Sensitivity Study of the Charged Lepton Flavour Violating Mode $B^+ \to K^+\tau^+e^-$ at Belle II.

Trevor Shillington

Department of Physics

McGill University

Montreal, Quebec, Canada

August, 2021

A thesis submitted to McGill University in partial fulfillment of the

requirements of the degree of Master of Science

Abstract

Belle II is a B factory experiment for the SuperKEKB electron-positron collider located at the KEK laboratory in Tsukuba, Japan, operating near the Upsilon(4S) resonance, at an energy of 10.58 GeV. In this thesis we outline our analysis searching for the ultra-rare charged lepton flavour violating (CLFV) decay $B^+ \to K^+\tau^+e^-$. This decay is far below experimental sensitivity if we assume the decay rate predicted by the Standard Model. However, many extensions of the Standard Model, specifically those attempting to incorporate the recent "B physics anomalies", predict much larger branching fractions which are potentially within the reach of experiments. Discovery of this mode would be explicit evidence of physics beyond the Standard Model, while a null result would allow us to place strict constraints on these models. A previous search was done at BaBar in 2012, setting a 90% CL upper limit branching fraction of a few x 10^{-5}. The much larger integrated luminosity dataset at Belle II can be exploited to improve the analysis sensitivity by at least an order of magnitude. A sensitivity study is performed based on a dataset of 200 fb^{-1}, and a data/MC comparison is done.

Abrégé

Belle II est une expérience ‹‹ usine à B ›› pour le collisionneur électron-positron SuperKEKB situé au laboratoire KEK à Tsukuba, au Japon, fonctionnant près de la résonance Upsilon(4S), à une énergie de 10,58 GeV. Dans cette thèse, nous décrivons notre analyse sur la recherche de la désintégration $B^+ \to K^+\tau^+e^-$ ultra-rare qui viole la saveur des leptons chargés (CLFV). Cette décroissance est bien au-dessous de la sensibilité expérimentale si nous supposons que le taux de décroissance prédite par le Modèle Standard. Cependant, de nombreuses extensions du Modèle Standard, en particulier celles qui tentent d'intégrer les récentes ‹‹ anomalies de la physique B ›› , prédisent des fractions de branchement beaucoup plus importantes qui sont potentiellement à la portée des expériences. La découverte de ce mode serait une preuve explicite d'une physique au-delà du Modèle Standard, tandis qu'un résultat nul permettrait d'imposer des contraintes strictes sur ces modèles. Une recherche précédente a été effectuée à BaBar en 2012, fixant une fraction de ramification limite supérieure de 90% CL de quelques x 10^{-5}. Le grand ensemble de données de luminosité intégré à Belle II peut être exploité pour améliorer la sensibilité de l'analyse d'au moins un ordre de grandeur. Une étude de sensibilité est réalisée sur la base d'un ensemble de données de 200 fb^{-1}, et une comparaison données/MC est effectuée.

Acknowledgements

I would like to thank the entire Belle II collaboration for all of the hard work that has been poured into it over the last decade. Over a thousand people have contributed to the experiment in some way and all of that work has had some impact on the work in this thesis.

I want to give special thanks to my supervisor, Steven Robertson, who has given me so much support throughout my Master's program. His guidance and expertise was paramount to my success, and I appreciate all of his patience and wisdom along the way. I also want to thank the McGill Belle II group, present and past, for all of the advice and assistance over the years. In no particular order, thank you to Andreas Warburton, Andrea Fodor, Hannah Wakeling, Kyle Amirie, Rob Seddon, Ryan MacGibbon, and Racha Cheaib.

I also want to give a huge thank you to my parents for their unconditional love and support throughout my education. It has been long and arduous, and I could not have done it without you. Additionally, I thank my brother, Brett, and all of my friends who have supported me. Last but not least, I would like to thank my partner, Olivia, for her love and support, and for translating the abstract.

Preface

Unless otherwise stated, the work in this thesis has been done solely by the author. All sections were written by the author and represent original content. All plots and tables, unless cited, were produced by the author.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Belle II is a particle physics experiment operating out of Tsukuba, Japan. It detects collisions at an electron-positron collider, operating at a specific energy to induce the production of a particle called a B meson. The in-depth study of the B meson and its decay channels are the primary focus of the Belle II experiment, and through them the goal is to understand in greater detail the Standard Model (SM) and probe for New Physics (NP). The SM has had great success in the past explaining particles and their interactions, to a very high degree of precision in many cases, but is ultimately incomplete. If we wish to move Beyond the Standard Model (BSM), to a theory that more accurately describes nature, we must probe the predictions of the SM in hopes of finding more hints that can guide us to new theories. One such SM prediction is the focus of this thesis.

The decay modes $B^{\pm} \to K^{\pm}\tau\ell$ ($\ell = e, \mu$) are theoretically allowed in the SM[1], but are so highly suppressed that they are not statistically observable at any conceivable B factory experiment. These are charged lepton flavour violating (CLFV) decay modes, something that has never been observed in any decay mode. However, many extensions to the SM predict enhancements of CLFV. In the recent years, various experimental results in semi-leptonic B physics have given hints of non-conservation of lepton universality (meaning the different flavours of leptons could have different coupling strengths, either

[1] If and only if non-zero neutrino masses are included.

to the known electroweak gauge bosons $W^{\pm}$ and Z^0, or to new particles). These results are particularly provocative for the prospects of CLFV, since there is no known way to include lepton flavour non-universality in a model without inducing CLFV [1]. Some models predict levels that may even be within reach of Belle II.

This thesis will outline the analysis of $B^+ \to K^+\tau^+e^-$ (and its charged conjugate) at Belle II. Belle II is still in its early stages of data collection, so the primary focus of this thesis will be a sensitivity study. In Chapter 2, there is theoretical overview of the Standard Model, charged lepton flavour violation, and an outline for the motivation for our search. A detailed description of Belle II and the SuperKEKB accelerator is given in Chapter 3. Chapter 4 will present the analysis, and Chapter 5 will present the results. A final conclusion will follow in Chapter 6.

Chapter 2

Background

2.1 The Standard Model

Particle physics today is understood through the Standard Model (SM). It describes the fundamental particles and their interactions via three of the four fundamental forces: electromagnetism, the weak force, and the strong force. The fourth fundamental force is gravity, and it is not incorporated into the SM. The Standard Model has been developed over the 20^{th} century, starting with the quantum revolution in the 1920s. Over the years, many new particles were theorized and discovered through experiment. In the latter half of the 20^{th} century, particles were described as quantum excitations in fields rather than independent structures. This was formalized by quantum electrodynamics (electromagnetism), quantum flavourdynamics (weak interactions) and quantum chromodynamics (strong interactions). Yang-Mills theory lies at the heart of the Standard Model, describing elementary particles with non-abelian Lie groups. Yang-Mills theory ultimately gave rise to the connection between electromagnetism and the weak force (unifying them as the electroweak force), and the strong force via the symmetry group $SU(3)_c$ x $SU(2)_L$ x $U(1)_Y$.

There are seventeen elementary particles in the SM today (See Figure 2.1). Elementary particles are truly non-composite particles. They can be grouped into two general

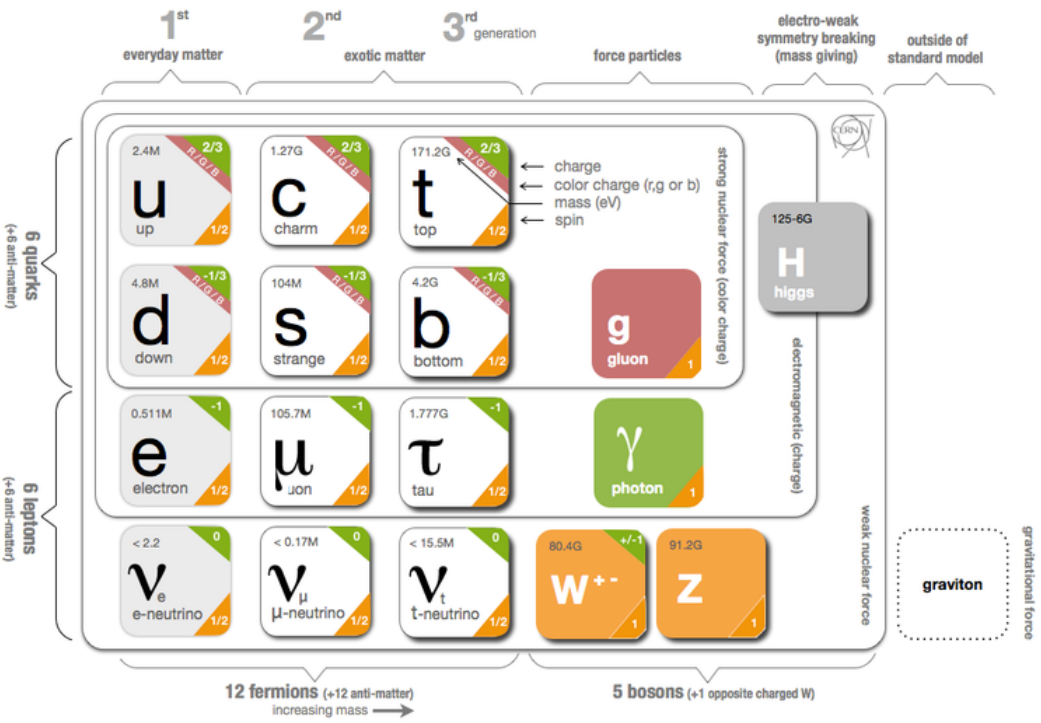

Figure 2.1: The elementary particles of the Standard Model.(Credit: CERN)

classes of particles: bosons and fermions. The fundamental difference between bosons and fermions is their spin, with bosons carrying integer spin (0, 1, 2, ...) and fermions carrying half-integer spin (1/2, 3/2, 5/2, ...). The bosons are comprised of the photon, the gluons, the $W^{\pm}$ bosons, the Z boson, and the Higgs boson. The spin 1 bosons, known as the gauge bosons, also happen to be the force carriers. The photon carries the electromagnetic force, the W and Z bosons carry the weak force, and gluons carry the strong force. The mass of the force carrier will determine the distance over which the force has effect (massless = infinitely far, the more massive = shorter). On the other side of it we have the fermions. Fermions are further broken into two sectors, the quark sector (strong interactions) and the lepton sector (no strong interactions). Each of these sectors have 3 generations, and each generation has two members. For example, in the first generation

of quarks you have the up and down quark, and in the first generation of leptons you have the electron and the electron neutrino.

2.1.1 Quark Sector

There are 6 quark flavours across 3 generations. The up and down quark make up the first generation, with the up quark having +2/3 charge and the down quark having -1/3 charge. In the second generation you have the charm and the strange quarks, and in the third generation there are the top and bottom quarks. Like the up quark, the charm and top quarks are +2/3 charge, and the strange and bottom quarks have -1/3 charge like the down quark.

In addition to these 6 quarks, each of them have an anti-quark companion which has the same mass but opposite quantum numbers (e.g. charge). On top of that, all of the quarks (anti-quarks) come in one of three colours: red (anti-red), green (anti-green) and blue (anti-blue). Quarks are not literally coloured, but these are the names given to the charges of the strong force. Much like electric charge is the charge of the electromagnetic force, colour charge is the charge of the strong force. Quarks interact via the strong force by exchanging colour-carrying gluons with other quarks.

Quarks can combine to make composite particles called hadrons. There are various ways to combine quarks but they must follow some rules. For one, all composite particles must be "colourless". You can achieve a colourless particle by combining a quark and an anti-quark of like-colour (e.g. a red quark and an anti-red anti-quark). These types of hadrons are called mesons, and they play a central role in high energy particle physics. The most common mesons are the pion, the kaon, the D meson, and the B meson. Each of these are characterized by their quark composition, with the pion just having up and down quarks, the kaon containing a strange quark, the D meson containing a charm quark, and the B meson containing a bottom quark. Specifically, B mesons are composed of a bottom quark and either an up (B^+), down (B^0), strange (B_s^0) or charm (B_c^+) anti-quark, each with an appropriate anti-particle. Another way to get a colourless

composite particle is to combine three quarks all of different colours (e.g. a red quark, a green quark and a blue quark). These are called baryons, with protons (uud) and neutrons (udd) being the most well-known examples. There are other, more exotic configurations that are possible, such as tetraquarks (two quarks and two anti-quarks) and pentaquarks (four quarks and an anti-quark). Quarks have never been observed alone due to colour confinement, and so always exist in hadronic form with other quarks.

Although quarks can interact via any of the four fundamental forces, they can only change flavour via the weak interaction. The Cabibbo-Kobayashi-Maskawa (CKM) [2] matrix is a 3x3 unitary matrix that contains the mixing strength of the flavour-changing weak interaction. Each element of the matrix, V_{ij}, provides information on the probability of transitioning from quark i to quark j. The CKM matrix only applies to charged current flavour-changing processes, (i.e. mediated by a $W^{\pm}$ boson), meaning that quarks have to change to a flavour of opposite charge. An example of this is $B^- \to D^0\pi^-$, where the negatively charged bottom quark in the B meson changes flavour to a positively charged charm quark via the W^-, creating the D meson. At tree level, flavour-changing neutral currents (FCNC) are not possible. The Z^0 boson only couples to quarks of the same charge, so the matrix element contains no mixing. However, higher order Feynman diagrams (e.g. box or penguin diagrams) allow for FCNC processes. An example of a higher order FCNC process is $B \to K\ell^+\ell^-$. The presence of a complex mixing phase in the CKM matrix allows for CP violation in the Standard Model, and direct evidence of this was found in B meson decays at both Belle [3] and BaBar [4] in 2001.

The B mesons are the heaviest class of mesons. You might expect that some meson containing a top quark, being the most massive quark, would be the most massive meson. However, the top quark is so massive that it exceeds the mass of W and Z bosons. Due to this fact, it can decay directly into a W or Z along with another quark. Due to this it has an extremely short lifetime (10^{-25} seconds). The time it takes to hadronize (form a hadron with other quarks) is roughly 20 times this, and therefore the top quark does not form bound states. The next heaviest quark is the bottom quark, and it is less massive than the

W / Z bosons. It does allow for hadronization and therefore the heaviest mesons available to study are the B mesons. This is the particle that is studied extensively at so-called "B factories", like Belle II, which are experiments that maximize the production of B mesons in collisions.

2.1.2 Lepton Sector

In the lepton sector there are also six flavours and three generations. There are three charged leptons, each with an associated neutral neutrino. The charged leptons consist of the electron, muon, and tau. These all have a charge of -1 (their anti-particles have a charge of +1, of course), and the SM assumes they differ only by mass, also known as Lepton Flavour Universality (LFU) (i.e. they couple to the electroweak gauge bosons with the same strength). Leptons do not interact with the strong force, but can interact with the other three fundamental forces.

The neutrinos are very low mass, and were originally assumed to have zero mass. Because of their small mass and neutral electric charge, the neutrinos rarely interact with other matter, and can only do so via the weak force. This makes them difficult (virtually impossible) to detect directly at experiments like Belle II. Only experiments specifically designed to detect neutrinos (which requires large mass, low backgrounds, and a strong neutrino flux) can do so, and even then the efficiency is low. Similar to the quarks, the neutrinos have a weak eigenstate that differs from their mass eigenstates, allowing them to mix. The Pontecorvo-Maki-Nakagawa-Sakata (PMNS) [5] matrix is a 3x3 unitary mixing matrix that contains the information of this mixing. Leptons, and thus neutrinos, are allowed to exist in isolation, and a free neutrino will oscillate between flavours, as a function of its energy and distance travelled. Since the neutrinos are allowed to change flavours under the PMNS matrix, they exhibit Lepton Flavour Violation (LFV). LFV is the non-conservation of individual lepton flavour numbers in interactions.

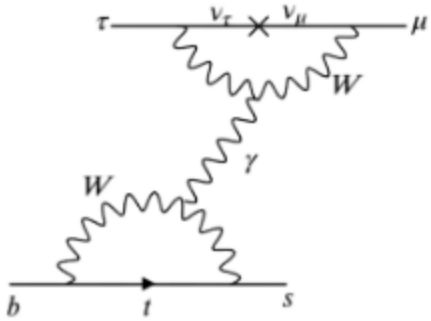

Figure 2.2: Allowed Standard Model Feynman diagram of a Charged Lepton Flavour Violating decay mode. It is highly suppressed and not statistically possible to detect at experiments.

The charged leptons do not mix and can only exhibit LFV in extremely suppressed higher-order (above tree level) Feynman diagrams, such as shown in Figure 2.2. Charged Lepton Flavour Violation (CLFV) will be discussed in detail in the following section.

2.1.3 Charged Lepton Flavour Violation

As stated in the previous section, we already know that lepton flavour violation happens directly via neutrino oscillation. This opens the question to LFV with charged leptons. Given neutrino oscillations, it is possible to have CLFV in the Standard Model (See Figure 2.2 for the Feynman diagram of an allowed SM process that exhibits CLFV). However, this process is suppressed by a factor of the neutrino mass squared over the W boson mass squared, which is $\mathcal{O}(10^{-48})$. Due to this, it is so highly suppressed that it is statistically not possible to observe experimentally. Therefore, any discovery of a CLFV mode is clear evidence of physics beyond the standard model (BSM).

2.2 Motivation

2.2.1 B-physics Anomalies

In the recent years there have been a handful of anomalies; disagreements between the Standard Model and experimental results at the 2-4σ level. Three of the most prominent anomalies are known as $R_{D^{(*)}}$, $R_{K^{(*)}}$, and P_5'. $R_{D^{(*)}}$ and $R_{K^{(*)}}$ are the ratios of branching fractions of semi-leptonic B decays, defined as

$$R_{D^{(*)}} = \frac{\Gamma(B \to D^{(*)}\tau\overline{\nu})}{\Gamma(B \to D^{(*)}\ell\overline{\nu})}, \tag{2.1}$$

with $\ell = e, \mu$, and

$$R_{K^{(*)}} = \frac{\Gamma(B \to K^{(*)}\mu^+\mu^-)}{\Gamma(B \to K^{(*)}e^+e^-)}, \tag{2.2}$$

where Γ indicates the branching fraction. The decays in $R_{D^{(*)}}$ are tree-level $b \to c\ell\nu$ transitions, while those in $R_{K^{(*)}}$ are loop-level FCNC $b \to s\ell^+\ell^-$ transitions. Experimental measurements of $R_{D^{(*)}}$ have been performed at BaBar [6], Belle [7], and LHCb [8] [9] [10], with all three experiments finding values consistent with one another, and generally exceeding the SM predictions (See Figure 2.3). The combined result (as of 2019) gives a 3.08σ tension with the SM [11].

