{B}$

Hadronic backgrounds at Belle II come from the events $e^+e^- \to q\bar{q}$ and $e^+e^- \to B\bar{B}$. The event rate for the energies at which we investigate is dominated by these processes, however their event signatures are distinct from leptonic processes in many variables allowing good separation.

HMMM DON'T KNOW WHAT TO WRITE ABOUT. SOMETHING ABOUT JETS?

Figure 5.11: Number of tracks in continuum and $B\bar{B}$ events.

6 | Preselection

Following reconstruction, preselection criteria were applied to the reconstructed events. Preselection criteria are defined as distinct from selection criteria in that they remove a minimal amount of signal while removing the more obvious background components; in choosing selection criteria we seek to maximise $S/\sqrt{S+B}$, which may necessarily involve the exclusion of a non-negligible amount of signal.

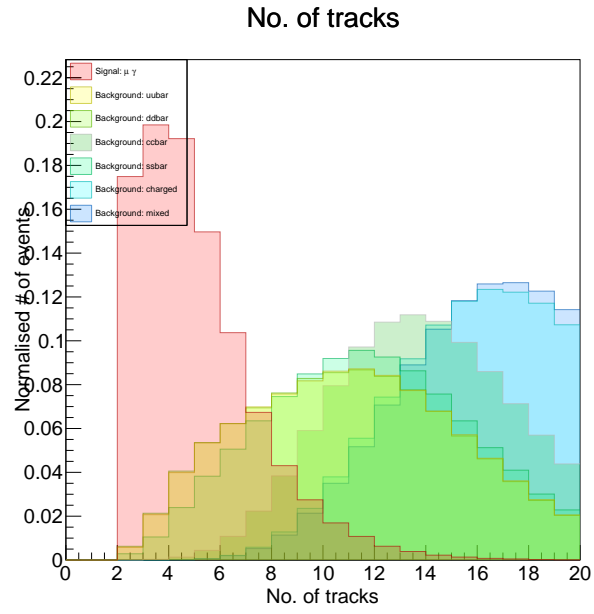The preselection criteria were selected by inspection of plots of various topology and energy based variables. Different preselection and selection criteria are chosen for each final state mode, due to the different event signatures. A total of 5 variables where chosen for preselection.

6.1 Muon mode

Figures xxx - xxx below show the variables on which preselection criteria were applied, with the specific values for preselection listed in Table xxx. Signal and background efficiencies after preselection are shown in Table XXXX.

98.06% of reconstructed $\tau \to \mu\gamma$ signal events pass the preselection criteria; around 1,000,000,000 background events out of 2,000,000,000 are removed through this process. Most notably affected by preselecton are the continuum events of which only XXXX% remain. This reduction is greatly attributed to the clear separation between signal and continuum background in the $E_{\text{total}}^{\text{CM}}$.

symbolic	description	lower	upper
$p_{\text{tag}}^{\text{CM}}$	CM momentum of tag track	—	$5.2\,\text{GeV}$
$\cos\theta_{\text{signal}}$	Cosine of polar angle of signal track	-0.9	—
$E_{\text{total}}^{\text{CM}}$	Center-of-mass energy of total system	—	$15\,\text{GeV}$
|thrust|	Magnitude of signal thrust vector*	0.92	—
$E_{\text{sum}}^{\text{CM}}$	Center-of-mass energy of photons and tracks	$4.5\,\text{GeV}$	—

Table 6.1: Preselection cuts (muon mode)

We also place a loose requirement of M_{inv}, requiring $1.65 < M_{\text{inv}} < 1.85$. Although this is one of our signal region variables, we are justfied in applying loose preselection on it because events outside this selection will lie outside the final signal region, and hence be unimportant to the analysis.

MC type	events in (scaled)	events out	ϵ_{ps}
$\tau \to \mu\gamma$	140	135	96.32 %
$\tau \to \mu\nu\nu$	678 130	631 235	93.08 %
$\tau \to \pi\nu$	1 076 758	822 028	76.34 %
$\tau \to e\nu\nu$	109 810	70 675	64.36 %
$\tau \to$ generic	2 304 225	636 954	27.64 %
$e^+e^- \to \mu^+\mu^-(\gamma)$	39 857 811	21 938 180	55.04 %
$e^+e^- \to e^+e^-(\gamma)$	7 111 247 601	1 457 005 758	20.49 %
$e^+e^- \to u\bar{u}$	66 447 114	29 767 142	44.80 %
$e^+e^- \to d\bar{d}$	16 384 978	7 307 320	44.60 %
$e^+e^- \to c\bar{c}$	6 017 200	23 401 910	38.99 %
$e^+e^- \to s\bar{s}$	14 579 500	6 299 797	43.21 %
$e^+e^- \to B^+B^-$	56 664 730	22 083 348	38.97 %
$e^+e^- \to B^0\bar{B}^0$	60 329 511	23 372 366	61.26 %

Table 6.2: Preselection efficiency

6.2 Electron mode

Separate preselection criteria was chosen for the electron mode. These criteria do not differ much, as expected, however due to the prevalence of bremsstrahlung in the electron mode the energy variables differ greatly. Due to this we set an upper limit on $E_{\text{sum}}^{\text{CM}}$ rather than a lower limit as for the muon mode.

symbolic	description	lower	upper
$p_{\text{tag}}^{\text{CM}}$	CM momentum of tag track	—	5 GeV
$\cos\theta_{\text{signal}}$	Cosine of polar angle of signal track	-0.975	—
$E_{\text{total}}^{\text{CM}}$	Center-of-mass energy of total system	—	14 GeV
|thrust|	Magnitude of signal thrust vector*	0.92	—
$E_{\text{sum}}^{\text{CM}}$	Center-of-mass energy of photons and tracks	4.5 GeV	—

Table 6.3: Preselection cuts (electron mode)

Again, a loose requirement of $1.65 < M_{\text{inv}} < 1.85$ is required for event preselection. Following reconstruction

Note the dramatic drop in the number of $\tau \to e\gamma$ signal events. This is due to the restriction of M_{inv}. We expect a peak around the invariant mass of the τ, $m_\tau \sim 1.777\,\text{GeV}/c^2$; see Figure 6.1. For the electron mode a peak occurs slightly after that mass, around $1.8\,\text{GeV}/c^2$. More important, however, is the high multiplicity of events at $M_{\text{inv}} = 0$. Obviously this is an unphysical value — we determine that this feature is due to the large number of signal electron tracks which are not reconstructed properly (**ELABORATE ON THIS**).