The decays represented in $R_{K^{(*)}}$ are rarer and governed by FCNC processes at loop-level. These are also sensitive to NP as new heavier particles could contribute at sizeable levels [12]. The measured values from LHCb [13] [14] are separated based on bins of q^2, the dimuon invariant mass squared, and they all deviated from the SM in the 2-3σ range. The measurements from Belle [15] and BaBar [16] were in better agreement with the SM. A related $b \to s\ell^+\ell^-$ decay, $B_s^0 \to \phi\mu^+\mu^-$, has also shown a discrepancy from the SM of more than 3σ [17]. The P_5' anomaly is also related to these, as it comes from the decay of $B \to K^*\ell^+\ell^-$. The angular distribution is of interest for NP, and can be parameterized in terms of eight angular observables [12]. One such observable is P_5', and the experimental findings of LHCb [18] show a 3.4σ tension with the SM, and Belle [19] and ATLAS [20]

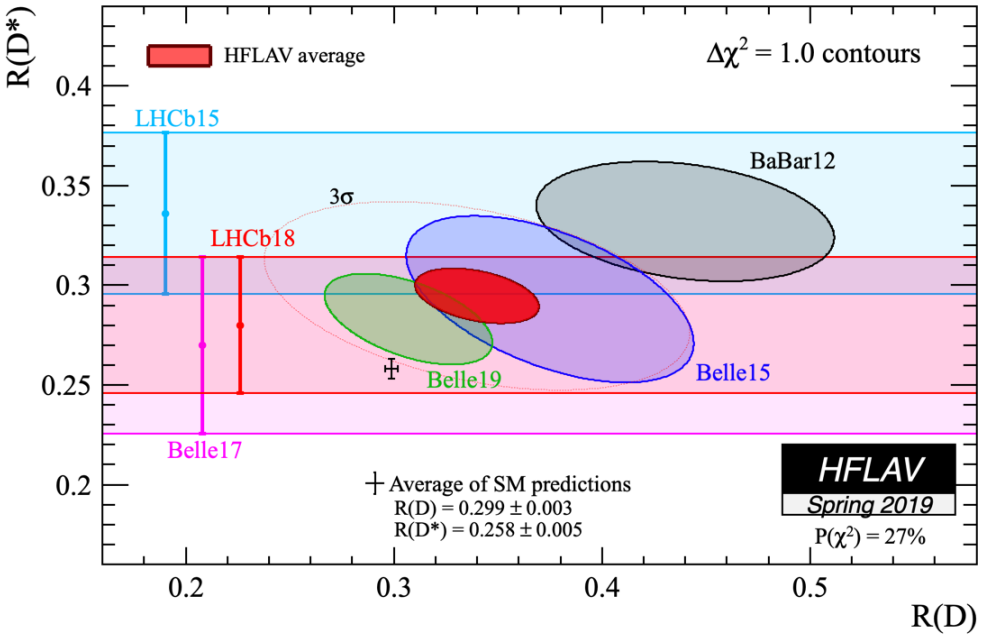

Figure 2.3: Plot of measurements of R_D and R_{D^*} from various experiments and the SM prediction. As of 2019, there is a 3.08σ tension of the combined experimental results and the SM. Taken from [11]

.

have found similar results, although with larger uncertainties [21] (See Figure 2.4). CMS has found results that are compatible with both the SM and the other experimental measurements [22]. The results of $R_{D^{(*)}}$ and $R_{K^{(*)}}$ both point towards violation of LFU, and due to this some BSM models are able to incorporate both simultaneously. Confirmation of either of these would be clear evidence of NP. The P_5' anomaly adds interest to the NP that is related to $b \to s\ell^+\ell^-$ transitions, and further studies will help understand the B anomalies and where NP might possibly come into play.

2.2.2 New Physics and Predictions

Along with these B anomalies have come many models that try to incorporate them. The anomalies which infer violation of LFU would bring about NP with couplings that have different strengths to the different lepton generations. As mentioned previously, and

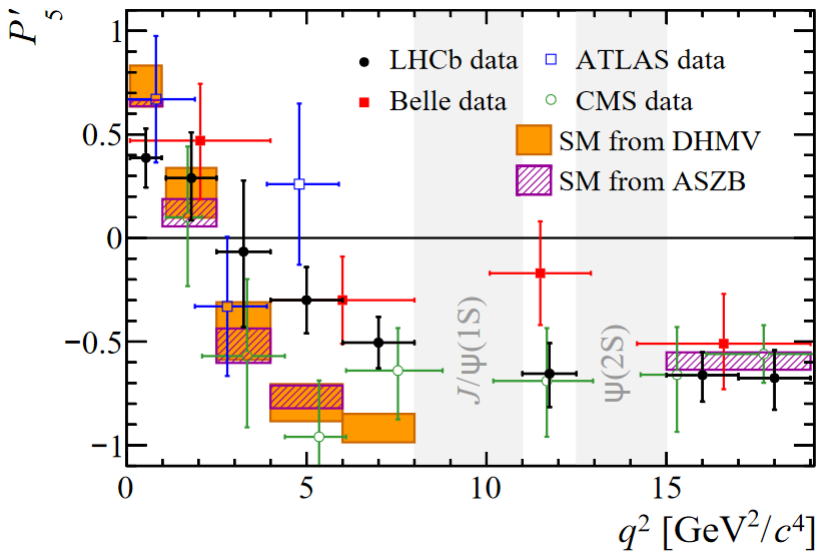

Figure 2.4: Angular observable P'_5 vs q^2 for $B \to K^*\ell^+\ell^-$ decays. LHCb finds tension with the SM predictions at the 3.4σ level, coming from the dimuon channel, while other experiments find similar results but with larger uncertainties. Taken from [21].

explicitly stated by Glashow et al. in 2014 [1], "any departure from lepton universality is necessarily associated with the violation of lepton flavour conservation". Many of the proposed models involve introducing leptoquark states, or new gauge bosons (Z' and W') [23]. Leptoquarks are new bosons that would couple to both quarks and leptons. For example, Angelescu et al. [24] show that a weak single vector leptoquark, U_1, could provide a solution to the anomalies if $m_{\mathrm{LQ}} \approx$ 1-2 TeV. They also predict BR($B \to K\tau\mu$) $\geq$ few x 10^{-7}. I note that many such models make predictions on BR($B \to K\tau\mu$) instead of BR($B \to K\tau e$) due to the higher contribution of NP to the second and third generation, but in general the enhancement of CLFV applies to $B \to K\tau e$ as well.

Various other leptoquark models exist, such as the one explored by Bečirević et al. (2016) [23] which is a triplet of vector leptoquarks, and Bečirević et al. (2018) [25] which is a two scalar leptoquark model based on SU(5) Grand Unified Theory. In these, the

authors note the correlation between the prediction of BR($B \to K\tau\mu$) and $R_{\nu\nu}$, which is also testable at Belle II and make them complementary analyses in the scope of NP.

Kumar et al. [26] looked at four models of tree-level NP that could account for the B anomalies. Three were leptoquark models and the other was a vector boson model. They agreed with the findings of Angelescu et al. that a U_1 vector leptoquark is most suitable to address the anomalies.

A three-site Pati-Salam gauge model (PS^3) has been proposed by Bordone et al. [27]. This model has the nice feature of naturally connecting the dynamical breaking of flavour symmetry to both the anomalies and the hierarchical structure of the quark and lepton mass matrices. This model is also interesting as it suppresses the decays in $B \to K\tau\mu$ when the τ has the opposite charge to the B and kaon, but highly enhances the case where the τ has the same charge, to levels within reach of Belle II. This highlights the importance of searching for the whole family of $B \to K\tau\ell$ decays. Another Pati-Salam model was looked at by Keeck and Teresi [28]. They show that a seesaw-motivated scalar leptoquark can also solve both LFU anomalies.

Guadagnoli et al. [29] use a gauged horizontal SU(2) symmetry to make predictions on $B \to K\tau\mu$ and show that BR($B \to K\tau^+\mu^-$) $\neq$BR($B \to K\tau^-\mu^+$), although which is larger depends on the value of a phase parameter.

The various predictions of these models on BR($B \to K\tau\mu$) span a wide range and can be as high as 10^{-6}. Many other models exist (e.g. see [30] [31] [32] [33] [34], to cite just a few), to explain the anomalies without explicit predictions on the decay modes we are interested in. The models mentioned here were not presented in great detail, and are just meant to show the diverse landscape of models attempting to explain the B anomalies with NP, while making explicit reference to an increase in CLFV and in particular to our family of modes, $B \to K\tau\ell$. It also demonstrates how making measurements on these decay modes can help to discriminate between models and help place stricter limits on them.

2.2.3 Previous Searches

The mode $B^+ \to K^+\tau^+e^-$ has so far only been searched for at BaBar (2012) [35]. They searched for the larger family of decay modes $B^\pm \to h^\pm\tau\ell$ ($h = K, \pi$ and $\ell = e, \mu$). The full 429 fb^{-1} BaBar data sample was used, amounting to 472 million $B\overline{B}$ pairs. They found no evidence for any of the decay modes, setting the upper limit on each branching fraction to a few times 10^{-5} at 90% confidence level (CL). Specifically for $B^+ \to K^+\tau^+e^-$, they set a branching fraction limit of < 1.5 x 10^{-5} at 90% CL. Due to the 50x increase in integrated luminosity at Belle II, it is expected that this limit can be improved by at least one order of magnitude. Similar to Belle II, BaBar was also a B factory that collided electrons and positrons at the $\Upsilon(4S)$ resonant energy. Due to this, our analysis employs the same overall strategy as used at BaBar, as will be discussed further in Chapter 4.

It is also worth mentioning that LHCb has searched for the related mode $B^+ \to K^+\mu^-\tau^+$ using B^{*0}_{s2} decays and set an upper limit on the branching fraction of < 3.9 x 10^{-5} at 90% CL [36], comparable to the findings of BaBar. As of time of writing, Belle has not published any results on these decay modes.

Chapter 3

The Belle II Experiment

The Belle II experiment is the successor to Belle (1999-2010), a Japan-based B-factory originally commissioned to probe the nature of CP violation in $\Upsilon(4S) \rightarrow B^0\overline{B^0}$ decays. It operates out of Tsukuba, Japan, at the High Energy Accelerator Research Organization (Japanese: Kō Enerugī Kasokuki Kenkyū Kikō), known simply as KEK. Belle II aims to explore the nature of the Standard Model (SM) by making precision measurements and searching for New Physics (NP). In tandem with the detector upgrade from Belle to Belle II, the collider, KEKB, was also approved for an upgrade to SuperKEKB. The Belle II detector and the SuperKEKB accelerator will be described in the following sections. Unless otherwise noted, the contents of this chapter are sourced from the Belle II Technical Design Report [37].

3.1 SuperKEKB Accelerator

SuperKEKB is the collider at which Belle II operates. It collides electrons and positrons at asymmetric energies ($E_{e^-} = 7\,GeV$, $E_{e^+} = 4\,GeV$), for a center-of-mass energy of 10.58 GeV. This energy corresponds to the mass resonance of the $\Upsilon(4S)$ meson (see Figure 3.1) which is a $b\bar{b}$ quarkonium state, and sits only 20 MeV/c^2 above the threshold for $B\overline{B}$ pair production. This will decay to a B meson pair (and nothing else, resulting in a low-momentum

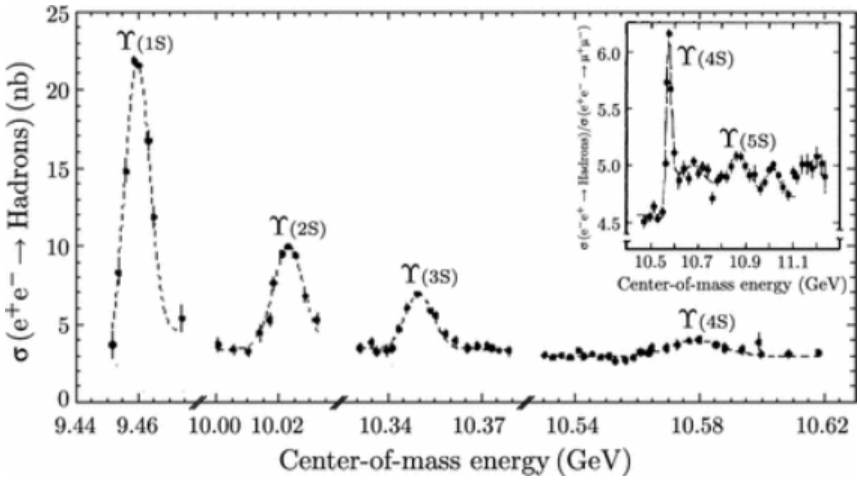

Figure 3.1: Cross-section of e^+e^- collisions, measured by CUSB and CLEO. The Υ(4S) is one resonance of a family of Υ-resonances. Figure from [38].

and clean environment) 96% of the time, with roughly half being charged B pairs and half being neutral B pairs. This is what makes it a (super) B-factory.

SuperKEKB has been upgraded from the original KEKB accelerator, with the two main improvements being a 2x increase in current and the novel 'Nano-Beam' scheme technique that allows for a 20x smaller beam spot (from 1 to 50 nm) at the collision point. This translates to a 40x higher luminosity than KEKB, a significant improvement. The peak luminosity is expected to be 8.0 x $10^{35}\, cm^{-2}s^{-1}$. See Figure 3.2 for the luminosity timeline. With a 40x higher peak luminosity, the target integrated luminosity is 50 ab^{-1} (50x Belle, 100x BaBar). This corresponds to roughly 40 billion B meson pairs.

The upgrade in luminosity came with an increased beam background, particularly from the Nano-Beam scheme. Intrabeam scattering is more prevalent in a smaller beam, causing emittance growth. The Touschek effect (loss of beam in a storage ring) also increases and shortens the beam lifetime. Compared to KEKB the positron beam energy is increased from 3.5 to 4.0 GeV to help mitigate these effects. The electron beam energy is decreased from 8.0 to 7.0 GeV to maintain the center-of-mass energy, but this also helps decrease the horizontal emittance and synchrotron radiation power [40]. See Figure 3.3 for injector beam parameters at SuperKEKB.

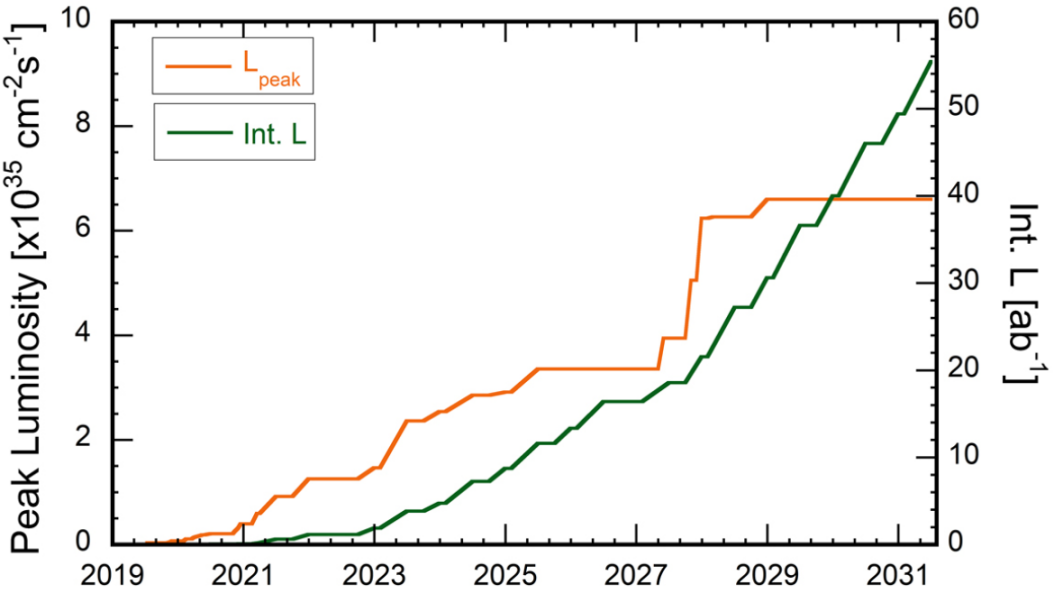

Figure 3.2: Projected luminosity timeline. Belle II is currently expected to reach the target dataset of 50 ab^{-1} by 2031. Sourced from the SuperKKEKB website [39].

	LER (e^+)	HER (e^-)	
Energy	4.000	7.007	GeV
Half crossing angle	41.5		mrad
Horizontal emittance	3.2	4.6	nm
Emittance ratio	0.27	0.25	%
Beta functions at IP (x/y)	32 / 0.27	25 / 0.30	mm
Beam currents	3.6	2.6	A
Beam-beam parameter	0.0881	0.0807	
Luminosity	8×10^{35}		$cm^{-2}s^{-1}$

Figure 3.3: Injector beam parameters at SuperKEKB. Taken from [40].

The collider itself consists of a 600 m injector linear accelerator (linac) with a 1.1 GeV positron damping ring (DR), the 7 GeV electron ring (known as the high energy ring, HER), and the 4 GeV positron ring (low energy ring, LER). The rings are side-by-side, located 10 m underground and have a circumference of 3016 m.

The electron beam is realized with a photocathode high-current radio-frequency (RF) gun, creating a bunch charge of 4 nC and an emittance of 10 $mm{\cdot}mrad$. This is then accelerated in the linac via 60 accelerating units and finally directed into the HER. To generate the positron beam, a 4 GeV electron beam is used to irradiate a 14 mm thick tungsten target located in the mid-linac region. The positrons are sent through a Flux Concentrator (FC) which uses a strong magnetic field to focus the beam before sending it into the linac. The beam is then sent through the 1.1 GeV damping ring to reduce the emittance. Finally, it is sent back into the linac and into the LER.

The HER and LER storage rings are horizontally separated and consist of four arcs and four 200 m long straight sections (see Figure 3.4). The rings are each comprised of many connected beam pipes with vacuum and cooling systems, surrounded by a multitude of various dipole, quadrupole and sextupole magnet systems to manipulate and guide the beams. RF cavities (accelerating sections) and wigglers (magnetic structures deflecting the beam in alternate directions but with zero overall net deflection, which reduce emittance and damping time) are installed in the straight sections. The beam pipes are constructed of either aluminum alloy or copper. Aluminum is cheaper and easier to manufacture, and when coated with TiN can perform similarly to copper in terms of suppressing electron cloud densities. However, in areas where synchrotron radiation (SR) is high (e.g. the wiggler sections) or close to the Belle II detector, copper is used. The majority of the beam pipes are of the antechamber-type, in which the beam pipe has wing-like structures extending on the horizontal axis. These antechambers are used to facilitate the high-power SR, specifically in the wiggler sections. In the LER arc sections, one of the antechambers house a strip-type nonevaporable getter (NEG) vacuum pump to achieve the target pressure of $\mathcal{O}(10^{-7}\ Pa)$ [41]. In the HER, the pumps are also NEGs but are lumped-type, located roughly every 3 m along the beam pipes. Auxiliary pumps are also located every 10 m along the ring.

At the interaction region (IR) of the Belle II detector, the two beams cross horizontally at an angle of 83 $mrad$. The beam pipe in the IR is double-walled and constructed of

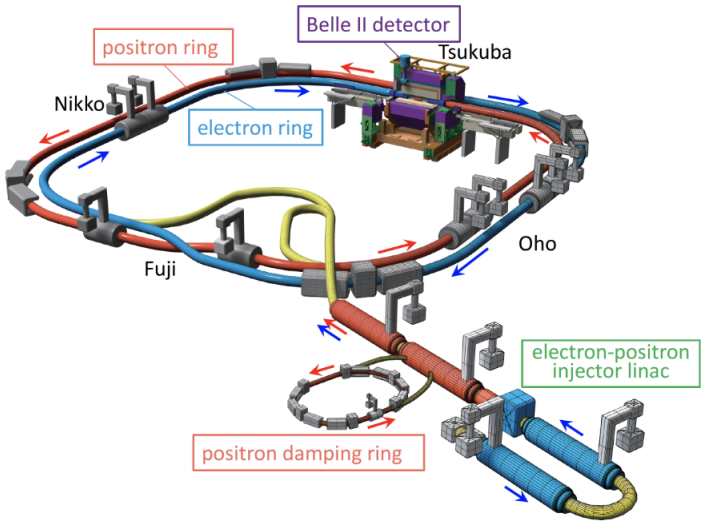

Figure 3.4: Schematic of SuperKEKB. The four straight sections are Oho, Fuji, Nikko, and Tsukuba. The Belle II detector is located at the Tsukuba region. Figure from [40].

beryllium in order to reduce multiple scattering and energy loss. The interior of the Be beam pipe is coated in gold and there is a $1\,mm$ layer of paraffin to act as a coolant between the double-wall. A final-focus superconducting magnet system (QCS) is used to realize the Nano-Beam scheme. It consists of 8 main quadrupole magnets, with 43 corrector magnets and 4 compensation solenoid coils. The corrector magnets fix any misalignments of the quadrupole magnets, tweak the beam orbit, and improve the dynamic aperture. The compensation solenoid coils are used to counter the Belle II detector's solenoid field such that each side of the interaction point (IP) has $\int B_z ds = 0$.

3.2 The Belle II Detector

The Belle II detector is cylindrical in shape and placed horizontally, surrounding the IR and covering most of the solid angle. The chosen components of Belle II satisfy the needs of a super B factory, which include excellent vertex resolution, high reconstruction efficiencies of charged particles, very good momentum resolution, precise photon energy

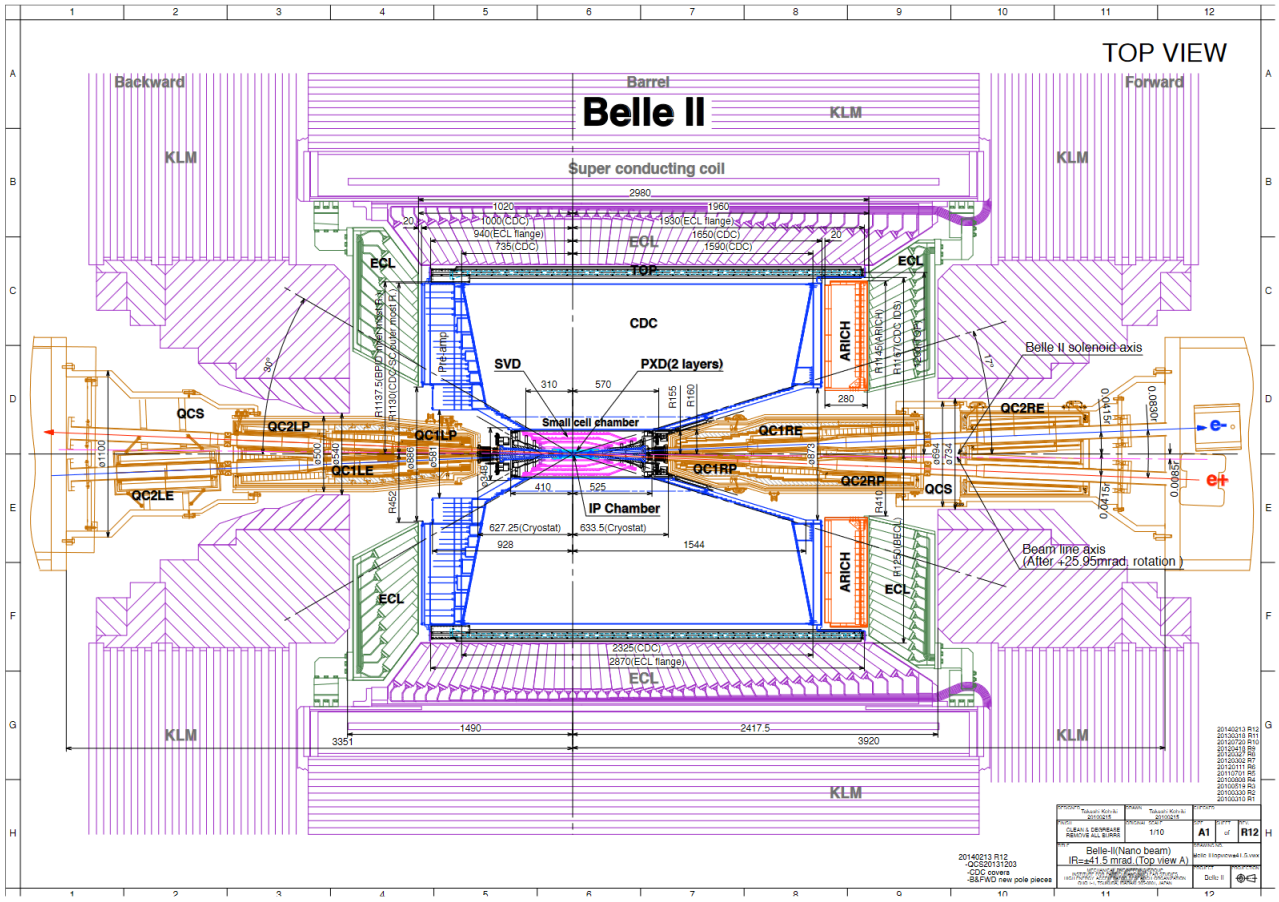

Figure 3.5: Top view schematic of the Belle II detector. Figure from [42].

measurements, high particle identification efficiencies, large coverage, and a trigger system that is fast, efficient, and accompanied by a data acquisition system that can handle large quantities of data. See Figure 3.5 for a schematic of the detector and its components. Each sub-component will be discussed in the following subsections, starting from the inner detector and moving outwards.

3.2.1 VXD

The innermost subdetector at Belle II is the vertex detector (VXD). The two components of the VXD are the Pixel Detector (PXD) and the Silicon Vertex Detector (SVD), which combine for six layers around the Be beam pipe. The primary purpose of the VXD is to

record trajectories of particles and enable reconstruction of the vertices from which particles originated. Excellent vertex resolution (roughly $50\,\mu m$) is required for a B factory. Both the PXD and SVD employ a cooling system using CO_2 circulating in thin pipes to remove the heat dissipated by the front-end readout chips and a continuous stream of dry Nitrogen across the whole VXD volume prevents condensation. The following subsections will describe the PXD and SVD in detail.

PXD

The Pixel Detector (PXD) is used for the precision vertex reconstruction of B-meson decays. The Be beam pipe extends to a radius of $10\,mm$, and the two layers of the PXD are found at r = $14\,mm$ and r = $22\,mm$ (see Figure 3.6). These are pixelated sensors based on DEPFET (DEPleted Field Effect Transistor) technology. At SuperKEKB, the high luminosities mean high background from beam-related effects (e.g. the Touschek effect) and from low-momentum-transfer QED processes (e.g. photon-photon reactions). These are especially important for the detectors close to the beam pipe as the background increases approximately with the inverse square of the radius. Strip detectors have a high occupancy (the fraction of channels hit in each triggered event) close to the IP due to this background, making the reconstruction of B-decay vertices impossible. At SuperKEKB luminosities, strip detectors can only be used at r $> 40\,mm$. Therefore, the PXD must use pixel sensors, which have many more channels than strips, in order to suppress the occupancy to low levels.

The inner layer of the PXD is comprised of 8 planar sensors, each one having a width of $15\,mm$ and an operating length of $90\,mm$. The outer layer consists of 12 modules, each with a width of $15\,mm$ and an operating length of $123\,mm$. As of time of writing, only two of the outer modules are currently installed for performance testing. The full PXD is scheduled for installation by 2022. The lengths are determined by the angular acceptance requirement to be in the polar angle range of 17 degrees forward and 150 degrees backwards. Pixel sizes for the inner and outer layer are 50x50 μm^2 and 50x75 μm^2,

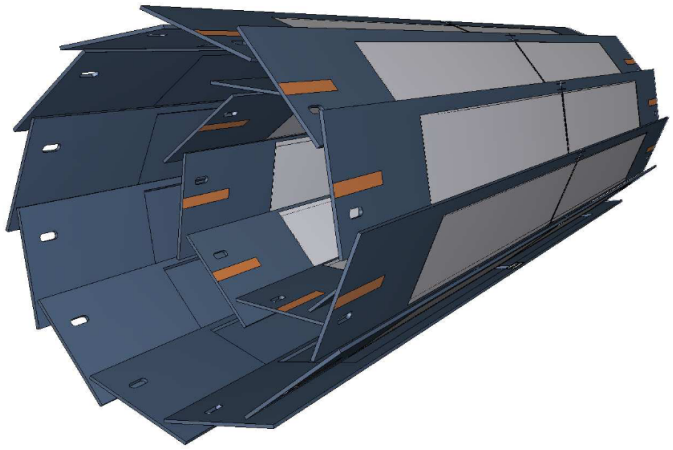

Figure 3.6: Schematic of the two-layer PXD. Light grey represents the DEPFET pixels. Figure from [37].

respectively. Both layers are read out in rolling shutter mode at 100 ns per pixel row. With 1600 pixel rows and four rows being read out in parallel, this translates to about 20 μs for an entire frame.

The pixel detectors used at Belle II are of the DEPFET type. This is a semiconductor detector that combines detection and amplification into a single device. They are very thin (75 μm) with a high signal to noise ratio. With close proximity to the beam pipe, there is a very high background rate and radiation damage for the PXD. The high backgrounds are suppressed by the tiny size of the pixels and 50 kHz readout rate, keeping occupancy low. The radiation damage can only be tackled by specifically fabricating the DEPFET and ASIC chips to withstand large radiation with minimal deterioration of performance.

SVD

The Silicon Vertex Detector (SVD) is also used to measure B meson decay vertices, but additionally can measure vertices of some processes involving D mesons and τ leptons. The outer four layers of the VXD comprise the SVD and are located at a radii of 38 mm, 80 mm, 115 mm, and 140 mm (see Figure 3.7). These are double-sided silicon strip sensors with a short integration time and are far enough away from the IP with fast enough read-

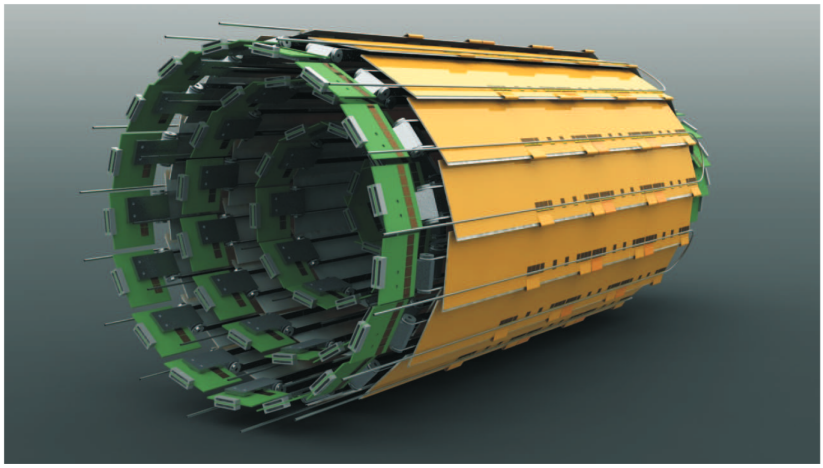

Figure 3.7: SVD barrel with all four concentric layers, cooling lines and hybrids. Figure from [37].

out chips to have sufficiently low occupancy (max of a few percent). Going from inner to outer, they have 16, 30, 56, and 85 sensors of rectangular or trapezoidal shape.

To reduce multiple scattering of the charged particles passing through the detectors (a serious concern at Belle II energies), the VXD needs to be designed with a goal of minimizing the mass. The SVD does this by using large sensors to minimize the amount of structural support required. Up to five sensors are fitted to a support structure called a ladder which uses light-weight styrofoam (Airex) with 65 μm thin carbon fiber on each side [43]. This gives an ultra light structure that is also very rigid. The inner ladder readout chips are fitted onto a flex hybrid circuit and then onto the sensors using an "Origami" style that allows for both sides of the sensor to be readout from the same circuit, ultimately reducing the support structure needed. The sensors at the ends of a ladder do not use the Origami design and use typical hybrid PCBs (printed circuit boards). Each sensor is readout individually with a total of 1902 readout chips.

3.2.2 CDC

The central drift chamber (CDC) is the primary device for tracking charged particles at Belle II. It has three critical functions: 1) reconstructing charged tracks and measuring

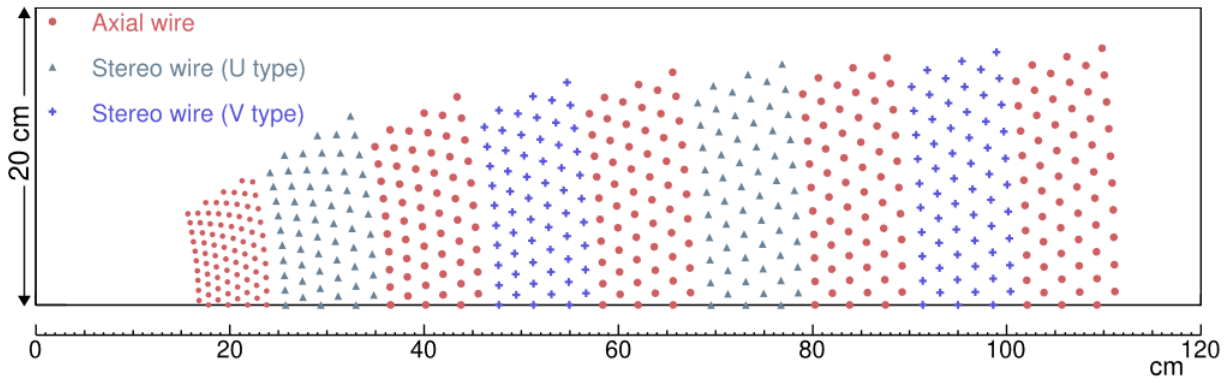

Figure 3.8: CDC sense wires organized into superlayers with alternating orientation. Figure adapted from [44].

their momenta precisely, 2) measuring energy loss to aide particle identification, especially for low momentum particles which do not make it to the other particle identification detectors, and 3) providing trigger signals for charged particles. Like the VXD, multiple scattering is the dominant effect on momentum resolution and low material is needed. This is achieved using thin Aluminum endplates, thin CFRP (carbon-fiber-reinforced polymers) inner and outer cylindrical barrels, a low-Z gas, and low mass wires.

The CDC is a 2.3 m long cylindrical unit with an inner radius of 160 mm and an outer radius of 1130 mm. It has 56 layers made up of a total of 14,366 gold-plated tungsten sense wires (each 40 μm in diameter) in both axial (aligned with the magnetic field) and stereo (skewed) orientations. The are nine 'superlayers', which are groups of 6 layers (8 for the innermost superlayer) in the same orientation. The superlayers alternate between axial and stereo orientation, with the axial orientation being the innermost and outermost superlayers (see Figure 3.8). Each sense wire is surrounded by eight field wires, creating a square cell. There are a total of 42,240 aluminum field wires. The radial cell size is 10 mm for the innermost superlayer, and roughly 18.2 mm for the rest. The smaller cell size for the innermost superlayer is due to the larger backgrounds and need to lower the occupancy. The combination of axial and stereo orientations allow for a full 3D reconstruction

of a helix track. The gas is a 50:50 He-C_2H_6 mixture with an average drift velocity of 3.3 $cm/\mu s$ and a max drift time of 350 ns (for a 17 mm cell size).

The CDC is inside of the 1.5 T magnetic field generated by the superconducting coil. This induces charged particles to curve and interact with various sense wires via ionization of the gas. The degree of curvature allows for determination of the momentum, while the energy deposited on the sense wires will allow for specific ionization, dE/dx, to be measured. With this setup, the CDC achieves a spatial resolution between 50-120 μm, depending on the layer and on the incident angles [44]. However, it is typically better than 100 μm. The transverse momentum, p_{T}, resolution for charged tracks is $0.127 p_{\mathrm{T}} \bigoplus 0.321\%$, where p_{T} is in GeV/c [44]. The resolution of dE/dx is between 8 and 12%, depending on the incident angle, with a 90 degree angle having the best resolution [44].

3.2.3 Particle Identification

The ability to identify charged particles (e, μ, π, K) over the entire kinematic range of the experiment is essential to Belle II. Particularly important to the experiment, and indeed to the analysis presented in this thesis, is the improved K/π separation. Surrounding the CDC are two different types of ring-imaging Cherenkov particle identification detectors. In the barrel region there is a time-of-propagation (TOP) counter and on the forward endcap there is the Aerogel Ring-Imaging Cherenkov (ARICH) detector. The basic operating principle of a Cherenkov detector is to take advantage of Cherenkov radiation, whereby a particle traversing through a medium faster than light will emit a cone of light with some angle θ_{C} about the direction of travel. The angle θ_{C} is dependent on the velocity of the traversing particle, and so if the momentum is known from the CDC, the particle mass can be deduced and thus identified. A brief overview of each will be delivered below.

TOP

The time-of-propagation (TOP) counter is used in the barrel region of the detector. It consists of 16 quartz bar detector modules which are 2.6 m long, 45 cm wide, and 2 cm tall.

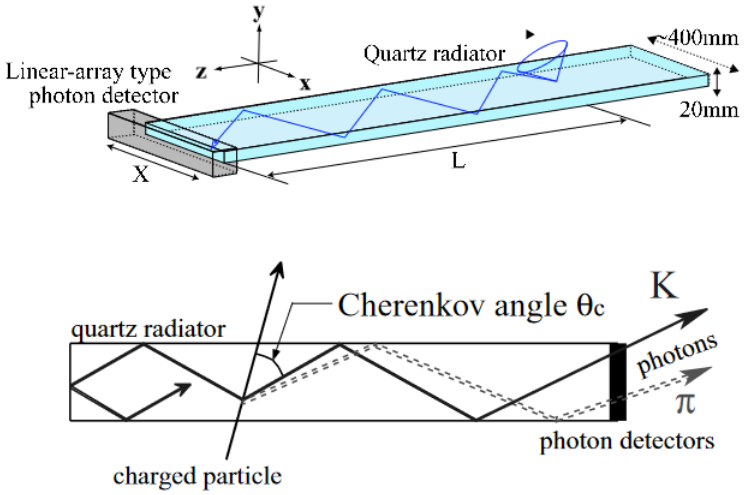

Figure 3.9: Top: general overview of the TOP counter. Bottom: schematic side-view showing difference in Cherenkov angles of photons generated from kaons and pions. Figures from [37].

Each bar expands into a wedge at the sensor end of the bar, and a spherical focus mirror is placed at the opposite end. A charged particle going through one of the quartz bars with sufficient velocity will emit Cherenkov radiation, and the photons will internally reflect to the sensors. The time of propagation and x - y coordinates of the photons are measured via two rows of sixteen micro-channel plate (MCP) photomultiplier tubes (PMTs), with single photon time resolution of at least $100\,ps$. This method also relies on the starting time (time of the initial e^+e^- collision), which must be known to within $50\,ps$. The precise timing and photon coordinates in x - y can then be used to reconstruct the Cherenkov ring image. See Figure 3.9.

ARICH

The ARICH is employed in the forward end-cap region of the detector and has been designed to distinguish kaons from pions over most of their momentum spectrum, and to discriminate between pions, muons and electrons below $1\,GeV/c$. The basic principle is to have an aerogel radiator be the medium to create Cherenkov photons, then have an

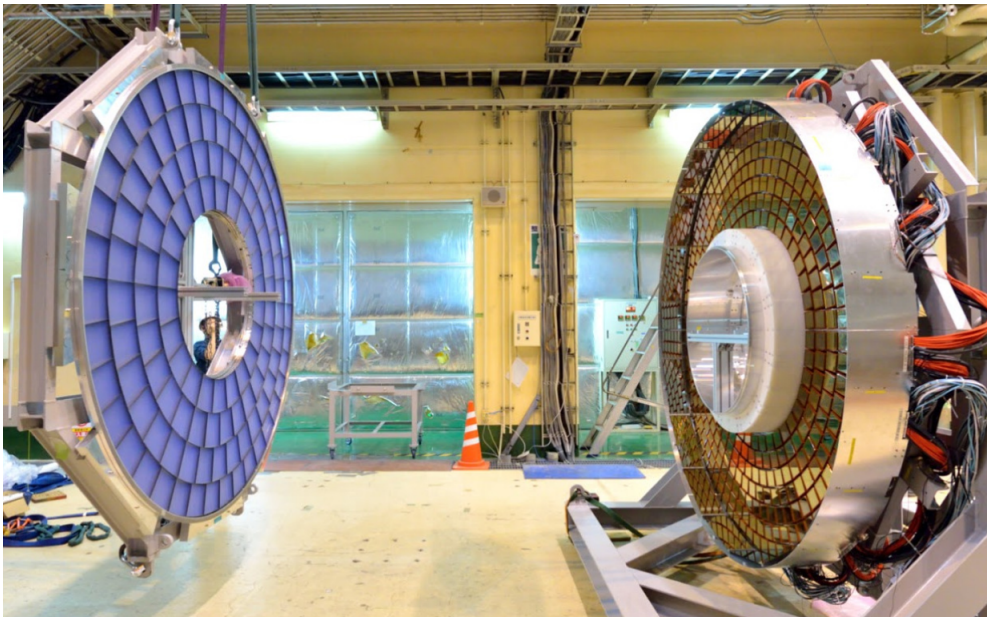

Figure 3.10: ARICH assembly in October 2017. Aerogel array on the left, HAPD array on the right. Figure from [46].

expansion volume of roughly 20 cm to allow the rings to form, followed by an array of position sensitive photon detectors, with high single photon efficiency and good resolution (see Figure 3.10). An important parameter to optimize is the number of photons, where the more detected the better. A novel approach to maximize this was taken at Belle II, where two 2 cm layers of aerogel with slightly different refractive index (n_1 = 1.045, n_2 = 1.055) are used, so that both layers create Cherenkov photons and the refractive indices and geometric positioning are optimized such that the rings will overlap at the photon detectors. About 124 hexagonal aerogel tiles are used for each layer, covering 3.5 m^2 of the forward endcap [45]. The photon detector plane is an array of 420 73 x 73 mm^2 hybrid avalanche photon detectors (HAPDs) with an overall gain of 70,000 [45]. In each HAPD, the photo-electrons are accelerated and detected in avalanche photodiodes.