After preselection, the signal electron mode has an efficiency of only 9.95 %. The total efficiency of the signal muon mode after all selection criteria have been applied (Section XXXXX) is, in comparison, 7.23 % - we would expect that the efficiency of the electron mode would get considerably worse after event selection. Due to this poor reconstruction and minimal number

MC type	events in (scaled)	events out	ϵ_{ps}
$\tau \to e\gamma$	140	29	13.25 %
$\tau \to \mu\nu\nu$	678 130	645 716	95.22 %
$\tau \to \pi\nu$	1 076 758	872 186	81.00 %
$\tau \to e\nu\nu$	109 810	0	0 %
$\tau \to$ generic	2 304 225	758 109	%
$e^+e^- \to \mu^+\mu^-(\gamma)$	39 857 811	18 676 083	%
$e^+e^- \to e^+e^-(\gamma)$	7 111 247 601	1 483 493 282	20.49 %
$e^+e^- \to u\bar{u}$	66 447 114	35 306 780	%
$e^+e^- \to d\bar{d}$	16 384 978	8 692 996	%
$e^+e^- \to c\bar{c}$	6 017 200	30 600 420	%
$e^+e^- \to s\bar{s}$	14 579 500	7 675 121	%
$e^+e^- \to B^+B^-$	56 664 730	28 912 567	%
$e^+e^- \to B^0\bar{B}^0$	60 329 511	30 716 377	61.26 %

Table 6.4: Preselection efficiency

of events remaining only after preselection, we do not continue to full event selection in MC for $\tau \to e\gamma$. Unless otherwise specified, the remainder of the analysis will concern the mode $\tau \to \mu\gamma$ only.

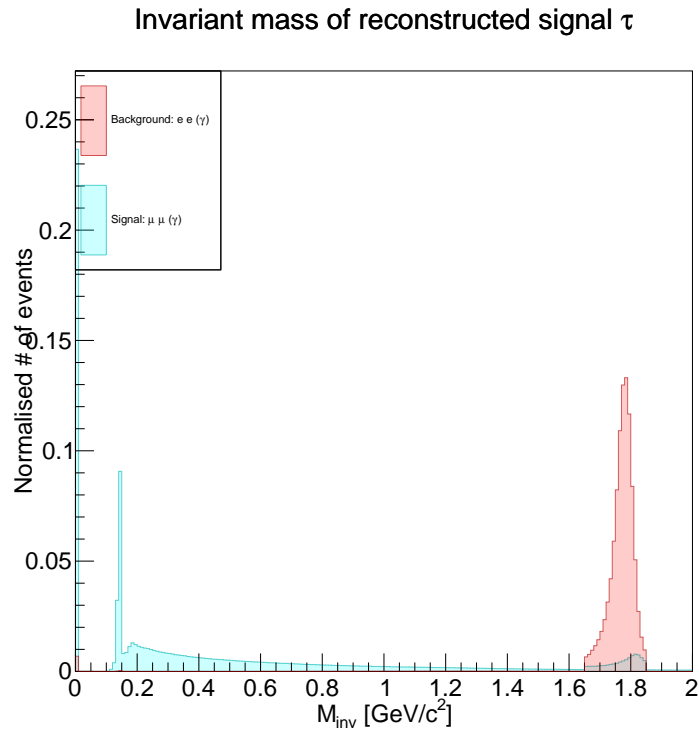**Note: should perhaps investigate/talk more about the underlying issues.**

Invariant mass of reconstructed signal τ

Figure 6.1: Comparison of M_{inv} for signal muon and signal electron modes. Note that we expect a peak around $m_\tau \sim 1.777\,\mathrm{GeV/c^2}$.

7 | Signal optimisation

There are many reasonable ways to quantify signal optimisation compared to background in a particle physics analysis. For this analysis, we have chosen the measure of signal optimisation $S/\sqrt{S+B}$. S is the expected number of signal events remaining after event selection has been performed; note that we scale for our signal events dependent on their associated LFV branching fractions, which are currently experimental upper bounds. Hence the measure $S/\sqrt{S+B}$ will vary for different branching fraction hypotheses. Similarly B is the expected number of background events remaining after event selection; specifically this is the sum of all events remaining across all background MC used in analysis. Since not all possible background processes have been included we would expect the empirical value of B to be higher; this is mostly unimportant for our analysis since we need only compare changes in $S/\sqrt{S+B}$ between different selection criteria, not to know the absolute value.

To maximise $S/\sqrt{S+B}$, the number of background events must be reduced while not removing too many sgnal events. This is especially important given the low branching fraction of our signal - taking the experimental upper limit of the $\tau \to \mu\gamma$ ($\tau \to e\gamma$) branching fraction, we expect only 135 (XXX) signal events, after reconstruction and preselection, in a $1000\,\text{fb}^{-1}$ dataset, compared to over $1\,500\,000\,000$ (1.5 billion) background events in the same luminosity. The actual number of signal events is very likely lower, due to double-counting in reconstruction (REPHRASE. Also, something to say about lower branching fraction?).

Forty selection criteria were used in selecting for $\tau \to \mu\gamma$ events. Variables chosen for selection originally came from the previous search for $\tau \to \ell\gamma$ at Belle (REF); more were added as the analysis progressed to achieve better greater signal separation. Some of these variables were not included in the final analysis. Notable were those related to missing momentum and mass, excluded due to poor reconstruction resulting in poor event separation. Others such as tag-track transverse momentum and signal-track momentum were not included due to significant overlap of background exclusion with other variables (most obviously center-of-mass tag track momentum and center-of-mass signal-track momentum) causing selection on these measureables to have little to no impact on signal optimisation.

The thresholds were chosen via a manuallly iterative process, continually improving the ratio $S/\sqrt{S+B}$ by adding or refining criteria. Initial criteria were informed by a combination

of automatic iteration over a range of threshold values for around 20 variables, and by visual inspection.

The accomplish the former, selection was repeatedly performed over all MC with only a single criteria applied. The lower or upper threshold of this criteria was modified by a small increment each time. The resulting value of $S/\sqrt{S+B}$ was then plotted against the selection threshold; this plot is called a *figure-of-merit* plot. An obvious peak, as in Figure XXXX, indicates a beneficial (??) value for selection. For many variables, no peak was apparent and so no initial threshold could be chosen through this method.
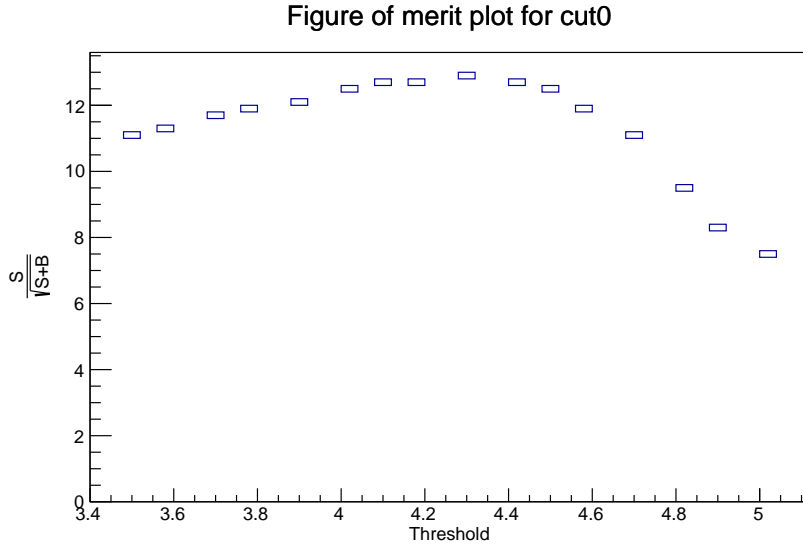

Figure 7.1: Figure-of-merit plot for $p_{\text{signal}}^{\text{CM}}$.