3.2.4 ECL

Extending beyond the particle identification detectors is the electromagnetic calorimeter (ECL). The ECL plays a very important role at Belle II. Neutral particles, mainly neutral pions, are abundant in B decays and deposit many photons over a large range of energies. One primary function of the ECL is the high efficiency detection of photons along with a precise measurement of their energy and angular coordinates. The ECL can additionally aide in the identification of electrons and low transverse momentum muons, and detection of K^0_{L} mesons. The ECL also plays a role in generating the signal for triggers and in calculating the online luminosity.

The ECL is comprised of 6624 thallium-doped caesium iodide CsI(Tl) crystals covering the barrel region and 2112 CsI(Tl) crystals covering the endcaps, for 90% coverage of the solid angle in the centre-of-mass system. The crystals are truncated pyramids with an average size of about 6 x 6 cm^2 in cross section and 30 cm in length. Two photodiodes, each with a water cooled preamplifier, are attached to the rear surface of each crystal for two independent output signals per crystal. These are then summed at an external shaper board, which also contain fast shaper amplifiers that generate signals for triggering. The shaper feeds the signal to an 18 bit, 1.76 MHz waveform digitizer for signal processing. An online FPGA (field-programmable gate array - an integrated circuit) fits the waveforms and extracts an amplitude and time. A novel approach also exploits the differences in hadronic and electromagnetic scintillation to allow for Pulse-Shape Discrimination (PSD).

3.2.5 KLM

The final layer of detector is the K^0_{L} and muon detector (KLM). The KLM covers both the barrel and the endcaps, with alternating layers of 4.7 cm thick iron plates and detector elements. In the barrel there are 15 detector layers and 14 iron plates, while each endcap has 14 detector layers and 14 iron plates. The iron layers, in addition to providing the

magnetic flux return for the solenoid, supply a minimum of 3.9 interaction lengths for the K_{L} to hadronically shower. K_{L} mesons can therefore be detected either in the ECL or KLM via hadronic showers. Muons with momentum above roughly $0.6\,GeV/c$ reach the KLM and can be detected as they deposit energy.

The detectors are glass-electrode resistive plate counters (RPCs) in the outer 13 layers of the barrel. In the inner barrel and the endcaps, ambient neutron background rates were too large for RPCs and instead scintillators were used. The scintillators are rectangular prisms internally coated with reflective TiO_2. At the center of the scintillator is a wavelength shifting (WLS) fibre that collects light created by traversing charged particles, and sends the light to a Silicon photomultiplier (SiPM) for detection.

3.2.6 Super Conducting Coil and Overall Structure

In addition to the detector components, a superconducting solenoid in between the ECL and KLM provides a $1.5\,T$ magnetic field. The iron structure integrated into the KLM acts as the return path for the magnetic flux of the solenoid. It also provides the overall structure and support system for all of the detector components. A liquid helium cryogenic system is used to cool the cryostat in which the solenoid is contained. The same system is used to cool the cryostats housing the QCS focusing magnets.

3.2.7 Data Acquisition

At SuperKEKB, bunches of electrons and positrons cross the interaction region at a high rate, as much as every 4 ns [1a]. Most crossings result in either no collision or are uninteresting interactions (e.g. $e^+e^- \to e^+e^-$, which is also the most common interaction). The need to selectively choose which events to record as data is paramount for an intensity frontier experiment, such as Belle II, in order to keep data storage manageable. This is done with the Belle II online system, consisting of the Data Acquisition (DAQ), Level 1 Trigger (TRG or L1), and the High Level Trigger (HLT). The TRG is the first system an

event must pass, which basically reads low resolution readout data live from the CDC, ECL and KLM. The TRG uses specialized fast electronics to do this, and if it detects an interesting event it will create a trigger signal. This trigger signal is then sent to the sub-detectors, and the DAQ system ensures that this is done synchronously. The DAQ also provides the high-speed data links that retrieve the data and send it to the HLT system. The HLT system is a computing cluster in close proximity to the detector. It is comprised of roughly 10,000 CPU cores and receives the full set of subdetector data for each TRG triggered event. The HLT system will quickly perform a reconstruction using the basf2 software and then determine whether the event should be kept or discarded. If kept, it gets stored on a local offline hard drive. The HLT reduces the amount of offline storage used by roughly 60% [47]. Once this data is stored it is further processed with calibrations and various skims are applied. The finished data is made available on a distributed computing grid.

Chapter 4

Analysis

4.1 Decay Mode and Method Overview

The analysis presented in this thesis is focused on the decay of $B^+ \to K^+\tau^+e^-$ and its charged conjugate. This mode is one of four in the broader family $B \to K\tau\ell$, where ℓ can be an electron or a muon, and either lepton can have opposite charge to that of the B and the K mesons. The other modes will be the focus of future work. Our general analysis method has been roughly adapted from BaBar (2012) [35].

The mode cannot be directly reconstructed in full due to the presence of the τ lepton, which will necessarily have neutrinos as a decay product. These neutrinos escape detection and their kinematic information is lost. An alternative approach is required, and has indeed been developed and used in B factory analyses involving missing energy (i.e. involving undetected neutrinos). The events at Belle II that we are interested in are $e^+e^- \to \Upsilon(4S) \to B^+B^-$. Let the B meson that decays into our signal mode be denoted as the signal B, B_{sig}. The other B meson will then be denoted as the tagged B, B_{tag}. The technique involves fully reconstructing the B_{tag} using hadronic modes (e.g. $B \to D\pi$; $D \to K\pi$) to precisely determine the three-momentum of the B_{sig} from the event kinematics. In the rest frame of the Υ(4S), the B_{sig} and B_{tag} decay back-to-back with equal and opposite momenta, each with energies equal to half of the center-of-mass energy of the

colliding e^+e^-. By recovering the $B_{\rm tag}$ momentum vector we also recover the full $B_{\rm sig}$ momentum vector. Using hadronic modes to reconstruct the $B_{\rm tag}$ ensures a high resolution on the momentum (high purity), although at the cost of efficiency. The τ four-momentum can then be indirectly reconstructed via

$$E_\tau = \frac{E_{CM}}{2} - E_K - E_e, \tag{4.1}$$

$$\vec{p}_\tau = -\vec{p}_{tag} - \vec{p}_K - \vec{p}_e, \tag{4.2}$$

$$m_\tau = \sqrt{E_\tau^2 - |\vec{p}_\tau|^2}, \tag{4.3}$$

where E_{CM} = 10.58 GeV, and $(E_K, \vec{p}_K)$ and $(E_e, \vec{p}_e)$ are the known four-momenta of the kaon and electron coming from the $B_{\rm sig}$. The τ mass will peak sharply at the nominal value for signal candidates and have a broad distribution for combinatorial backgrounds.

We perform a "cut and count" analysis using Monte Carlo simulations. In this type of analysis, we simply look at a signal region (which we define as the region of reconstructed tau mass within 60 MeV of the nominal tau mass) and count the number of expected background events and the number of expected signal events. Once the process (i.e. what cuts to make) is determined, we can look at real data and count the events. Based on the number of events in real data and the expected signal efficiency and number of background events, we can calculate the significance that there is signal present. While we are defining the process we do not look at data within an extended signal region (within 175 MeV of the nominal tau mass) in order to remain unbiased, also known as a blinded study. For this thesis, we keep the tau mass signal region blinded. We use the blinded data to look at various distributions in order to compare MC/data agreement and estimate systematic uncertainties.

"One-pronged" decays are those that have just one charged decay product. In our reconstruction of the tau, we consider four one-pronged decays (branching fractions indicated in parentheses [48]):

- e channel: $\tau \to e\nu\overline{\nu}$ (17.83%),
- mu channel: $\tau \to \mu\nu\overline{\nu}$ (17.41%),
- pi channel:$\tau \to \pi^+\nu$ (10.83%), and
- rho channel: $\tau \to \pi^+\pi^0\nu$ (25.52%).

These channels combine to total 71.59% of tau decays. Between the pi and the rho channel we expect a high cross-feed (i.e. events from one channel being identified as candidates in a different channel). These channels will be treated separately (potentially different cuts on different channels), combining the results at the end. It is also worth noting that $\tau \to \pi^+\pi^0\pi^0\nu$ has a branching fraction of 9.30% and $\tau \to \pi^+\pi^0\pi^0\pi^0\nu$ has a branching fraction of 1.05%. These are not included in the analysis (except as cross-feed) as they have very high combinatorics and take a lot of computational time to process, but should be considered in future studies. Additionally, the three-pronged channel $\tau \to \pi^+\pi^-\pi^-\nu$ has a 9.8% branching fraction, and could also be considered in future studies.

4.2 Data Sample and Simulations

4.2.1 Data Sample

For this analysis we use a dataset of 62.8 fb^{-1} to study the MC/data relation. As mentioned, we do not look at events in the data within $\pm 175\, MeV/c^2$ of the nominal tau mass (signal region) in order to keep the study blinded. Unblinding the data is not within the scope of this thesis.

4.2.2 Monte Carlo Samples

For this analysis we used Monte Carlo (MC) generated samples of generic and signal events. These are produced by the Belle II data production team. In general, MC is produced first via generators (e.g. EvtGen [49]), which create the particles (positions and

four-vectors) of a specified collision. Then the generated events are simulated in a virtual Belle II detector (e.g. Geant4 [50]) in order to convert the events into detector data. The event generator used for generic and signal MC in our analysis is EvtGen. For our signal B meson, the decay uses PHOTOS [51] to account for radiative corrections and PHSP to generate a uniform three-body phase-space model, as the NP process is unknown. The simulation of detector interaction is done with Geant4.

The generic MC contains events of the types

- $e^+e^- \to \Upsilon(4S) \to B^+B^-$,
- $e^+e^- \to \Upsilon(4S) \to B^0\overline{B^0}$,
- $e^+e^- \to q\bar{q}$ $(q = u, d, c, s)$ and
- $e^+e^- \to \tau^+\tau^-$,

with the daughters decaying generically into all known modes. The generic MC used in this analysis is equivalent to $2\,\text{ab}^{-1}$ of data and is broken up into charged, mixed, uubar, ddbar, ccbar, ssbar, and taupair samples. The signal MC corresponds to 50 million $\Upsilon(4S) \to B^+B^-$ events, with one of the B mesons decaying into our signal mode ($B^\pm \to K^\pm\tau^\pm e^\mp$), and one of them decaying generically. The tau is allowed to decay generically and is not specified to be one of our targeted tau channels.

4.3 FEI

The Full Event Interpretation (FEI) algorithm is an exclusive tagging algorithm developed for the Belle II experiment [52]. It is designed to fully reconstruct the tag-side B meson, allowing for inferred kinematics of the signal-side B. The FEI utilizes a multivariate algorithm in a neural network to reconstruct the B_{tag} in more than 100 decay channels, both semileptonic and hadronic, leading to $\mathcal{O}$(10,000) distinct decay chains. For reasons described above, we only use the hadronic FEI. The FEI starts with detector level information (tracks, clusters, vertices) to construct final-state particles and builds up intermediate

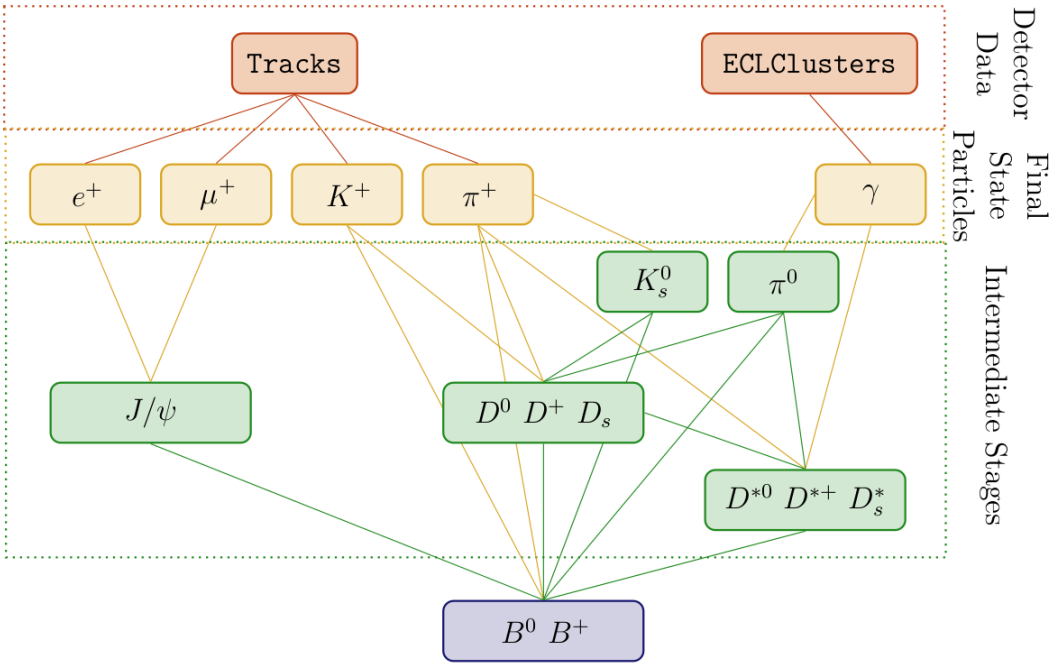

Figure 4.1: Overview of the FEI hierarchy. The algorithm takes final-state particles and gradually reconstructs intermediate objects until a B meson candidate is obtained. Taken from the internal Belle II Confluence.

states hierarchically until it has reconstructed a plausible B meson candidate (See Figure 4.1). Charged final-state particles are reconstructed as tracks in the CDC and VXD, while neutral final-state particles get reconstructed as clusters in the ECL. Once the FEI has reconstructed a B candidate, the remaining tracks and clusters are left for the signal side B reconstruction, and can be used to reject reducible background (e.g. for our decay we require precisely three tracks remain after the B_{tag} reconstruction). The FEI can reconstruct multiple candidates for any single event, and employs gradient-boosted decision trees to calculate the probability that the decay chain it reconstructs correctly describes the true process. It delivers this probability as a variable called *Signal Probability*, which can be used to discriminate correctly tagged candidates and incorrect ones.

The performance of the FEI can be measured with the tag-side efficiency (the fraction of all events with a correct tag), and the tag-side purity (the fraction of tagged events

with a correct tag). Hadronic tagging suffers from low tag-side efficiency, but has a comparatively high tag-side purity. The inverse is true for semileptonic tagging. The FEI is directly implemented in the Belle II Analysis Software Framework (basf2). The multivariate classifiers used by the FEI are trained on MC simulated events, and are retrained with every new major central production of MC. Within the FEI, cuts are made in two stages, pre-cuts and post-cuts. The pre-cuts are made before the multivariate classifiers have been applied in order to reduce computing time and save on memory consumption. The first pre-cut is loose and fast, generally on a quantity such as invariant mass or the B meson beam-constrained mass. Once that is made, the candidates will be ranked and only 10-20 best-candidates in each decay channel are considered. Once this is done, the more computationally expensive parts of the reconstruction can continue. These are the matching of reconstructed particles to generated particles (if MC), vertex fitting, and the multivariate classification. After this the post-cuts will be applied. The first post-cut is a loose cut on the signal probability. The second is ranking of all candidates based on signal probability and only keeping the 10-20 overall best-candidates.

Improvements to the FEI are continually happening and the version of the FEI used in this analysis is FEIv4_2020_MC13_release_04_01_01, which is trained on MC samples from the same campaign (MC13) and the same basf2 release version (release-04) that we use. Additionally, the FEI can be adapted generically or specifically. In the generic FEI, it is trained independently of any specific signal-side decay mode. Specific FEI is also possible, where the FEI is trained only on the remaining tracks and clusters after a signal event has been specified. For the purposes of this thesis, the generic FEI was used. It may be of future interest to explore the specific FEI in the context of our analysis.

4.4 Event selection

Once we have our MC/data samples, we go through our analysis process. We write and process our scripts using basf2. basf2 uses modules, usually written in C++, that each

do some specific process (e.g. reading data, particle reconstruction, etc.) These modules are linked together in a linear order via a path in a steering script, written in python, thereby completing some task (e.g. event reconstruction). For this analysis we use basf2 throughout the entire process, outlined here:

1. Run the FEI skim, which applies the FEI and makes some additional loose cuts.
2. Run the reconstruction script, which builds our signal B candidates and the $\Upsilon(4S)$ and makes event cuts.
3. Run a boost script which returns a set of user-defined kinematic variables in the rest frame of the tau for use later in the analysis.
4. Run a candidate selection script, which selects one best candidate per event.
5. Run a final script that applies the rest of the cuts and splits the output files based on tau channel.

These steps will be detailed further in the following sections.

4.4.1 FEI Skim

To apply the FEI, a skim script is used which also makes additional cuts. Skimming is used at Belle II to reduce the amount of data that analysts have to run over. As most of the events will be cut by analysts anyway, skims are created and used by analysts who work on similar analyses to reduce wasting computational resources. The hadronic FEI skim for charged B is outlined below.

Definitions

- Cleaned tracks: $d_0 < 0.5\,cm$, $|z_0| < 2\,cm$, and $p_{\mathrm{T}} > 0.1\,GeV$ [1]
- Cleaned ECL clusters: $0.296706 < \theta < 2.61799$ and $E > 0.1\,GeV$

[1] See [53] for Belle II coordinate system

Event pre-cuts

- $R_2 < 0.4$ (FoxWolframR2 [54], calculated using cleaned tracks and clusters)
- $n_{\text{tracks}} \geq 4$
- $n_{\text{cleanedtracks}} \geq 3$
- $n_{\text{cleanedECLclusters}} \geq 3$
- Visible energy of event $> 4\,GeV$ (CMS frame)
- $2\,GeV < E_{\text{cleanedtracks\&clustersinECL}} < 7\,GeV$

Where n refers to the multiplicity of the specified detector objects, and E refers to reconstructed energies of the specified objects.

Tag-side B cuts

- $m_{\text{bc}} > 5.24\,GeV$
  - where $m_{\text{bc}} = \sqrt{(\frac{E_{\text{CM}}}{2})^2 - p^2_{B_{\text{tag}}}}$
- $|\Delta E| < 0.2\,GeV$
  - where $|\Delta E| = E_{B_{\text{tag}}} - \frac{E_{\text{CM}}}{2}$
- $SignalProbability > 0.001$ (omitted for decay mode 25)

Hadronic Modes used in FEI See Table 4.1 for the list of hadronic decay modes used by the FEI to reconstruct B_{tag}.

After running the hadronic FEI skim, roughly 5% of signal MC events and 5.6% of generic MC events (up to 12% for charged B generic MC) remain.

ID	Decay Mode
1	$B^+ \to \overline{D^0}\pi^+$
2	$B^+ \to \overline{D^0}\pi^+\pi^0$
3	$B^+ \to \overline{D^0}\pi^+\pi^0\pi^0$
4	$B^+ \to \overline{D^0}\pi^+\pi^+\pi^-$
5	$B^+ \to \overline{D^0}\pi^+\pi^+\pi^-\pi^0$
6	$B^+ \to \overline{D^0}D^+$
7	$B^+ \to \overline{D^0}D^+K^0_S$
8	$B^+ \to \overline{D^{0*}}D^+K^0_S$
9	$B^+ \to \overline{D^0}D^{+*}K^0_S$
10	$B^+ \to \overline{D^{0*}}D^{+*}K^0_S$
11	$B^+ \to \overline{D^0}D^0K^+$
12	$B^+ \to \overline{D^{0*}}D^0K^+$
13	$B^+ \to \overline{D^0}D^{0*}K^+$
14	$B^+ \to \overline{D^{0*}}D^{0*}K^+$
15	$B^+ \to D^+_s\overline{D^0}$
16	$B^+ \to \overline{D^{0*}}\pi^+$
17	$B^+ \to \overline{D^{0*}}\pi^+\pi^0$
18	$B^+ \to \overline{D^{0*}}\pi^+\pi^0\pi^0$
19	$B^+ \to \overline{D^{0*}}\pi^+\pi^+\pi^-$
20	$B^+ \to \overline{D^{0*}}\pi^+\pi^+\pi^-\pi^0$
21	$B^+ \to D^{+*}_s\overline{D^0}$
22	$B^+ \to D^+_s\overline{D^{0*}}$
23	$B^+ \to \overline{D^0}K^+$
24	$B^+ \to D^-\pi^+\pi^+$
25	$B^+ \to D^-\pi^+\pi^+\pi^0$
26	$B^+ \to J/\psi K^+$
27	$B^+ \to J/\psi K^+\pi^+\pi^-$
28	$B^+ \to J/\psi K^+\pi^0$
29	$B^+ \to J/\psi K^0_S\pi^+$
30	$B^+ \to \Lambda^-_c p\pi^+\pi^0$
31	$B^+ \to \Lambda^-_c p\pi^+\pi^-\pi^+$
32	$B^+ \to \overline{D^0}p\bar{p}\pi^+$
33	$B^+ \to \overline{D^{0*}}p\bar{p}\pi^+$
34	$B^+ \to D^+p\bar{p}\pi^+\pi^-$
35	$B^+ \to D^{+*}p\bar{p}\pi^+\pi^-$
36	$B^+ \to \Lambda^-_c p\pi^+$

Table 4.1: Hadronic decay modes used by the FEI to reconstruct B_{tag}.

4.4.2 Reconstruction

Once the FEI script has ran, we run a reconstruction script. The reconstruction script is ran using the same basf2 release version as the FEI, to maintain consistency (release-04-02).

The first thing that is done in the script is importing the ROOT files from the FEI, which include the list of reconstructed *B+:tag*. We also import the lists of charged (e, μ, K, π^+) and neutral particles (π^0) using a pre-defined selection. For charged particles, we use the selection known as 'good' (e.g. *K+:good*), which imposes track cuts and a particle ID (pID) cut:

- $thetaInCDCAcceptance$
  - requires the particle polar angle θ to be within the range $17° < \theta < 150°$. Computed using only the initial particle momentum.
- $nCDCHits > 20$
- $dr < 0.5\,cm$
- $|dz| < 2\,cm$
- $pID > 0.5$.

For the neutral pions we use the selection known as 'eff60', which uses a photon selection along with a mass range cut of $0.075 < M < 0.175\,GeV/c^2$ to achieve roughly 60% π^0 efficiency.

We impose further track cuts and a stricter pID cut for the charged lists:

- $|d0| < 1.5\,cm$
- $|z0| < 2.5\,cm$
- $p_\mathrm{t} > 0.05\,GeV$ (transverse momentum, lab frame)
- $p < 10\,GeV$ (momentum, lab frame)

- $pID > 0.6$ (e, μ, K)
- $pID > 0.5$ (π^+).

We also employ a hierarchical pID cut system to ensure that no particles are featured in more than a single list. We start with the kaon list and just require kaonID > 0.6. Then the muon list will require muonID > 0.6, and additionally that kaonID < 0.6. The electron list will require electronID > 0.6, and additionally both the kaonID and muonID be < 0.6. Finally, the charged pions require pionID > 0.5, and additionally kaonID, muonID, and electronID be < 0.6.

Once these lists are created, we begin reconstructing the signal side. The four tau channels are reconstructed by requiring the daughter is exactly one charged track passing the proper pID requirements. The rho channel is directly reconstructed as $\tau \to \pi^+\pi^0$, with basf2 automatically reconstructing the intermediate rho resonance. The signal B is then reconstructed from a kaon and tau with the same sign and an electron with the opposite sign. The $\Upsilon(4S)$ then gets reconstructed with the tag B and the signal B of opposite charge.

Once we have the $\Upsilon(4S)$, there should be nothing left over, so we ensure this by building the "Rest of Event" (ROE), and require that there are no tracks (charge) in it. Finally, we declare all the variables we wish to include in the output and process the path.

4.4.3 Best Candidate Selection

At this point, each event can have multiple candidates for how it was precisely reconstructed. We choose the best candidate by first boosting the particles on the signal side into the rest frame of the tau. The candidate chosen is the one where the primary electron (e^- from the signal B) has the highest momentum in this reference frame. Multiple rankings were compared, including ranking by the 'SignalProbability', and although there were minimal differences in the result, we chose the one that fared the best in reconstructing the tau mass in signal MC.

4.4.4 Final Selection with Punzi Optimization

Applying all of the above we are left with approximately 0.40% of signal MC events, and 0.006% of background (generic MC) events.

We then look at several variables for potential background suppression. In order to maximize sensitivity, we find cut values that optimize the Punzi figure of merit (FOM) [55], which is defined as

$$FOM_{\text{Punzi}} = \frac{\epsilon_{\text{signal}}}{\frac{n_\sigma}{2} + \sqrt{N_{\text{B}}}}, \tag{4.4}$$

where ϵ_{signal} is the signal efficiency, n_σ is the number of sigmas corresponding to a one-sided Gaussian test (we set n_σ = 3, i.e. 3σ significance), and N_{B} is the number of background events. This is a beneficial figure of merit to use in the case of NP as it does not rely on the signal branching fraction as input, which is unknown.

The following variables have been studied and optimized by choosing values that maximize the Punzi FOM for a dataset of 200 fb^{-1}, which is roughly the expected dataset available at Belle II by the end of 2021. The results will be shown in the section that follows.

B_{tag} m_{bc} The beam constrained mass of the tagged B is a good variable to reduce continuum background, as it should peak around the B meson mass for signal events. Note that the FEI skim already employs a cut of B_{tag} $m_{\text{bc}} > 5.24\, GeV/c^2$.

momentum Various particle momenta were also looked at. The momentum of the primary electron and the tau daughter, both in the CM frame were looked at and determined to not be discriminating between signal and background. When boosted to the tau frame, the momentum of the daughter of the tau can become a useful discriminating variable. The $\tau \to \pi$ channel is a two-body decay, and thus the pion will have a discrete momentum. As mentioned earlier, there is significant cross-feed between the pion and rho channel, so this variable can also be useful in the rho channel.

particle ID The pID of the charged particles on the signal side can also be good discriminating variables. This was found to be most useful for the leptons, and especially the primary electron. In the end, the electronID of the primary electron was the only additional pID cut necessary.

Signal Probability The Signal Probability is returned by the FEI and can theoretically be used for the discrimination of good and bad reconstructed B_{tag} candidates. It is already cut on loosely by the FEI skim (> 0.001). In practice, most of the values are very low and for the low statistics of 200 fb^{-1}, it was not powerful enough to use. However, as the FEI improves and with larger datasets it will likely be useful to make a tighter cut on this variable.

Extra Energy The extra energy is the remaining ECL energy in the ROE. Since we remove all events with leftover tracks, this energy is due strictly to neutrals. This was found to be a useful discriminating variable.

$m_{\ell\ell}$ **and** $m_{K\pi}$ B decays with charmonium, $B \to K(c\overline{c})$; $(c\overline{c}) \to \ell^+\ell^-$, can get through our signal selection in the $\tau \to e$ channel. To remove these background events, we take the invariant mass of the electron and positron, $m_{\ell\ell}$, and remove events that fall into the mass range of the J/Ψ ($3.04 < m_{\ell\ell} < 3.14\,\mathrm{GeV/c^2}$). For larger datasets it would also make sense to remove events in the mass range of the $\Psi(2S)$ ($3.60 < m_{\ell\ell} < 3.75$), but for 200 fb^{-1} it does not make any significant improvements. Additionally, requiring $m_{\ell\ell} > 0.1\,\mathrm{GeV/c^2}$ removes photon conversion events, also in the $\tau \to e$ channel.

Semileptonic D decays, such as $B^+ \to \overline{D}^{(*)0}X^+$; $\overline{D}^0 \to K^+\ell^-\overline{\nu}_\ell$, are one the most prominent backgrounds remaining. We calculate the invariant mass of the kaon and the electron, which is assumed to be a pion, for which there is a strong peak around the mass of the D meson in the background. By requiring $m_{K\pi} > 1.89\,GeV/c^2$, we remove the majority of the background at the much lesser cost of signal events.

Tau Mass The reconstructed tau mass is our signal variable. We do not look within $\pm 175\, MeV/c^2$ of the tau mass in data, and consider the signal region to be $\pm 60\, MeV/c^2$ of the nominal tau mass.

Chapter 5

Results

5.1 Optimization

The following histograms are plotted based on tau channel, before final cuts are applied and with no cut on the tau mass range. B_{tag} m_{bc} is shown in Figure 5.1. The signal MC peaks strongly at the B meson mass, while the combinatoric background is visible below the B peak. The background content is dependent on tau channel, with the leptonic channels having a higher contribution from the mixed and charged samples, and the hadronic channels having a higher continuum contribution. The optimized cut line is shown with a red dashed line, which retains most of the signal.

$m_{K\pi}$ is plotted in Figure 5.2. This is a very strong discriminating variable as can be seen by the distribution of generic vs signal MC. Although signal MC is somewhat broadly distributed, the background peaks at a much lower value. This cut removes a lot of the background in every channel, at a modest cost of signal events.

$m_{\ell\ell}$ is shown in Figure 5.3. The only channel that this cut is applied to is the electron channel, due to the previously mentioned photon conversion events below 0.1 GeV/c^2, and the charmonium resonance between 3.04 and 3.14 GeV/c^2.

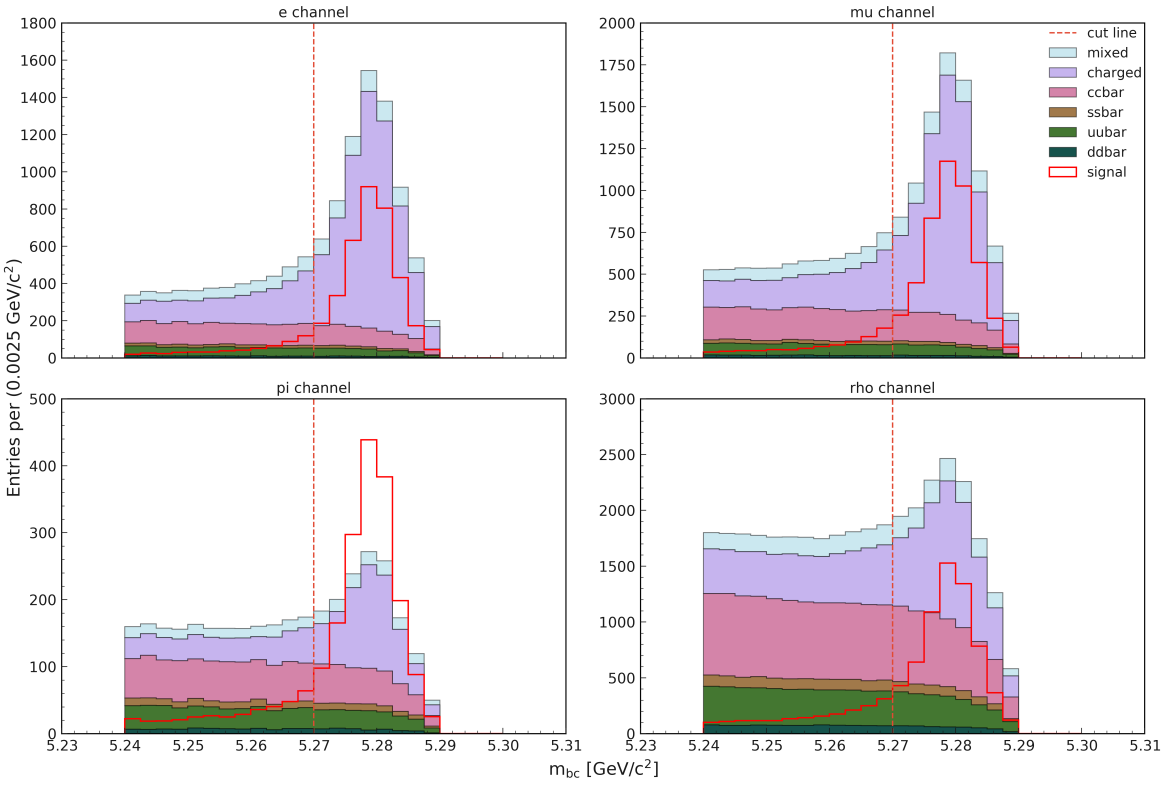

Figure 5.1: Histograms of B_{tag} m_{bc}, based on tau decay channel. The red dashed line indicates the cut line, for which we retain events above that value ($>$5.27 GeV/c^2). The generic MC is scaled to 200 fb^{-1}, and the signal MC is arbitrarily scaled.

The primary electron electronID is presented in Figure 5.4. For most events this value is close to 1, but there is a small tail of background below the cut line of 0.98 that gets removed.

The ROE extra energy is shown in Figure 5.5. This is the sum of ECL cluster energies that are left over after reconstruction. Only clusters with energy $> 20\,MeV$ are considered. For signal MC, the source of this extra energy are π^0s that either come from the tau and were not used in its reconstruction or that were from the B_{tag} and got missed in the FEI reconstruction. In the electron and muon channel, this variable is strong at rejecting continuum as ROE extra energy tends to peak at smaller values for signal and have a

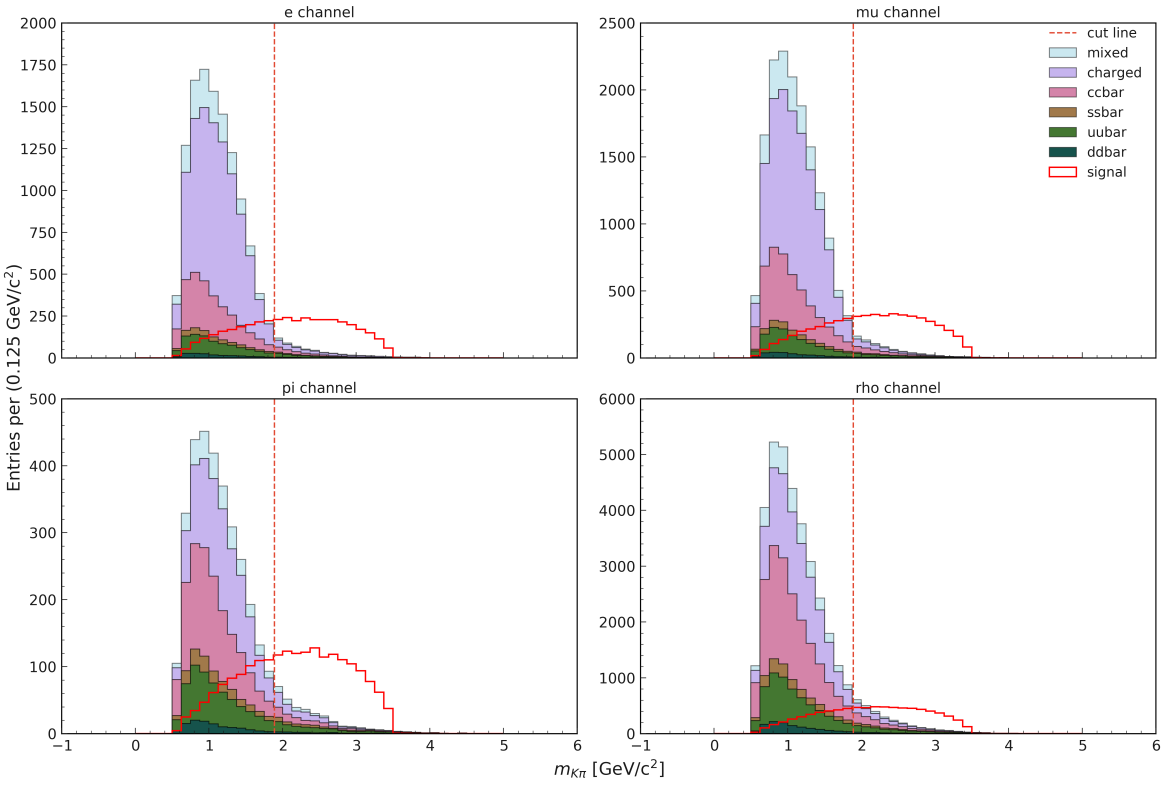

Figure 5.2: Histograms of $m_{K\pi}$, based on tau decay channel. The red dashed line indicates the cut line, for which we retain events above that value ($>$1.89 GeV/c^2). The generic MC is scaled to 200 fb^{-1}, and the signal MC is arbitrarily scaled.

broader, higher value for continuum. In the pion and rho channel, the ROE extra energy cut is powerful in rejecting all types of background, as all samples tend to peak at higher values, not just continuum.

The tau daughter momentum in the tau rest frame is plotted in Figure 5.6. The cut applies only to the rho channel. It might appear useful to make cuts in other channels, but the other variables we use get rid of most of the background in those channels, and so cutting on this variable doesn't increase our sensitivity. You can clearly see the peak in the pion channel, as it is a two-body decay. Additionally, the rho channel shows a small peak, indicating cross-feed from the pion channel.

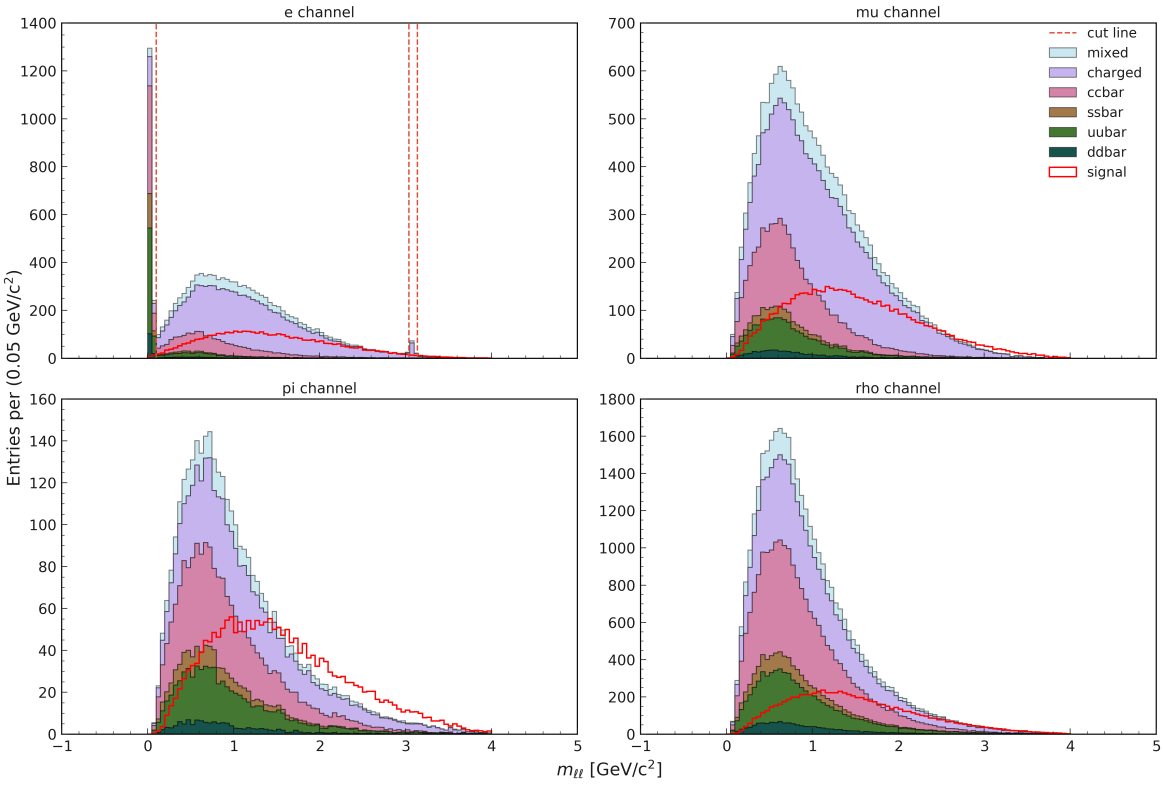

Figure 5.3: Histograms of $m_{\ell\ell}$, based on tau decay channel. The red dashed lines indicate the cut lines, only applying to the electron channel, for which we retain events above $>$0.1 GeV/c^2 (gets rid of photon conversion events) and additionally veto events in the range $3.04 < m_{\ell\ell} < 3.14\,GeV/c^2$ (remove charmonium events). The generic MC is scaled to 200 fb^{-1}, and the signal MC is arbitrarily scaled.

Lastly, the tau mass is plotted, both before the final cuts (Figure 5.7) and after the final cuts (Figure 5.8). The nominal tau mass is shown with a red dashed line. The tau mass peaks strongly at the nominal value for signal, and tends to have broader distribution for background. For reference, each bin is 100 MeV and the signal region (not shown) is 120 MeV.