Initial visual inspection was performed by finding obvious separation between signal and many background variables. This was informed by knowledge of event topologies (Section XXXX), as well as some figure-of-merit plots as described above. Figures XXXX and XXXXX are examples of good and poor selection variables. In Figure XXXX, large background peaks occuring at less than $1\,\text{GeV}/c$ and greater than $4.3\,\text{Gev}$ lead to great separation from signal, which does not peak anywhere in those regions. Figure XXXX offers less than optimal threshold positions, as there is no apparent way to remove any amount of background without also removing a large amount of signal; note that this variable was not included during event selection.

7.1 Muon mode

Figure XXXX shows the number of signal and background events remaining after selection has been applied. Specific discussion and justification of selection criteria is given below.
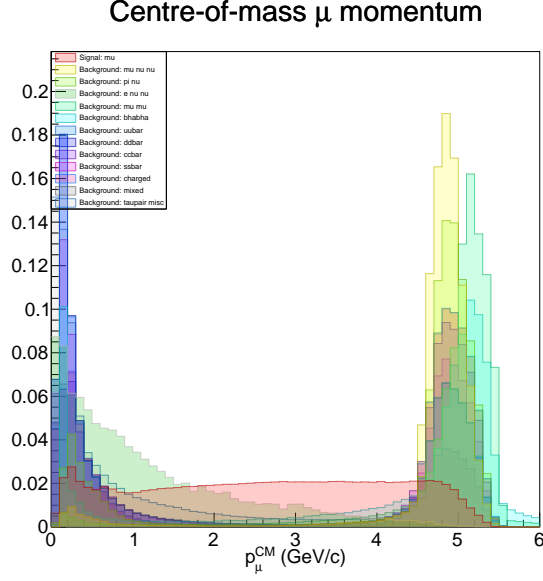
Centre-of-mass μ momentum

Figure 7.2: Center-of-mass frame momentum of signal track.

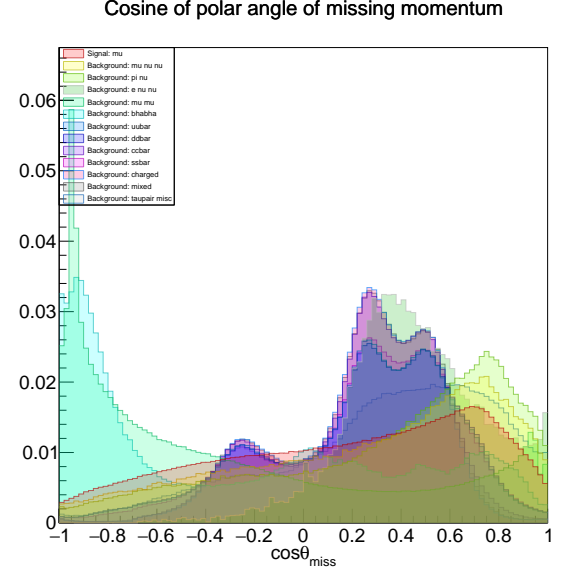**Cosine of polar angle of missing momentum**

Figure 7.3: Cosine of polar angle of missing momentum.

7.1.1 Energy-momentum

I need some way to say this without saying "we require" repeatedly. Could borrow from Hayasaka-san

We require p_μ^{CM} greater than $1\,\mathrm{GeV}/c$ and less than $4.3\,\mathrm{GeV}/c$, to remove low-momentum tracks from continuum events and to remove high-momentum tracks from leptonic processes without high-energy final state photons, that is, generic tau-pair processes, mu-pair processes, and bhabha. Muon transerve momentum p_μ^{t} is required to be greater than $0.1\,\mathrm{GeV}/c$; tracks with a lesser transverse momentum could not reach the CDC sub-detector component so we do not select these events. Low-energy tracks from continuum events often do not satisfy this criteria, however most low-energy tracks are excluded through selection on p_μ^{CM}. To suppress $\mu^+\mu^-$ and Bhabha backgrounds we require center-of-mass momentum of the tag track $p_{\mathrm{tag\ track}}^{\mathrm{CM}}$ less than $2.5\,\mathrm{GeV}$. Through conservation of momentum we expect the signal and tag tracks of the processes $e^+e^- \to \ell^+\ell^-(\gamma)$ we to have momentum peaks around $5\,\mathrm{GeV}$, which are excluded by this selection. Since most background processes reconstruct the signal photon from low-energy photons, we require the energy of the signal photon E_γ to be $0.8\,\mathrm{GeV}$. This selection criteria is effective in removing background events across different processes. Total energy in the center-of-mass frame $E_{\mathrm{total}}^{\mathrm{CM}}$ is required to be less than $12\,\mathrm{GeV}$ to suppress $\mu^+\mu^-$ events.

NOTE: variables I didn't use but could have used $= p_{\mathrm{tag\ track}}^{\mathrm{t}}$, p_μ, $E_{\mathrm{sum}}^{\mathrm{CM}}$

Event process	events in	events out
$\tau \to \mu\gamma$	135 ± 1	6
$\tau \to \mu\nu\nu$	$631\,235$	163
$\tau \to \pi\nu$	$822\,028$	40
$\tau \to e\nu\nu$	$70\,675$	0
$\tau \to$ generic	$636\,954$	0
$e^+e^- \to \mu^+\mu^-(\gamma)$	$21\,938\,180$	15
$e^+e^- \to e^+e^-(\gamma)$	$1\,457\,005\,758$	0
$e^+e^- \to u\bar{u}$	$29\,767\,142$	9
$e^+e^- \to d\bar{d}$	$7\,307\,320$	3
$e^+e^- \to c\bar{c}$	$23\,401\,910$	0
$e^+e^- \to s\bar{s}$	$6\,299\,797$	3
$e^+e^- \to B^+B^-$	$22\,083\,348$	0
$e^+e^- \to B^0\bar{B}^0$	$23\,372\,366$	0

Table 7.1: Events remaining after selection (electron mode)

7.1.2 Angular relations

As discussion in Section XXXX, angular distributions of reconstruction particles depend strongly on the decay process and hence can be used as discriminating variables. For this analysis, $\cos\theta_\mu$ is required to be in the range -0.8 to 0.9 to suppress continuum, mupair and Bhabha processes. Cosine of the signal photon polar angle $\cos\theta_\gamma$ is required to be less than 0.85. We require the polar angle of the reconstructed signal τ in the center-of-mass frame to be between 1 and 2.2. Note that these variables are correlated and overlap in the background which is removed by this selection. Selection is also made on the cosine of the helicity angle $\cos\theta_\mathrm{H}$, requiring $-0.8 < \cos\theta_\mathrm{H} < 0.8$. The helicity angle is the opening angle between the signal muon in the signal τ frame, and the boost direction of the τ in the center-of-mass frame.