Additional plots of the FEI modes that MC were reconstructed in (Figure A.1), the Signal Probability (Figure A.2), and the primary electron momentum (Figure A.4) and

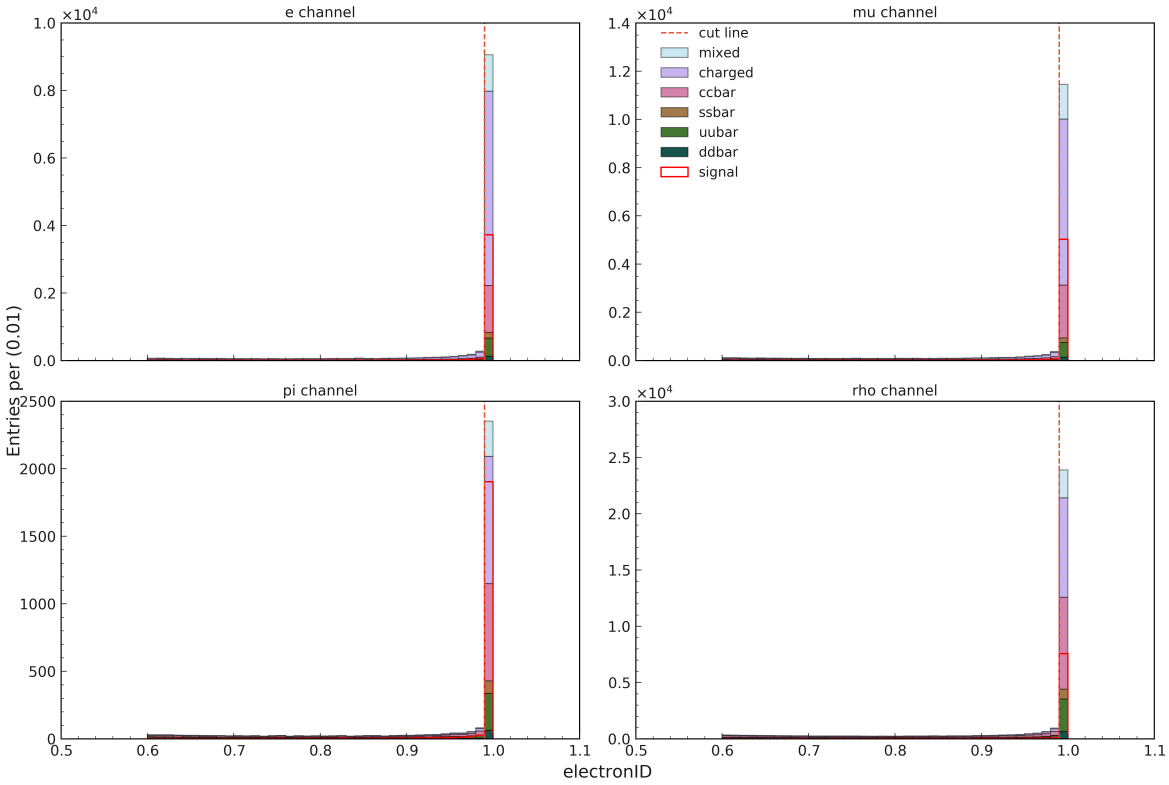

Figure 5.4: Histograms of the primary electron electronID, based on tau decay channel. The red dashed line indicates the cut line, for which we retain events above that value (>0.98). The generic MC is scaled to 200 fb^{-1}, and the signal MC is arbitrarily scaled.

tau daughter momentum (Figure A.3), both in the CMS frame, can be found in Appendix A.

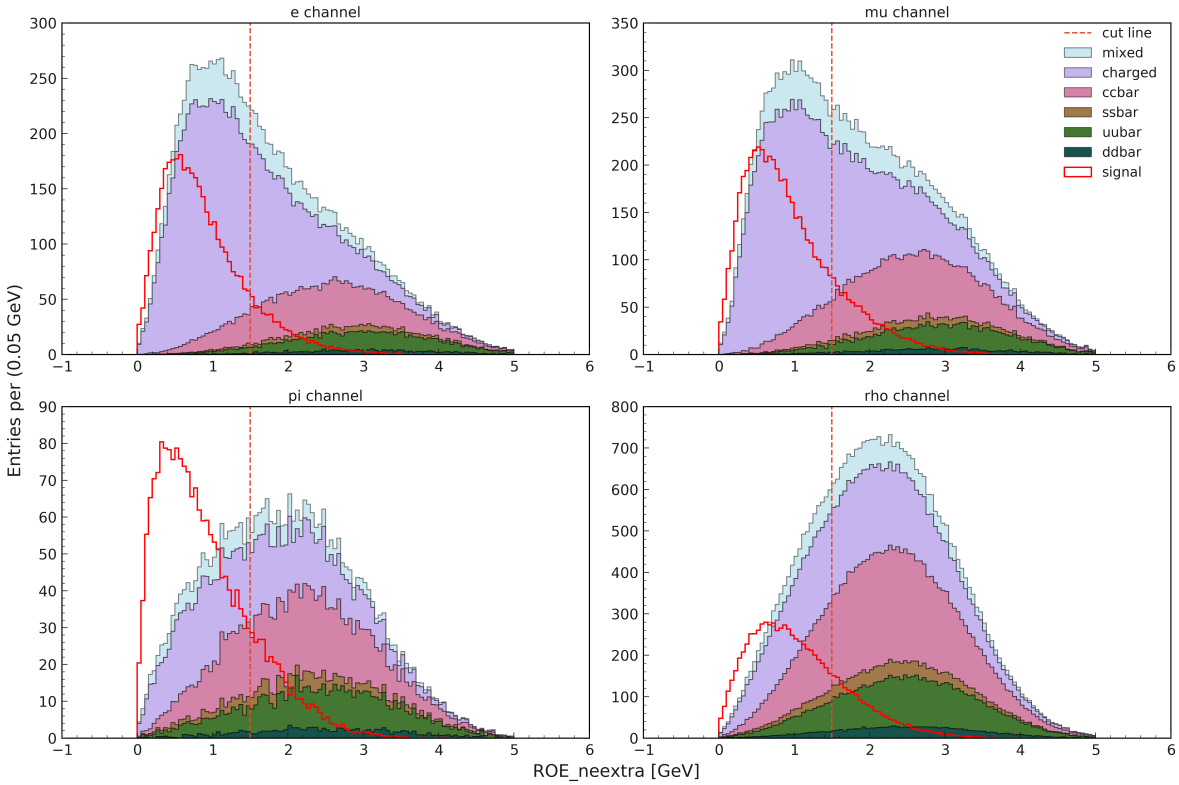

Figure 5.5: Histograms of the ROE extra energy, based on tau decay channel. The red dashed line indicates the cut line, for which we retain events below that value ($< 1.5\,GeV$). The generic MC is scaled to 200 fb^{-1}, and the signal MC is arbitrarily scaled.

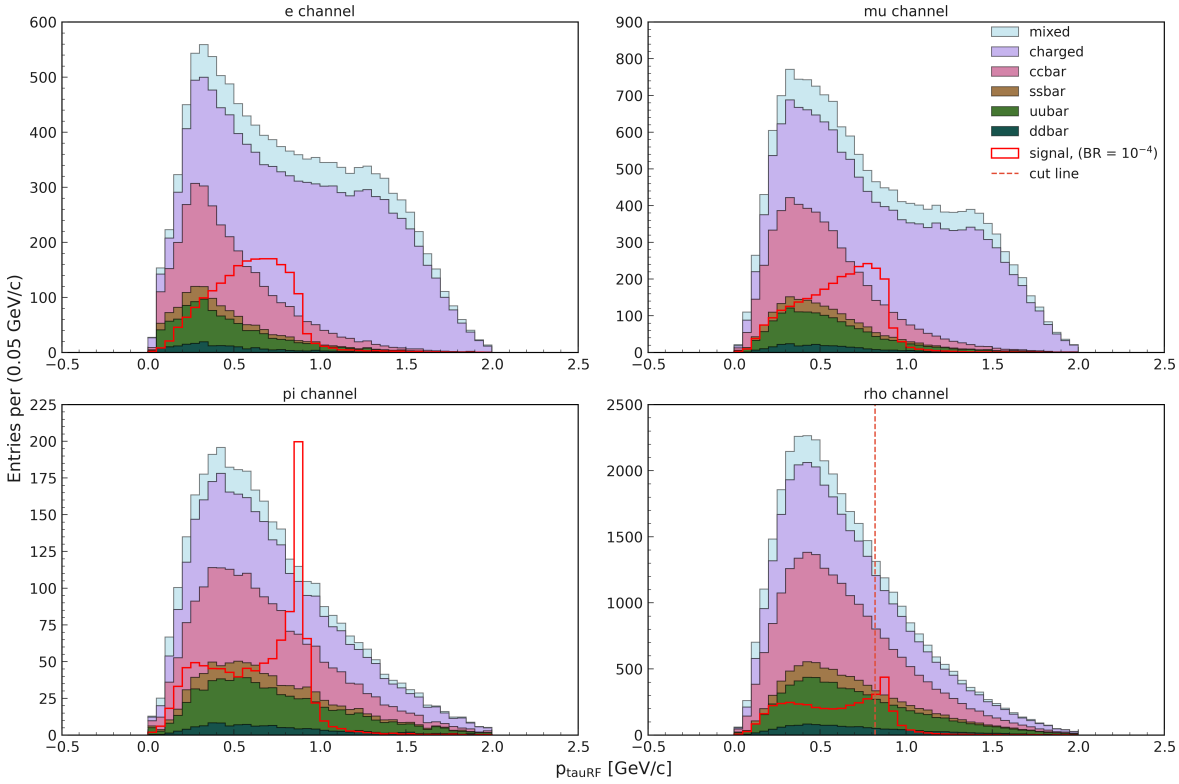

Figure 5.6: Histograms of the tau daughter momentum in the tau rest frame, based on tau decay channel. The red dashed line indicates the cut line, only applying to the rho channel, for which we retain events above that value ($> 0.82\,GeV/c$). The generic MC is scaled to 200 fb^{-1}, and the signal MC is arbitrarily scaled.

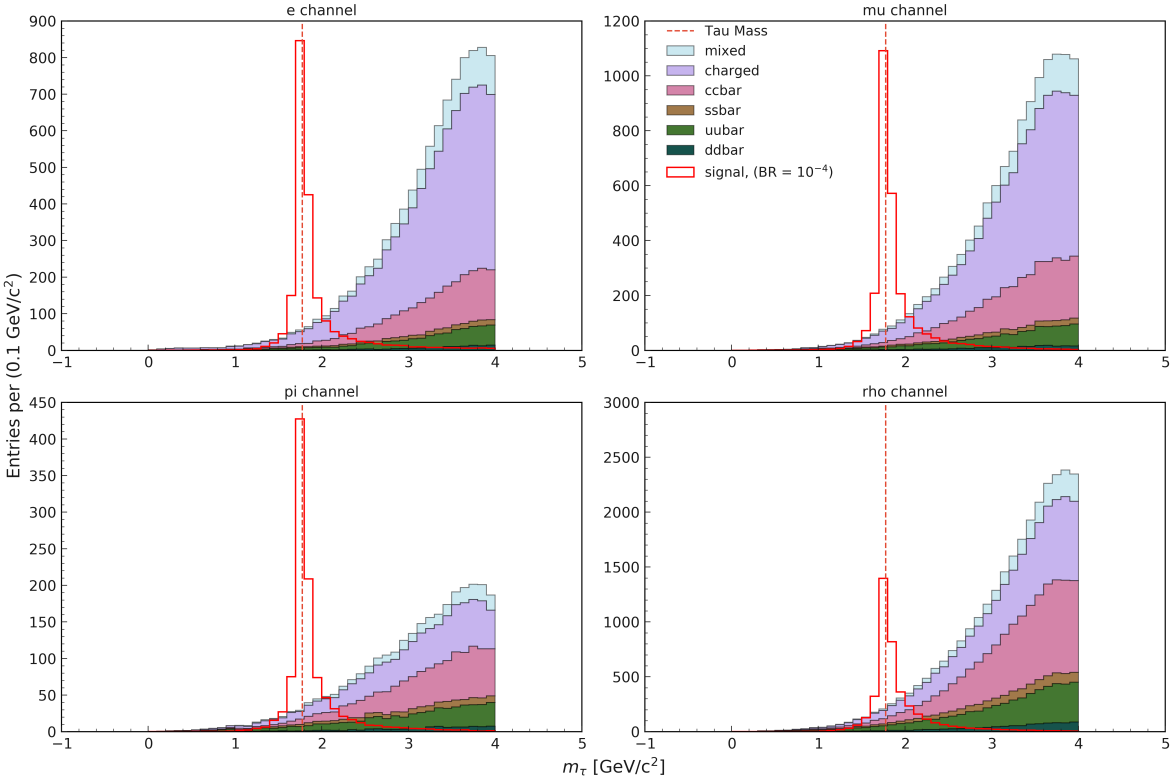

Figure 5.7: Histograms of the tau mass, based tau decay channel. The nominal tau mass is represented with the dashed line. The signal region is $\pm 60 MeV/c^2$ of the nominal tau mass. The generic MC is scaled to 200 fb^{-1}, and the signal MC is arbitrarily scaled.

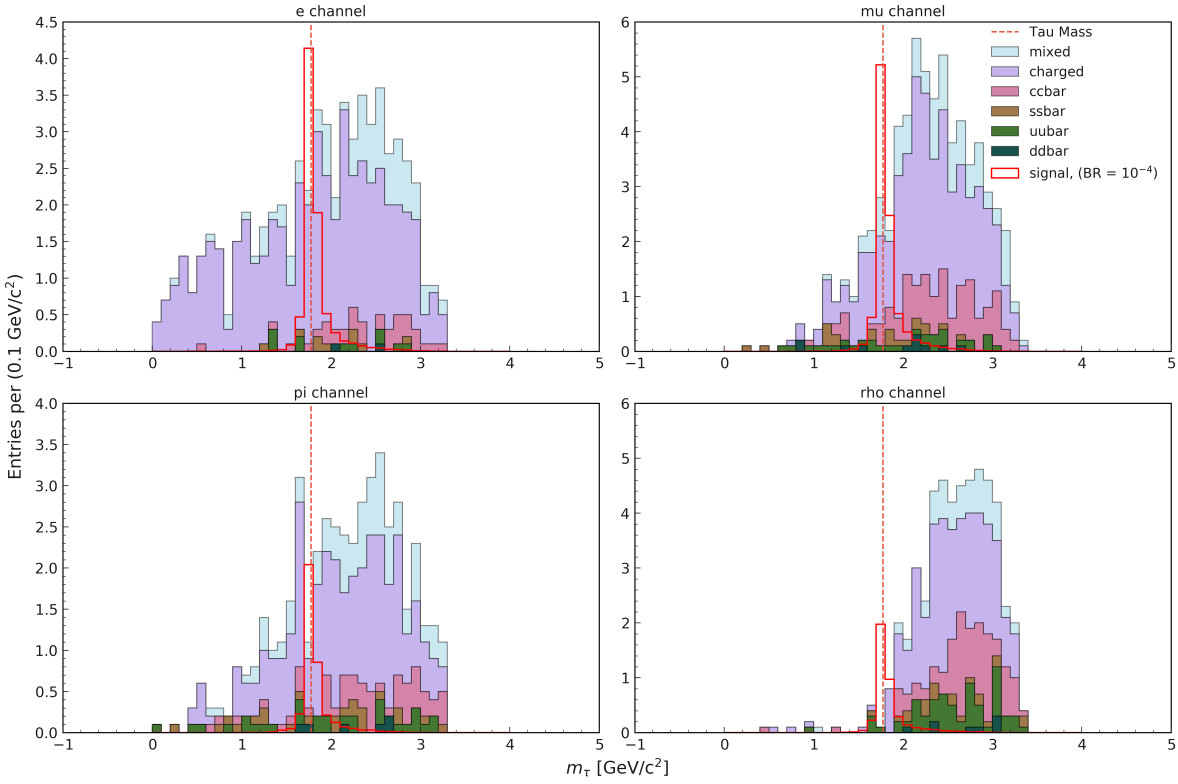

Figure 5.8: Histograms of the tau mass after all final cuts have been made, based tau decay channel. The nominal tau mass is represented with the dashed line. The signal region is $\pm 60 MeV/c^2$ of the nominal tau mass. Both generic and signal MC are scaled to 200 fb^{-1}, and assuming a branching fraction of 10^{-4} for the signal mode.

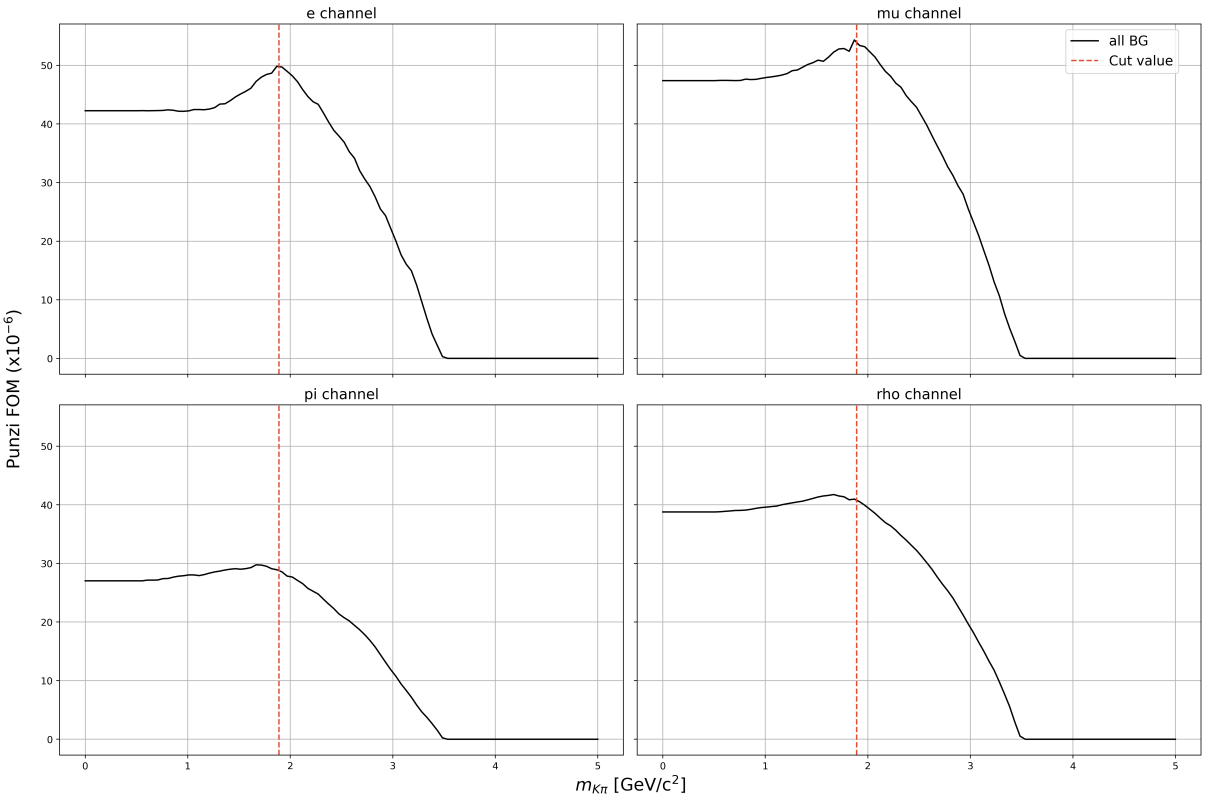

Figure 5.9: Plot of the Punzi FOM (defined in Equation 4.4) vs cut value (retaining events above the value) for the variable $m_{K\pi}$, with only events from within the tau mass signal region ($\pm 60 MeV/c^2$ the nominal tau mass). The dashed line represents the optimized cut value for this variable. The generic MC is scaled to 200 fb^{-1}.