7.1.3 PID likelihoods

Signal-side tracks are identified as muons by requiring $\mathcal{L}_\mu > 0.8$. To reduce the effect of double-counting, mostly for signal events, we require $\mathcal{L}_K < 0.08$ and $\mathcal{L}_e < 0.005$. Due to how events are reconstructed in `basf2` we also require $\mathcal{L}_\pi > 0.8$ (WHY IS THIS??). On the tag-side, we require $\mathcal{L}_\mu < 0.85$ to identify the tag-track as not a muon.

SOMETHING ABOUT $\mathcal{L}_\mu$ being not perfect? Very rough on the tag-side.

7.1.4 Background suppression using event shape

The thrust axis for a collection of N particles with momenta $\mathbf{p}_i$ $(i = 1, \ldots, N)$ is defined as the unit vector along with the total projection of these momenta is maximised. The thrust scalar, or thrust, is defined as

$$T = \frac{\sum_{i=1}^{N}|\mathbf{T} \cdot \mathbf{p}_i|}{\sum_{i=i}^{N}|\mathbf{p}_i|}. \tag{7.1}$$

Signal thrust and rest-of-event thrust variables are used for selection. The signal thrust axis is constructed from all particles used to reconstruct the signal-side τ; for our signal mode these are the signal muon and signal photon. On average these particles have similar momentum and a small opening between them, so that signal thrust has a clear peak around 0.942. We consider the signal thrust axis for the other dominant leptonic processes. For most events the signal-side is reconstructed from a high-energy charged track and a low-energy photon generated from bremsstrahlung, beam background or similar (radiative mu-pair or Bhabha events can generate higher energy final-state photons). Regardless of opening angle, the difference in momentum between signal track and signal photon leads to signal thrust for this event peaking around 1. A requirement for signal thrust to be in the range 0.936 to 0.944 to suppress leptonic backgrounds as shown in Figure XXXX.
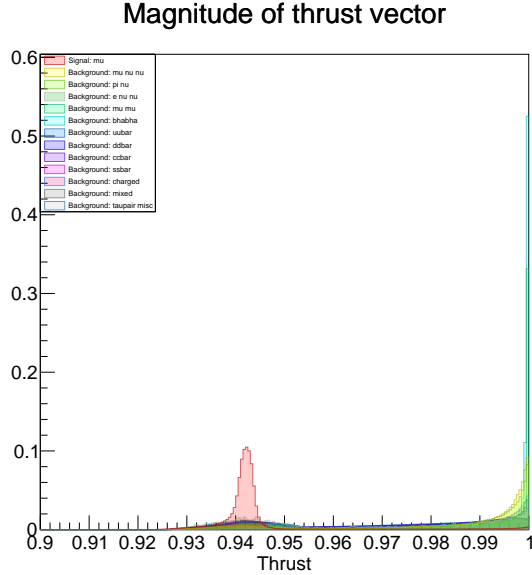

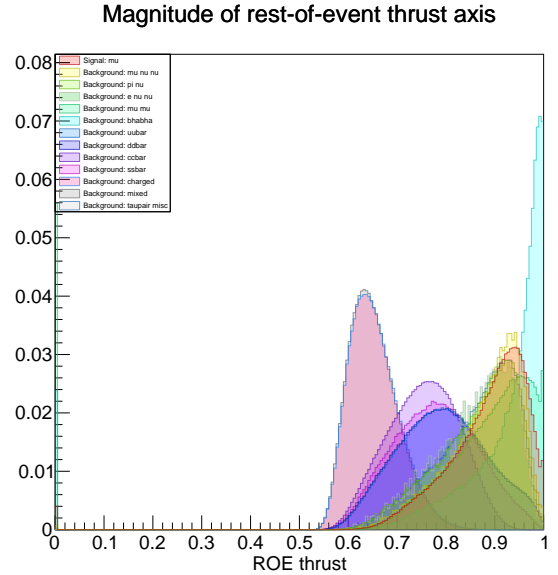Figure 7.4: Signal thrust.

Figure 7.5: Rest-of-event thrust.

We construct another thrust axis is also built using rest-of-event (ROE) data. ROE comprises all tracks and clusters not associated with the signal-side reconstruction, and so ROE thrust is influenced by tag-side events. Hadronic processes such as $B\bar{B}$ events and continuum have on average a lower rest-of-event thrust than leptonic processes due to the number of tracks and photons produced due to interactions with the detector. The magnitude of this ROE thrust vector is required to be in the range from 0.85 and 0.98. As shown in Figure XXXX, events produce very clean Gaussian-like distributions with little trailing tail. As such, while this selection removes a non-negligible amount of signal, it suppresses continuum events very strongly and almost completely excludes $B\bar{B}$ events. Event selection is also performed using the angle between thrust axes. We require the cosine of the angle between the signal thrust vector and rest-of-event thrust vector to be greater than 0.7.

Several continuum suppression variables were chosen by inspection after many more obvious criteria had been applied. These are CLEO cones, Super Fox-Wolfram moments, and the reduced Fox-Wolfram moment (the ratio of the second to the zeroth Fox-Wolfram moments).

Inherent differences in event shape can be used to suppress backgrounds. Some useful disciminant variables are the CLEO cones, named as such due to being introduced in a 1995 paper by the CLEO collaboration (REF). To calculate these variables, the space around the signal thrust axis is divided into cones at nine polar angle intervals of $10°$ each, with the ith interval covering angles from $(i-1) \times 10°$ to $i \times 10°$ from the thrust axis (see Figure XXXX). Forward and backward intervals are combined. CLEO cone variables cci are defined as the lab-frame momentum flow of the ith cone; momentum flow in each cone is calculated as the scalar sum of all tracks and clusters in that cone. Requirements on CLEO cones are given in Table XXXX.
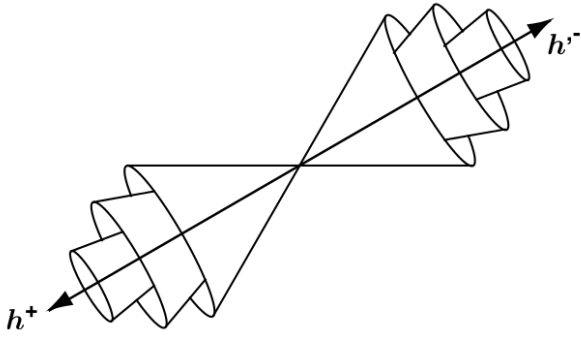

Figure 7.6: Illustration of cleo cones.