Plots of the Punzi FOM vs cut value are also produced, made after the tau mass signal region has been selected, but before any of the final cuts. The FOM plot for $m_{K\pi}$ is shown in Figure 5.9 as an example. The Punzi FOM is optimized globally, and you can see that the cut value can be more optimized for some channels and slightly less optimized for others. Plots for the remaining variables are shown in Appendix B. Additionally, some of the optimizations might appear to be non-optimal in all channels, but that is because the variables were optimized in tandem, not one at a time. In other words, cutting on one variable will change what the optimization of another variable will be.

The results of the Punzi FOM optimization result in the following cut values:

- $B_{\mathrm{tag}}\ m_{\mathrm{bc}} > 5.27\,GeV/c^2$
- Primary electron electronID > 0.98
- ROE Extra Energy $< 1.5\,GeV$
- $m_{K\pi} > 1.89\,GeV/c^2$
- $0.1 < m_{\ell\ell} < 3.04$ or $m_{\ell\ell} > 3.14\,GeV/c^2$ (electron channel only)
- $p_{\pi^+} > 0.82\,GeV/c$, in the tau rest frame (rho channel only)

MC type	Initial Events	After FEI Skim	After Cuts (all)	After Cuts ($\tau \to e$)	After Cuts ($\tau \to \mu$)	After Cuts ($\tau \to \pi$)	After Cuts ($\tau \to \rho$)
Signal	50M	2.5M	37,785	11,640	14,803	5,731	5,611
Signal (BR = 10^{-4} and 200 fb^{-1})	21,588	1,082	16.3	5.0	6.4	2.5	2.4
All Generic (200 fb^{-1})	1,137M	64M	8.3	2.8	3.7	1.8	0
All Generic (2 ab^{-1})	11,374M	640M	83	28	37	18	0
Mixed	1,020M	81.5M	15	3	8	4	0
Charged	1,080M	128M	55	23	21	11	0
ccbar	2,658M	187M	7	1	5	1	0
ssbar	766M	31M	4	1	2	1	0
uubar	3,210M	171M	1	0	1	0	0
ddbar	802M	41M	1	0	0	1	0
taupair	1,838M	0	0	0	0	0	0

Table 5.1: Number of events remaining in signal and generic MC. The second row are signal events that have been corrected to 200 fb^{-1} if we assume a branching fraction of 10^{-4}. The third row is 2 ab^{-1} scaled to 200 fb^{-1}. The individual generic samples are equivalent to 2 ab^{-1}. The taupair sample does not survive the FEI skim at all.

The statistics of signal and generic MC are compiled in Table 5.1. We calculate the significance, Z, for a number of potential yields by first calculating the p-value of the $N_{signal} = 0$ hypothesis,

$$p = \sum_{n=n_{obs}}^{\infty} \frac{b^n}{n!} e^{-b}, \tag{5.1}$$

where b is the number of expected background events (b=8.3 for 200 fb^{-1}), and n is the number of observed events. Z is then calculated as

$$Z = \Phi^{-1}(1 - p), \tag{5.2}$$

Yield ($n_{\rm obs}$)	p-value	Significance (Z)
10	0.321	0.46
12	0.135	1.10
14	0.0439	1.71
16	0.0113	2.28
18	2.35e-3	2.83
20	4.00e-4	3.35
22	5.67e-5	3.86
24	6.79e-6	4.35
26	6.96e-7	4.83
28	6.16e-8	5.29

Table 5.2: Table for significance based on potential yields, calculated from Equations 5.1 and 5.2. Z=1.645 corresponds to a 90% confidence level. With $n_{\rm obs} = b + s$, and b=8.3 for a dataset of 200 fb^{-1}, we would need to expect s=6 to reach a 90% CL. We expect s=6 if the branching fraction of our signal mode is roughly 4x10^{-5}. The current world limit is set to $<$1.5x10^{-5} at a 90% CL by BaBar.

where Φ^{-1} is the inverse of the cumulative distribution of the standard Gaussian. Table 5.2 compiles the p-value and significance for a variety of different n. Since $n = b + s$, where s is the number of expected signal events, we can check the expected significance for our dataset and for various signal branching fractions. Note that a confidence level of 90% corresponds to Z=1.645, and that BaBar set the current limit on this signal mode at < 1.5x10^{-5} with a 90% CL. It is also worth reminding the reader that BaBar worked with a dataset of 429 fb^{-1}. Referring to Table 5.1, for a 200 fb^{-1} dataset we expect b=8.3, therefore we need an expected signal yield of approximately s=6, assuming a branching fraction of 1.5x10^{-5}, in order to compete with the current limits, ignoring uncertainties for the moment. Currently, we expect a signal yield of roughly 2.5 for 200 fb^{-1} and a branching fraction of 1.5x10^{-5}. If we assume a branching fraction of 10^{-4}, we expect s=16.3 (Table 5.1) giving n=24.6, which we can exclude with $Z > 4.35$. The branching fraction which we can exclude with 90% confidence, corresponding to s=6, is just below 4x10^{-5}.

5.2 Data/MC Comparison

Here we show some results from our data/MC comparison. These comparisons help determine the level of data/MC agreement in the current dataset in order for us to estimate the systematic uncertainties in the background and signal efficiency estimates. Additionally, this helps us identify aspects of our analysis which may need further study in order to make this a viable scientific result for publication. A 62.8 fb^{-1} dataset is used, along with 2 ab^{-1} of generic MC which is then scaled to 62.8 fb^{-1}. Due to this scaling, we expect some smoothing of the MC compared to the data, but we are more interested in the overall distribution. We exclude all events that fall within $\pm 175\, MeV/c^2$ of the tau mass in order to remain unbiased. The following figures are for the electron channel only, and the same figures for the other channels can be found in Appendix C. In Figure 5.10 we plot B_{tag} m_{bc}. There is some disagreement between data and MC, with the data having a shift towards the lower values and a weaker peak. In Figure 5.11 we plot B_{tag} ΔE, which has reasonable agreement, although the pull indicates a sinusoidal pattern showing that the sidebands are slightly increased and the peak is lower. In Figures 5.12 and 5.13 the primary electron momentum and the charged tau daughter momentum, respectively, are plotted in the CMS frame. These both show a similar pattern, having data significantly exceed MC in the peaking region, around 0.2-0.3 GeV/c.

These discrepancies could come from a variety of sources, but are likely tracking and vertex performance related. Particle ID corrections could also help account for the disagreements. Overall, these discrepancies are well studied and can be corrected for, although applying those corrections are beyond the scope of this thesis. In the systematic uncertainty section below, these issues will be discussed in more detail.

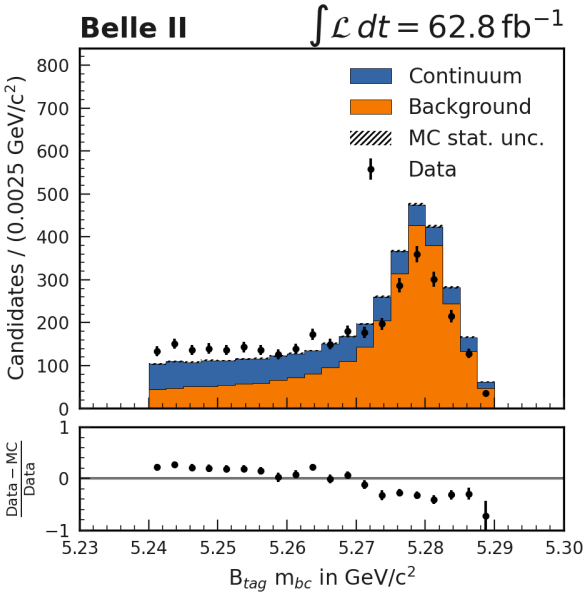

Figure 5.10: Data vs generic MC for B_{tag} m_{bc} in the e channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

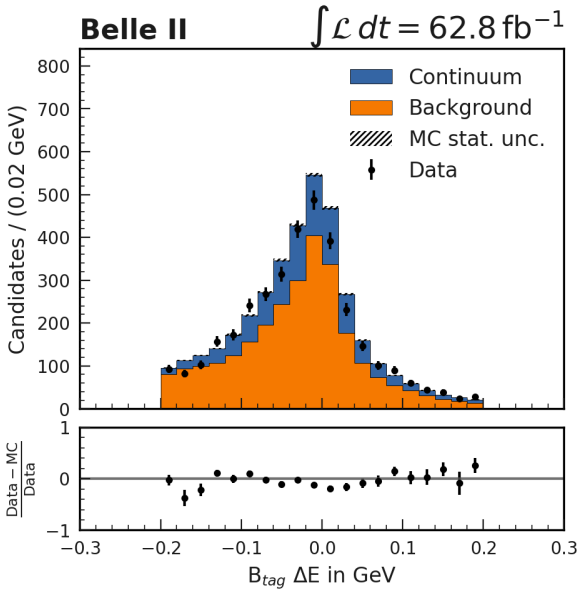

Figure 5.11: Data vs generic MC for B_{tag} ΔE in the e channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

5.3 Uncertainties

5.3.1 Statistical

The uncertainty for this analysis is dominated by statistics at the dataset size of 200 fb^{-1}. Since we estimate 8.3 background events for this dataset, the uncertainty on the mean of the underlying Poisson distribution is $\sqrt{8.3} \approx 2.9$. If the signal mode has a branching ratio of 10^{-5} (just below the current limit), we expect 1.6 signal events. Therefore the statistical uncertainty alone makes it difficult to compete with current limits with a dataset of 200 fb^{-1}. The existing limits are based on more than 2x the integrated luminosity, so this is not surprising.

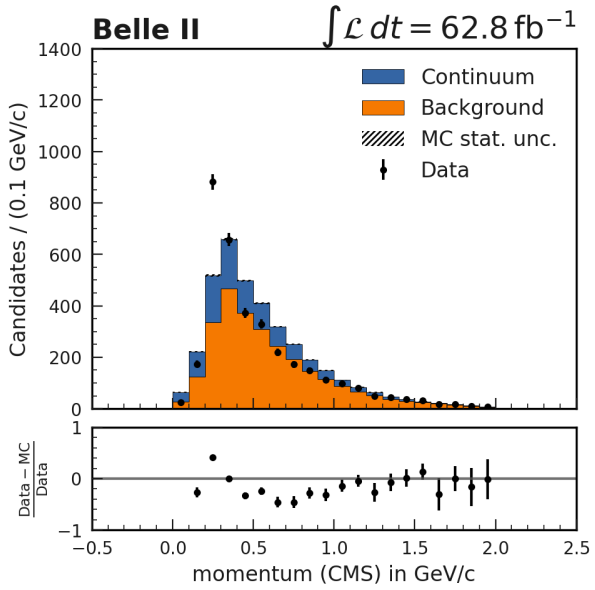

Figure 5.12: Data vs generic MC for the primary electron momentum in the CMS frame, in the e channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

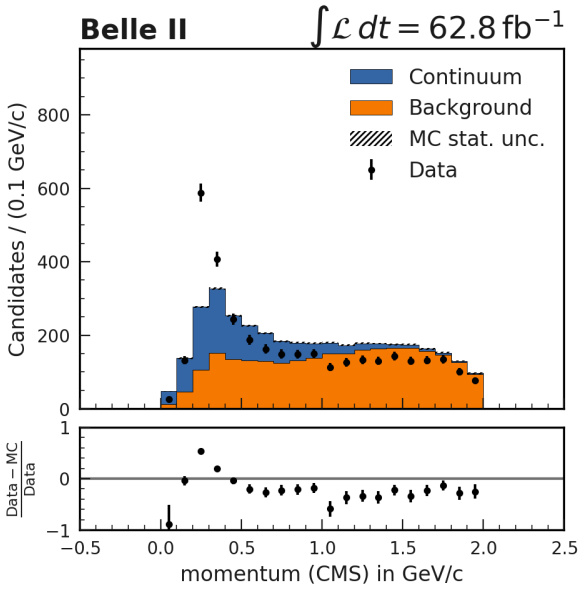

Figure 5.13: Data vs generic MC for the charged tau daughter momentum in the CMS frame, in the e channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

5.3.2 Systematics

Tracking and Vertex Performance The main tracking and vertex systematics relating to this analysis are the track momentum scaling, slow pion tracking efficiency, mid to high momentum tracking efficiency, and boost vector scaling (relating to the 4-momentum of the $\Upsilon(4S)$). The slow pion tracking efficiency covers charged pion tracks between 0.05 and 0.20 GeV/c. It has been studied and a scale factor correction can be applied based on momentum, with an associated uncertainty (e.g. in our dataset, for p in range [0.05,0.12) GeV/c, the scale factor is 0.990±0.068). Mid to high momentum tracking efficiency has likewise been studied on a per-dataset basis and a systematic uncertainty of 0.69% per track can be applied to our generic MC. Track momentum scaling applies a global scale factor to the data, and an associated systematic uncertainty can be evaluated

by varying the scale factor within a small range around the recommended value. There is also tracking momentum resolution correction that involves a smearing factor, which is still under study. The boost vector correction and associated uncertainty are also still under study but can help correct data/mc discrepancies.

Lepton ID The lepton ID corrections have also been studied, paramaterized by momentum, p, and polar angle, θ, and are charge-dependent. Figure 5.14 shows the data/MC correction factors for electron efficiency in the different (p, θ) bins, with an electronID $>$ 0.9, for our dataset. Large polar angles tend to require the largest corrections.

Hadron ID Likewise, hadron ID corrections for kaons and pions are also studied. Similar correction tables as presented for the lepton ID are also available for hadron ID, also presented in terms of polar angle and momentum. The corrections also tend to be higher for large angles, and also for high momentum.

5.4 Discussion

The results obtained from the sensitivity study of $B^+ \to K^+\tau^+e^-$ show a promising future for this mode at Belle II. The study was performed under an assumed dataset of 200 fb^{-1} and utilized the Punzi FOM to optimize the search. At this dataset size, our analysis is unlikely to be competitive with the current limits set by BaBar. However, it has been shown in Table 5.2 that we would be within one order of magnitude of BaBar. Data/MC comparisons and systematics have been qualitatively discussed and no major issues are present. Improvements on Data/MC agreement are expected with the proper corrections applied. Future datasets and MC campaigns will naively have improvements over what has been presented here. The FEI can also be expected to improve over future MC campaigns. The full Belle II dataset is expected to be 250 times larger than the dataset studied in this thesis, and so a major improvement over the current branching fraction limit can

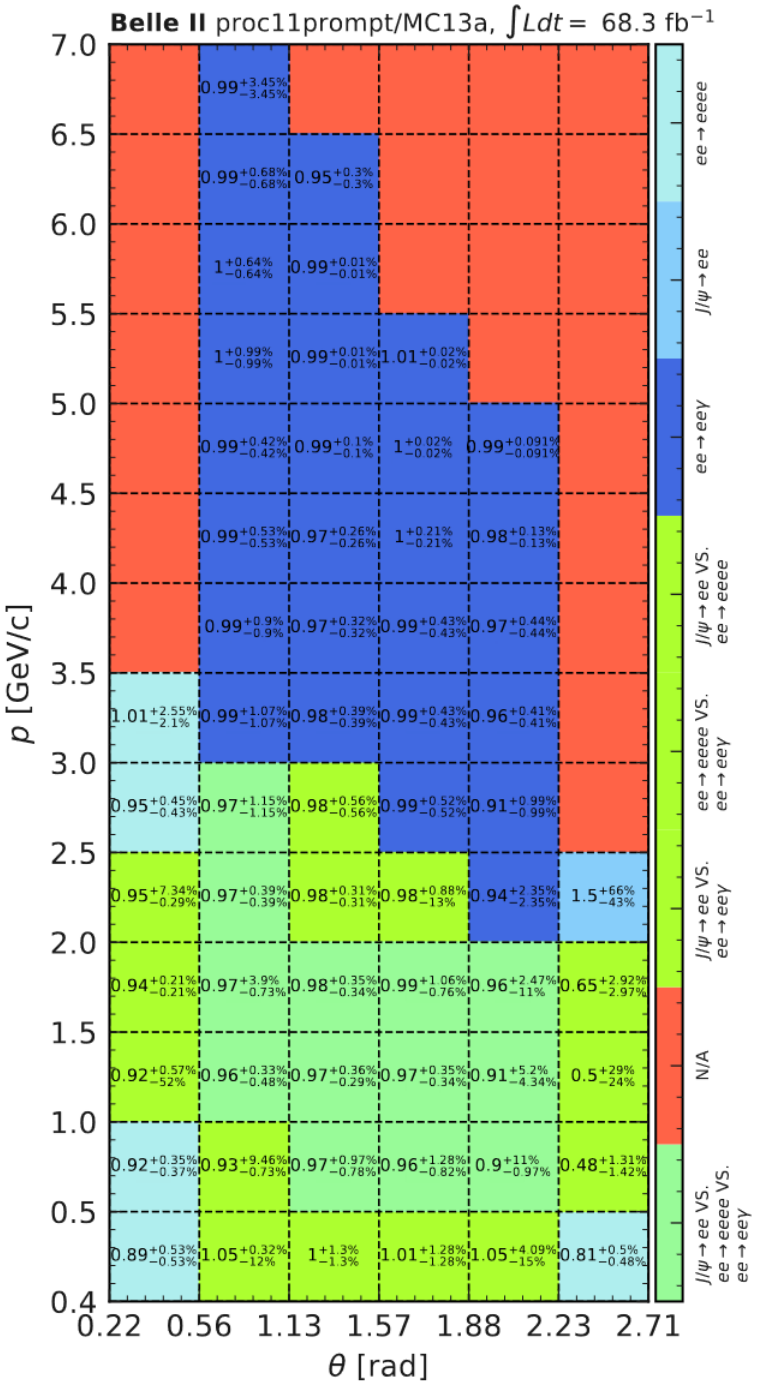

Figure 5.14: Electron ID correction table for negatively charged electrons, with electron ID > 0.9, binned by momentum, p, and polar angle, θ. Taken from the internal Belle II Confluence page [56].

be expected. The optimized cut values in this thesis will also vary as we increase the dataset, and in general the cuts will tend to tighten as that dataset size goes up.

5.4.1 Future Prospects

The sensitivity study performed in this thesis lays the ground work for a full $B \to K\tau\ell$ analysis at Belle II. With increasing Belle II performance, including that of the FEI, coupled with the very large statistics of the Belle II target dataset, the search for these New Physics decay modes will be able to reach novel sensitivities. This analysis focused on hadronic tagging methods, but semi-leptonic or semi-inclusive tagging methods should also be explored. Applying other analysis techniques such as an unbinned likelihood analysis, or using machine learning to better discriminate background should also be examined in order to maximize the sensitivity of these searches. Other NP CLFV modes of interest are $B \to \pi\tau\ell$ and $B \to \tau\ell$, which can also be analysed in depth at Belle II.

Chapter 6

Conclusion

In this thesis we performed a sensitivity study of the decay mode $B^+ \to K^+\tau^+e^-$. We based the study on a 200 fb^{-1} dataset, which is roughly what will be available at Belle II by the end of the 2021. By optimizing the Punzi Figure of Merit, we found variables at their associated cut values that were best suited to maximize the sensitivity. The outcome of this sensitivity study is promising for publishing a competitive result on the branching fraction of $B^+ \to K^+\tau^+e^-$ on the timescale of the next couple of years.

Bibliography

[1] Sheldon L. Glashow, Diego Guadagnoli, and Kenneth Lane. Lepton flavor violation in B decays? *Physical Review Letters*, 114(9), Mar 2015.

[2] Makoto Kobayashi and Toshihide Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[3] K. Abe et al. Observation of large CP violation in the neutral B meson system. *Physical Review Letters*, 87(9), Aug 2001.

[4] B. Aubert et al. Observation of CP violation in the B^0 meson system. *Physical Review Letters*, 87(9), Aug 2001.

[5] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. Remarks on the Unified Model of Elementary Particles. *Progress of Theoretical Physics*, 28(5):870–880, 11 1962.

[6] J. P. Lees et al. Measurement of an excess of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ decays and implications for charged higgs bosons. *Physical Review D*, 88(7), Oct 2013.

[7] M. Huschle et al. Measurement of the branching ratio of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ relative to $\overline{B} \to D^{(*)}\ell^-\overline{\nu}_\ell$ decays with hadronic tagging at belle. *Physical Review D*, 92(7), Oct 2015.

[8] S. Hirose et al. Measurement of the τ lepton polarization and R(D*) in the decay $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$. *Physical Review Letters*, 118(21), May 2017.