CLEO cone	lower	upper
cc1	–	5
cc2	2.4	–
cc3	–	–
cc4	–	1.7
cc5	–	0.9
cc6	–	0.7
cc7	–	0.5
cc8	–	–
cc9	–	0.4

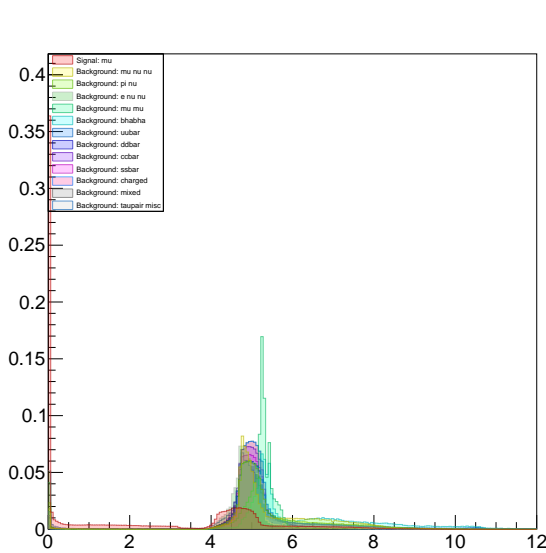Figure 7.7: CLEO cone selection criteria.
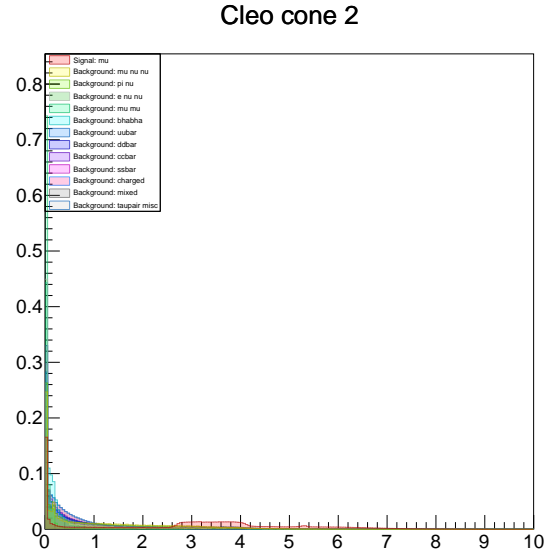

Figure 7.8: CLEO cone 1.

Figure 7.9: CLEO cone 2.

42

Fox-Wolfram moments (FWM) are event shape variables which describe energy flow from high-energy particle collision events, introduced to describe e^+e^- annihilation event shapes. Variables h_i^{so} ($i = 2, 4$) and h_j^{oo} are the normalised FWM, defined as

$$h_l^k = \frac{\sum_{m,n} |\vec{p_m}||\vec{p_n}| P_l(\cos\theta_{mn})}{\sum_{m,n} |\vec{p_m}||\vec{p_n}|},$$ (7.2)

where $\vec{p_m}$ and $\vec{p_n}$ are the momenta of particles m and n and $P_l(\cos\theta_{mn})$ is the l-th other Legendre polynomial of cosine of the angle θ_{mn} between $\vec{p_m}$ and $\vec{p_n}$. We categorise the type of FWM with $k = so, oo$. For h_i^{so}, m is from signal particles and n is from ROE; for h_j^{oo} both m and n are from ROE.

FWM	lower	upper
Hso(0,0)	0.05	1
Hso(0,1)	−0.05	0.3
Hso(0,2)	–	0.48
Hso(0,3)	−0.1	0.25
Hso(2,0)	−0.1	1
Roo(1)	−0.018	0.08
Roo(3)	−0.01	0.007

Table 7.2: Super Fox-Wolfram moment selection criteria

Note: not entirely sure what Hso(0,0) etc. means

Two other event shape variable requirement are required on the reduced Fox-Wolfram moment R_2 — the ratio of the second to the zeroth Fox-Wolfram moments — and the cosine of the angle between the signal thrust axis and the beam axis $\cos\theta_{\text{TBz}}$. We require $R_2 > 0.4$ and $\cos\theta_{\text{TBz}} < 0.8$, with both variables providing good background suppression as seen from Figures XXXX and XXXX.

7.1.5 Remaining criteria

We require the number of reconstructed tracks to be less than 15 to suppress hadronic backgrounds; discussion of track counting is found in Section XXXXX. Selection is also made on the total energy deposited across all ECL clusters by neutral particles, in the range of 2 to 6 GeV to suppress XXXXXX. The ratio of energies in the inner 3x3 and 5x5 cells of the ECL, or E9E25, peaks close to 1 for the signal mode, since the signal photon carries a large fraction of total energy deposited in the ECL. We require E9E25 greater than 0.95 to suppress tau-pair backgrounds.

The emission of a photon after the short τ lifetime via $\tau \to \mu\gamma$ gives cluster timing peak around zero with a tight width, so we require this variable to be in the range -1 ns to 1 ns. POCA STUFF: Point-of-closest-approach (POCA); $-0.05 < d_0 < 0.05$, $-0.06 < z_0 < 0.06$;
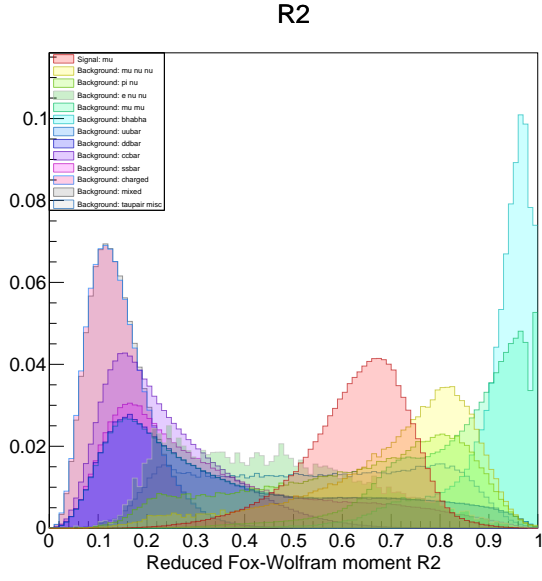
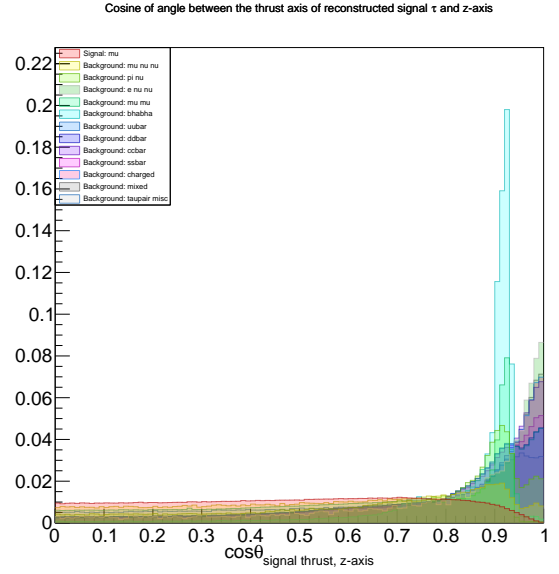Figure 7.10: Reduced Fox-Wolfram moment R_2.

Figure 7.11: Cosine of the angle between the thrust axis of the signal track and z-axis.

signed distance to the POCA in the r-phi plane; z coordinate of the. These impact parameter variables become more important when analysing actual data rather than MC, due to XXXXXX. Hence while these selection criteria to not discriminate background in a major way in MC, they may become more important when running over Belle II data.
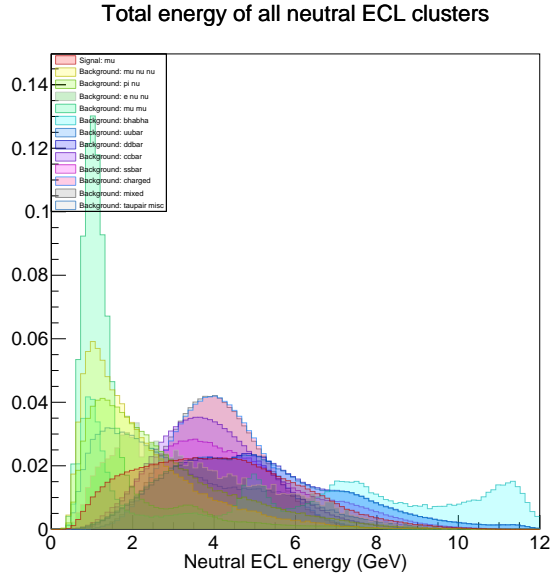
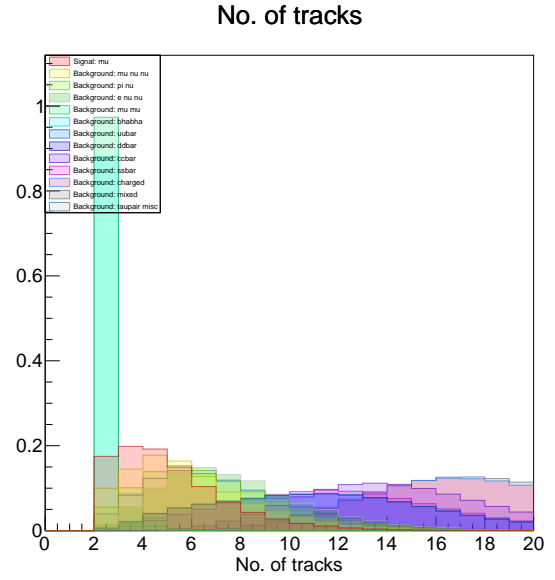Figure 7.12: Total energy deposited in ECL clusters by neutral particles.

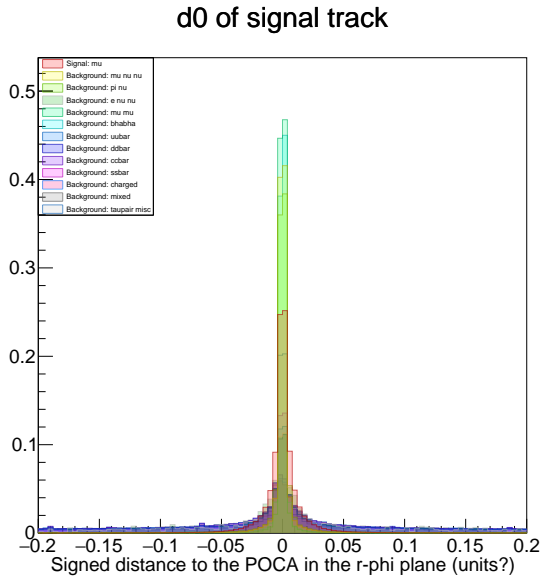Figure 7.13: Number of reconstructed tracks.
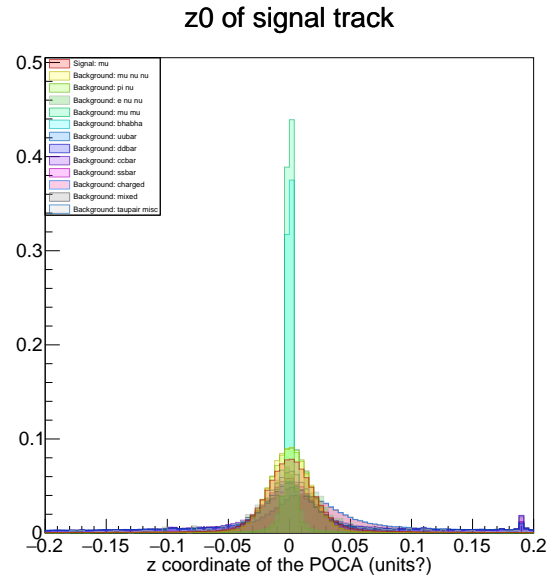

Figure 7.14: Point-of-closest approach - D0.

Figure 7.15: Point-of-closest approach - Z0.

variable	description
$p_{\text{signal}}^{\text{CM}}$	CM momentum of signal track
$p_{\text{tag}}^{\text{CM}}$	CM momentum of tag track
$p_{t\ \text{signal}}$	Transverse momentum of signal track
$p_{t\ \text{tag}}$	Transverse momentum of tag track
$\cos\theta_{\text{signal}}$	Cosine of polar angle of signal track
$\cos\theta_{\text{tag}}$	Cosine of polar angle of tag track
$E_{\text{total}}^{\text{CM}}$	Center-of-mass energy of total system
$|\text{thrust}_{\text{signal}}|$	Magnitude of signal thrust vector*
$\mu\text{-ID}_{\text{signal}}$	μ PID of signal track
$\mu\text{-ID}_{\text{tag}}$	μ PID of tag track
E_γ	Energy of signal photon
$\cos\theta_\gamma$	Cosine of polar angle of signal photon
$\cos\theta_{\text{signal}-\gamma}^{\text{CM}}$	Cosine of angle between signal track and signal photon in CM
$E_{\text{sum}}^{\text{CM}}$	Center-of-mass energy of photons and tracks
$\cos\theta_{\text{signal, tag track}}^{\text{CM}}$	Opening angle between signal and tag track
$\cos\theta_{\text{H}}$	Cosine of the helicity angle
n_{tracks}	Number of charged tracks
K-ID$_{\text{signal}}$	K PID of signal track
π-ID$_{\text{signal}}$	π PID of signal track
e-ID$_{\text{signal}}$	e PID of signal track
clusterTiming	timing of this cluster
cosTBz	Cosine of the angle between the thrust axis of the signal track and z-axis
cosTBTO	Cosine of the angle between the thrust axis of the signal track and ???
thrustRoe	Thrust roe
tauCMtheta	Center-of-mass frame polar angle of reconstructed signal τ
$E_{\text{neutral ECL}}$	Energy in ECL clusters deposited by neutral particles
sigD0	D0 of signal track
sigZ0	Z0 of signal track
cc1	Cleo cone 1
cc2	Cleo cone 2
cc4	Cleo cone 4
cc5	Cleo cone 5
cc6	Cleo cone 6
cc7	Cleo cone 7
cc9	Cleo cone 9
hso00	Hso(0,0)
hso01	Hso(0,1)
hso02	Hso(0,2)
hso03	Hso(0,3)
hso20	Hso(2,0)
hoo1	R(1)
hoo2	R(2)
hoo4	R(4)

Table 7.3: Selection criteria

8 | Signal region analysis

After selection, we are left with a far reduced number of background events and a non-zero amount of expected signal events. We can now analyse events within the signal region. This is a region in ΔE vs M_{inv} space, with $-0.4 < \Delta E < 0.2$, and $1.65 < M_{\text{inv}} < 1.85$.

8.1 Muon mode

ΔE and M_{inv} are fit with asymmetric Gaussians to determine mean and σ. We find mean $\Delta E \approx 47\,\text{MeV}$ and $M_{\text{inv}} \approx 1.79\,\text{GeV/c}^2$; this is consistent with our expections of $\Delta E \sim 0$ and $M_{\text{inv}} \sim m_\tau$. Signal variable resolutions are $\sigma_{\Delta E}^{\text{high/low}} = (5 \pm 5)\,\text{MeV}/(10 \pm 10)\,\text{MeV}$ and $\sigma_{M_{\text{inv}}}^{\text{high/low}} = (5 \pm 5)\,\text{GeV/c}^2/(10 \pm 10)\,\text{GeV/c}^2$.
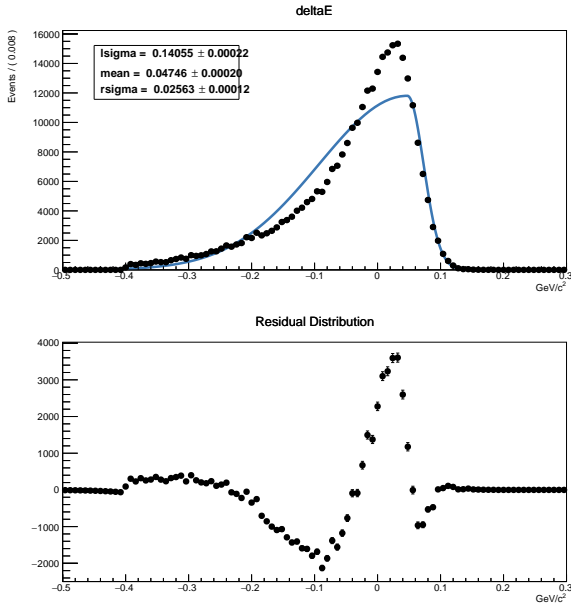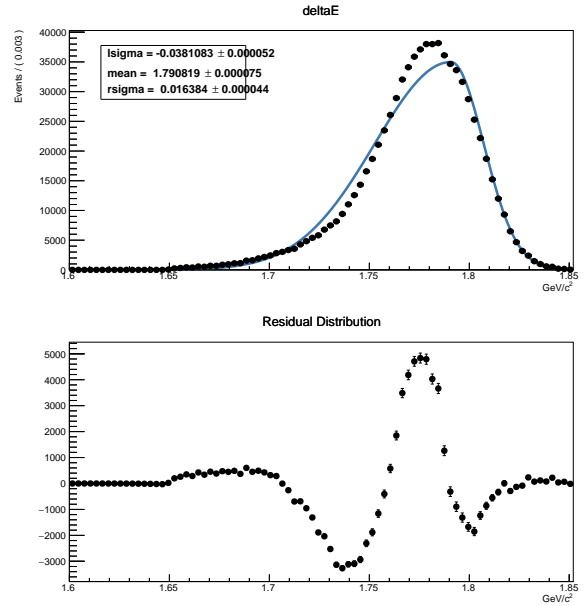

Figure 8.1: ΔE asymmetric Gaussian fit

Figure 8.2: M_{inv} asymmetric Gaussian fit

After accounting for the correlation between signal region variables we produce 2σ and 3σ bounds on event selection. The signal region is shown in Figure XXXX, with signal and background MC distributions overlayed.

In the entire signal region we find 6 $\tau \to \mu\gamma$ events, and 163 $\tau \to \mu\nu\nu$ events, 40 $\tau \to \pi\nu$ events, 15 $e^+e^- \to \mu^+\mu^-(\gamma)$ events, 9 $e^+e^- \to u^+u^-$ events, 3 $e^+e^- \to d^+d^-$ events, and 3
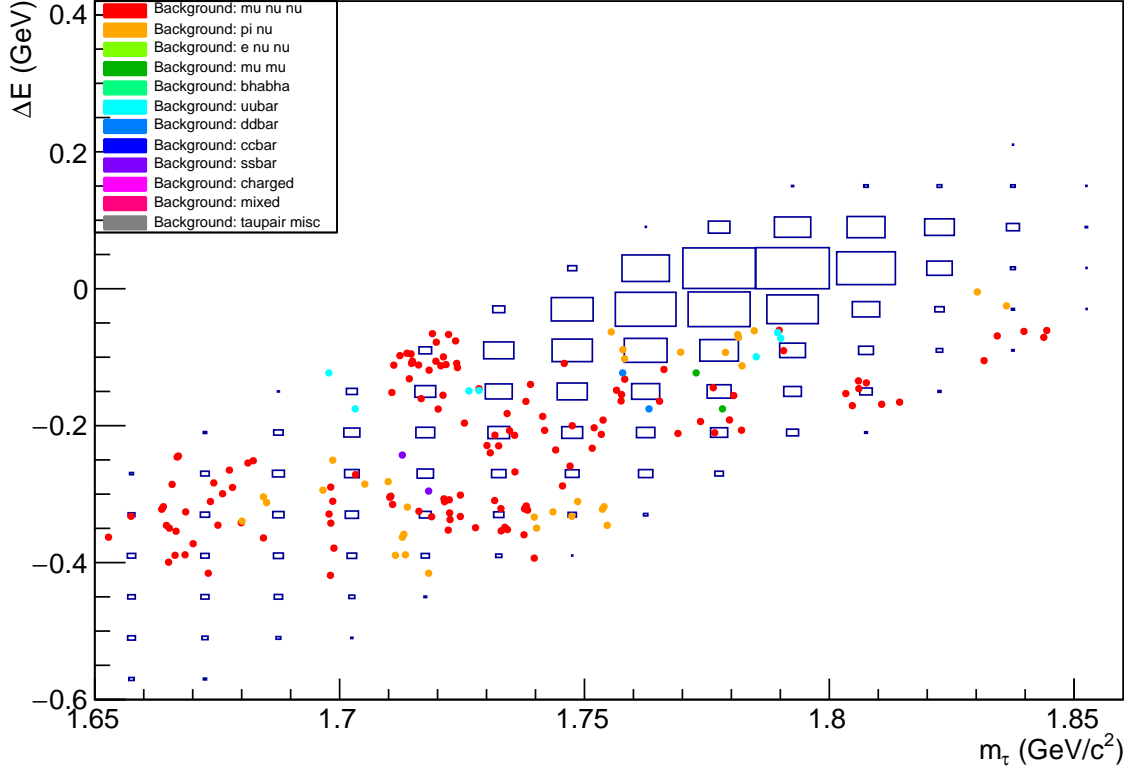
Figure 8.3: $M_{\mathrm{inv}} - \Delta E$ signal region for $\tau \to \mu\gamma$. Shaded boxes is the event distribution for $\tau \to \mu\gamma$; dots are unscaled background events coloured according to the legend.

$e^+e^- \to s^+s^-$ events. This totals to 6 signal events and 232 ± 1 background events. Within the 2σ ellipse, we find 5 ± 5 signal events and 13 ± 13 background events. Of these background events we have 10 ± 10 $e^+e^- \to u^+u^-$ events and 3 ± 3 $\tau \to \mu\nu\nu$ events.

To determine an upper limit on the branching fraction of $\tau \to \mu\gamma$, we select a smaller section of the total signal region. We then plot a one-dimensional projection of events in this region, shown in Figure8.5. A PDF is fit to the background events, and we find areas of overlap with the signal distribution. SOMETHING SOMETHING an upper limit on the branching fraction is found as

$$\mathcal{B}(\tau \to \mu\gamma) < XXXXXX \tag{8.1}$$

8.2   New bounds and new physics

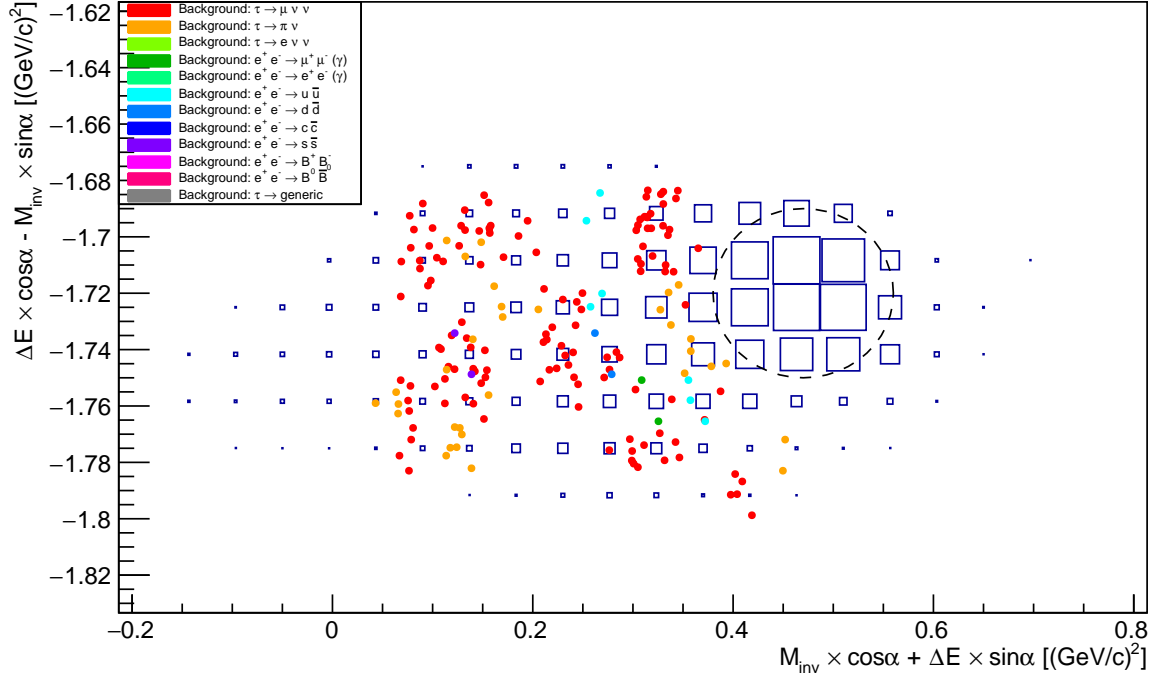COMMENT ON THE MASS SCALE THESE PROBE IN NEW PHYSICS SCENARIOS (2HDM, Babar's looked at it)

Figure 8.4: Rotated $M_{\mathrm{inv}} - \Delta E$ signal region for $\tau \to \mu\gamma$. Shaded boxes is the event distribution for $\tau \to \mu\gamma$; dots are unscaled background events coloured according to the legend. $\alpha = 72°$. The dotted ellipse represents a region of 3.811(59) signal events and 0 background events.
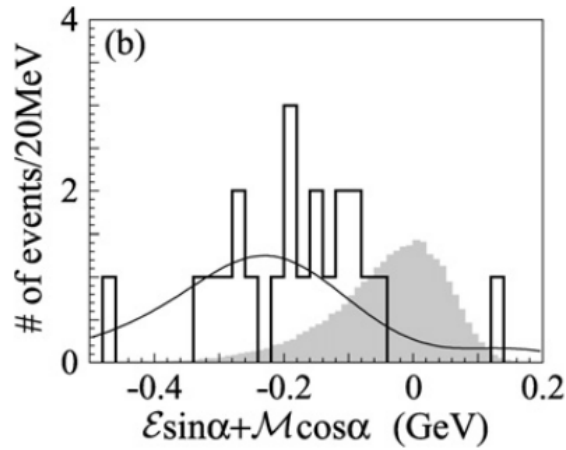

Figure 8.5: One-dimensional projection of signal region plot. $\mathcal{E}\sin\alpha + \mathcal{M}\cos\alpha$ against scaled number of events in selection region. $\mathcal{E} \equiv \Delta E - \Delta E^{(0)}$ and $\mathcal{M} \equiv M_{\mathrm{inv}} - M_{\mathrm{inv}}^{(0)}$ where $\Delta E^{(0)}$ and $M_{\mathrm{inv}}^{(0)}$ are the most probable values. **NOT MY PLOT**

9 | Conclusion

We analysed an $1\,\mathrm{ab}^{-1}$ sample of Belle II MC comprising a range of common backgrounds which were reconstructed and simulated in basf2 using SuperKEKB geometries. We were able to provide estimates on possible improves to $\tau \to \ell\gamma$ branching fractions;

$$\mathcal{B}\nabla(\tau \to \mu\gamma) = XXX \pm XXX, \tag{9.1}$$

$$\mathcal{B}\nabla(\tau \to e\gamma) = XXX \pm XXX. \tag{9.2}$$

These are consistent with predictions made by the Belle II Collaboration [REF].

9.1 Future improvements

This analysis could be extended to run over actual Belle data. The full dataset of $1\,\mathrm{fb}^{-1}$ is available, and it is possible to this data to be reconstructed in basf2. Previous searches at Belle have used $200\,\mathrm{fb}^{-1}$ (???) and $515\,\mathrm{fb}^{-1}$; a search over higher luminosity would likely improve the accuracy of some τ LFV branching fractions.

Could use TMVA (at all? or more? maybe I'll do some). Could use more topological variables such as some of the ones I recently added. Could reconstruct electrons better. Could investigate beam backgrounds and ways to remove it.

9.2 Working group status

This study of $\tau \to \ell\gamma$ has been performed as part of the *Tau and low multiplicity* working group at Belle II. TALK ABOUT THE FIVE GOLDEN OBSERVABLES?

9.3 Expected impact on Belle II

This research is the first to investigate τ LFV at Belle II using full analysis of MC. UHHHH

Bibliography

[1] T. Ohshima. Study of LFV in tau decay at Belle. *Nucl. Phys. Proc. Suppl.*, 169:174–185, 2007. [,174(2007)].

[2] Passemar.