[9] R. Aaij et al. Measurement of the ratio of branching fractions B($\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$)/B($\overline{B} \to D^{(*)}\mu^-\overline{\nu}_\mu$). *Physical Review Letters*, 115(11), Sep 2015.

[10] R. Aaij et al. Measurement of the ratio of the $B \to D^{*-}\tau^+\nu_\tau$ and $B \to D^{*-}\mu^+\nu_\mu$ branching fractions using three-prong τ-lepton decays. *Physical Review Letters*, 120(17), Apr 2018.

[11] Y. Amhis et al. Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018. *The European Physical Journal C*, 81(3), Mar 2021.

[12] Ying Li and Cai-Dian Lü. Recent anomalies in B physics. *Science Bulletin*, 63(5):267–269, Mar 2018.

[13] R. Aaij et al. Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays. *Physical Review Letters*, 113(15), Oct 2014.

[14] R. Aaij et al. Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays. *Journal of High Energy Physics*, 2017(8), Aug 2017.

[15] J.-T. Wei et al. Measurement of the differential branching fraction and forward-backward asymmetry for $B \to K^*\ell^+\ell^-$. *Physical Review Letters*, 103(17), Oct 2009.

[16] J. P. Lees et al. Measurement of branching fractions and rate asymmetries in the rare decays $B \to K^{(*)}\ell^+\ell^-$. *Physical Review D*, 86(3), Aug 2012.

[17] R. Aaij et al. Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$. *Journal of High Energy Physics*, 2015(9), Sep 2015.

[18] R. Aaij et al. Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^1 of integrated luminosity. *Journal of High Energy Physics*, 2016(2), Feb 2016.

[19] A. Abdesselam et al. Angular analysis of $B^0 \to K^*(892)^0\ell^+\ell^-$. In *LHC Ski 2016: A First Discussion of 13 TeV Results*, Apr 2016.

[20] The ATLAS Collaboration. Angular analysis of $B_d^0 \to K^* \mu^+ \mu^-$ decays in pp collisions at $\sqrt{s} = 8\,TeV$ with the ATLAS detector. Technical report, CERN, Geneva, Apr 2017.

[21] Elena Graverini. Flavour anomalies: a review. *Journal of Physics: Conference Series*, 1137:012025, Jan 2019.

[22] CMS Collaboration. Measurement of the P_1 and P_5' angular parameters of the decay $\mathrm{B}^0 \to \mathrm{K}^{*0} \mu^+ \mu^-$ in proton-proton collisions at $\sqrt{s} = 8\ \mathrm{TeV}$. Technical report, CERN, Geneva, 2017.

[23] Damir Bečirević et al. Palatable Leptoquark Scenarios for Lepton Flavor Violation in Exclusive $b \to s \ell_1 \ell_2$ modes. *Journal of High Energy Physics*, 11:035, 2016.

[24] A. Angelescu et al. Closing the window on single leptoquark solutions to the B-physics anomalies. *Journal of High Energy Physics*, 2018(10), Oct 2018.

[25] Damir Bečirević et al. Scalar leptoquarks from grand unified theories to accommodate the B-physics anomalies. *Physical Review D*, 98(5), Sep 2018.

[26] Jacky Kumar, David London, and Ryoutaro Watanabe. Combined explanations of the $b \to s \mu^+ \mu^-$ and $b \to c \tau^- \overline{\nu}$ anomalies: A general model analysis. *Physical Review D*, 99(1), Jan 2019.

[27] Marzia Bordone et al. Low-energy signatures of the PS^3 model: from B-physics anomalies to LFV. *Journal of High Energy Physics*, 10:148, 2018.

[28] Julian Heeck and Daniele Teresi. Pati-salam explanations of the b-meson anomalies. *Journal of High Energy Physics*, 2018(12), Dec 2018.

[29] Diego Guadagnoli, Méril Reboud, and Olcyr Sumensari. A gauged horizontal SU(2) symmetry and $R_{K^{(*)}}$. *Journal of High Energy Physics*, 2018(11), Nov 2018.

[30] S. Fajfer and N. Košnik. Vector leptoquark resolution of r_{K} and $r_d^{(*)}$ puzzles. *Physics Letters B*, 755:270–274, Apr 2016.

[31] F.F. Deppisch et al. Leptoquark patterns unifying neutrino masses, flavor anomalies, and the diphoton excess. *Physical Review D*, 94(1), Jul 2016.

[32] Lorenzo Calibbi, Andreas Crivellin, and Tianjun Li. Model of vector leptoquarks in view of the b-physics anomalies. *Physical Review D*, 98(11), Dec 2018.

[33] Andreas Crivellin. SUSY flavour / Flavour sector in BSM scenarios. *PoS*, FPCP2017:030, 2017.

[34] Ilja Doršner et al. The role of the S_3 gut leptoquark in flavor universality and collider searches. *Journal of High Energy Physics*, 2017(10), Oct 2017.

[35] J. P. Lees et al. Search for the decay modes $B^{\pm} \to h^{\pm}\tau\ell$. *Physical Review D*, 86(1), Jul 2012.

[36] R. Aaij et al. Search for the lepton flavour violating decay $B^+ \to K^+\mu^-\tau^+$ using B^{*0}_{s2} decays. *Journal of High Energy Physics*, 2020(6), Jun 2020.

[37] T. Abe et al. Belle II Technical Design Report. Nov 2010. https://arxiv.org/abs/1011.0352.

[38] Thomas E. Browder and Klaus Honscheid. B mesons. *Progress in Particle and Nuclear Physics*, 35:81–219, Jan 1995.

[39] SuperKEKB. Luminosity projection, 2021. https://www-superkekb.kek.jp/Luminosity_projection.html.

[40] Kazunori Akai, Kazuro Furukawa, and Haruyo Koiso. SuperKEKB collider. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 907:188–199, Nov 2018.

[41] Yusuke Suetsugu et al. Superkekb main ring vacuum system status until the end of phase-2 commissioning. *Journal of Vacuum Science & Technology A*, 37(2):021602, 2019.

[42] E. Kou et al. The Belle II physics book. *Progress of Theoretical and Experimental Physics*, 2019(12), Dec 2019.

[43] M. Friedl et al. The Belle II silicon vertex detector. *Physics Procedia*, 37:867–873, 2012. Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2011).

[44] T.V. Dong et al. Calibration and alignment of the Belle II central drift chamber. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 930:132–141, 2019.

[45] S. Sandilya. Particle Identification with the TOP and ARICH Detectors at Belle II. *Springer Proc. Phys.*, 203:563–565, 2018.

[46] L. Burmistrov. Belle II aerogel rich detector, 2019. 15th Vienna Conference on Instrumentation.

[47] Umberto Tamponi et al. Data taking. https://b2-master.belle2.org/software/development/sphinx/online_book/fundamentals/02-datataking.html.

[48] P.A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020.

[49] Anders Ryd et al. EvtGen: A Monte Carlo Generator for B-Physics. 5 2005. https://evtgen.hepforge.org/doc/EvtGenGuide.pdf.

[50] S. Agostinelli et al. GEANT4–a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250–303, 2003.

[51] N. Davidson, T. Przedzinski, and Z. Was. PHOTOS Interface in C++; Technical and Physics Documentation. *Comput. Phys. Commun.*, 199:86–101, 2016.

[52] T. Keck et al. The full event interpretation. *Computing and Software for Big Science*, 3(1), Feb 2019.

[53] Kuhr T. Hara, T. and Y. Ushiroda. Belle II coordinate system and guideline of belle ii numbering scheme, 2011.

[54] Geoffrey C. Fox and Stephen Wolfram. Event Shapes in e+ e- Annihilation. *Nucl. Phys. B*, 149:413, 1979. [Erratum: Nucl.Phys.B 157, 543 (1979)].

[55] Giovanni Punzi. Sensitivity of searches for new signals and its optimization. *eConf*, C030908:MODT002, 2003.

[56] Marco Milesi. Lepton ID corrections. https://confluence.desy.de/display/BI/Recommendations+ +Moriond+2021.

Appendix A

Histograms

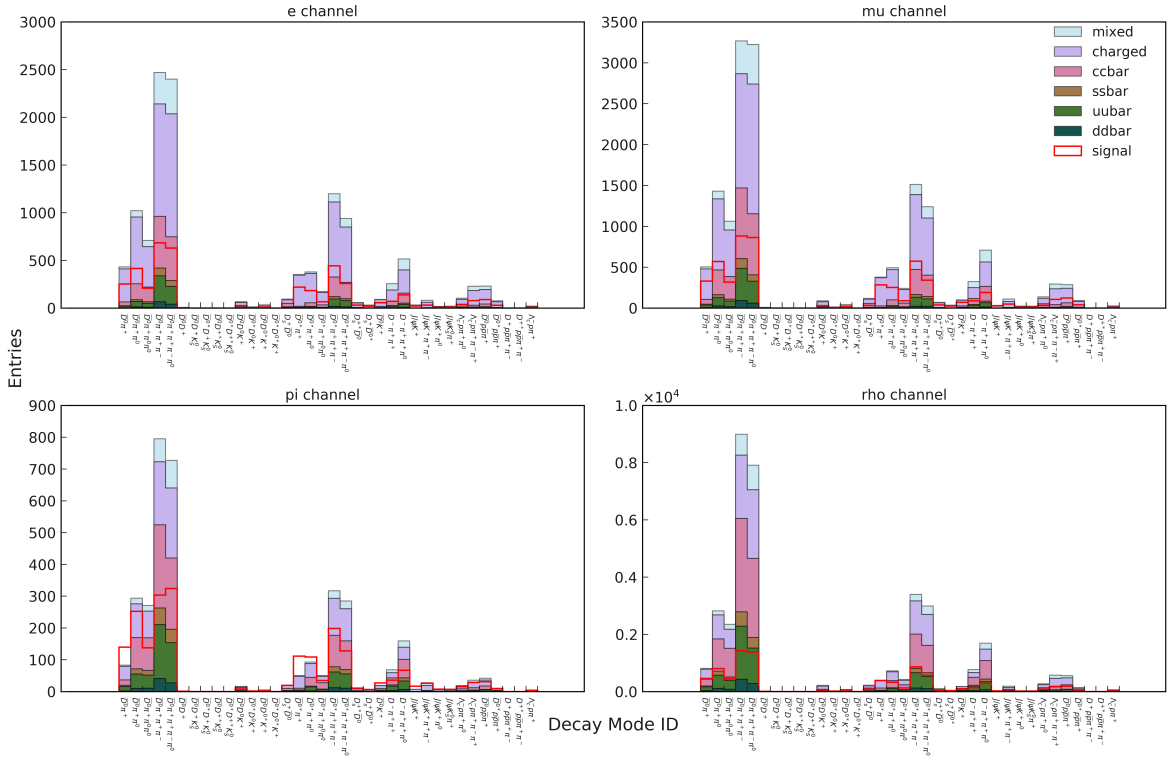

Figure A.1: Histograms of B_{tag} FEI mode, based on tau channel. The generic MC is scaled to 200 fb^{-1}, and the signal MC is arbitrarily scaled.

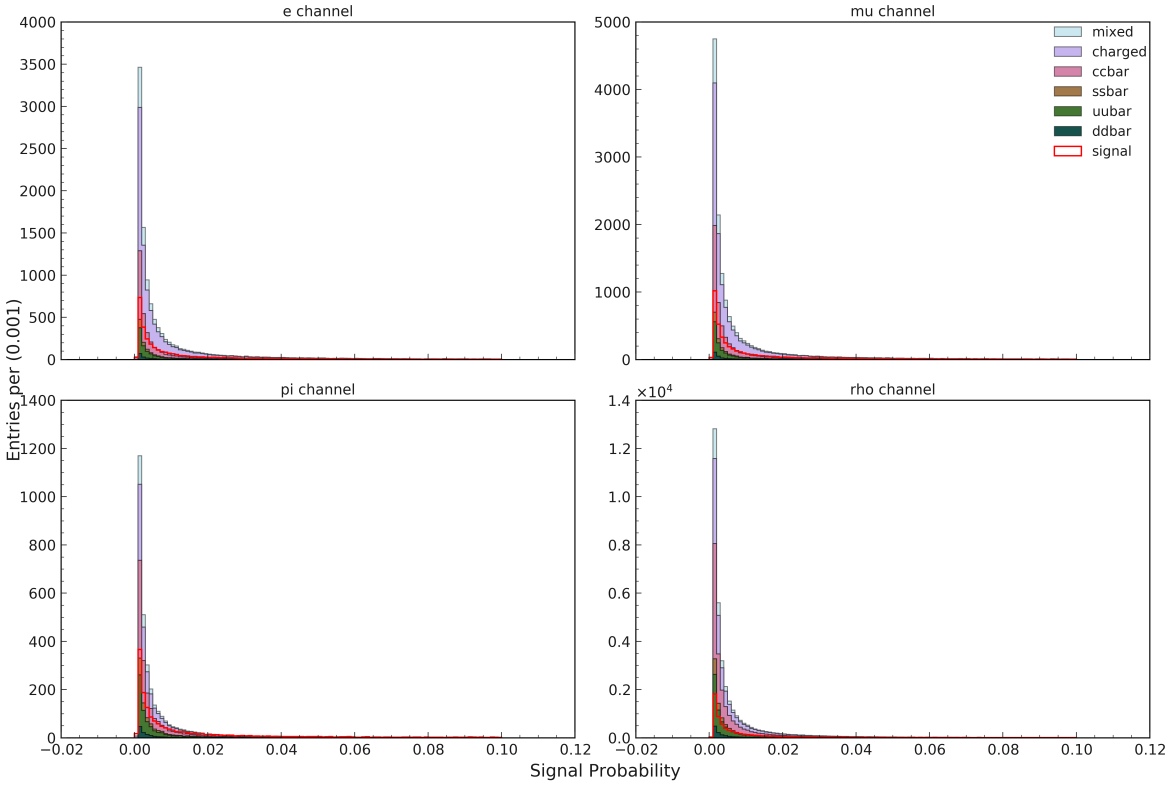

Figure A.2: Histograms of Signal Probability, based on tau decay channel. The generic MC is scaled to 200 fb^{-1}, and the signal MC is arbitrarily scaled.

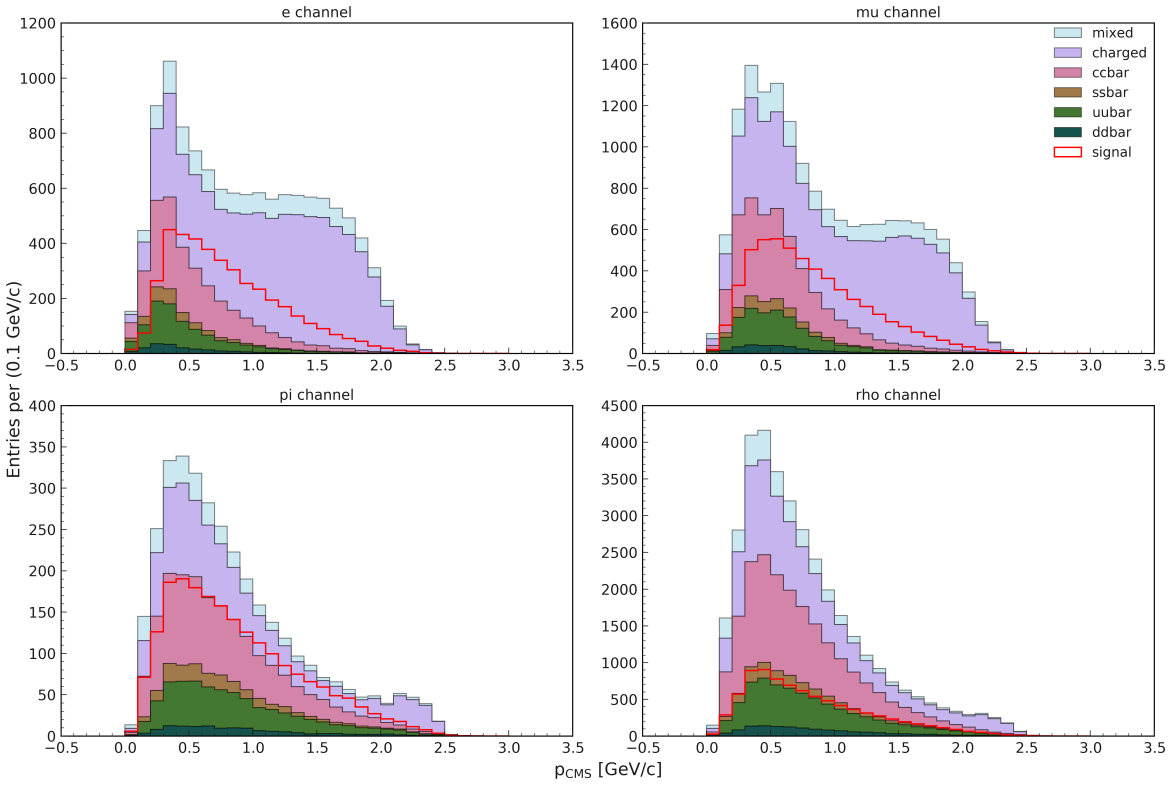

Figure A.3: Histograms of the tau daughter momentum in the CMS frame, based on tau decay channel. The generic MC is scaled to 200 fb^{-1}, and the signal MC is arbitrarily scaled.

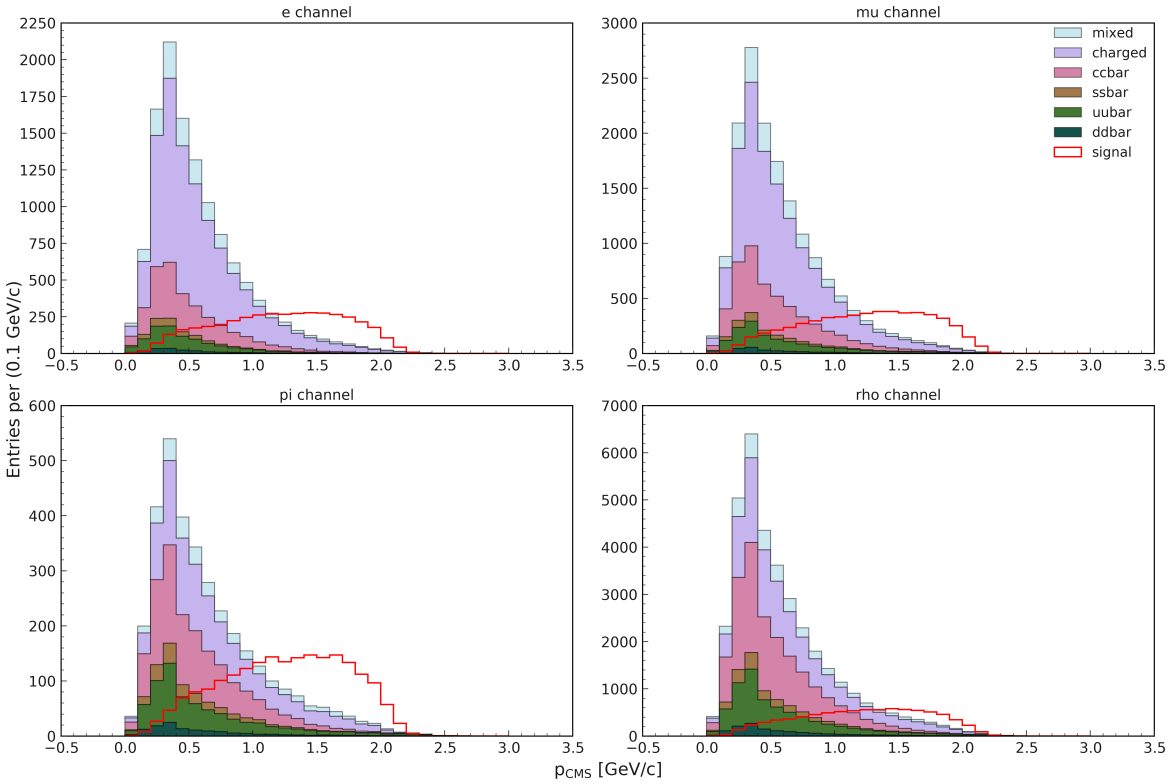

Figure A.4: Histograms of the primary electron momentum in the CMS frame, based on tau decay channel. The generic MC is scaled to 200 fb^{-1}, and the signal MC is arbitrarily scaled.

Appendix B

Punzi FOM Plots

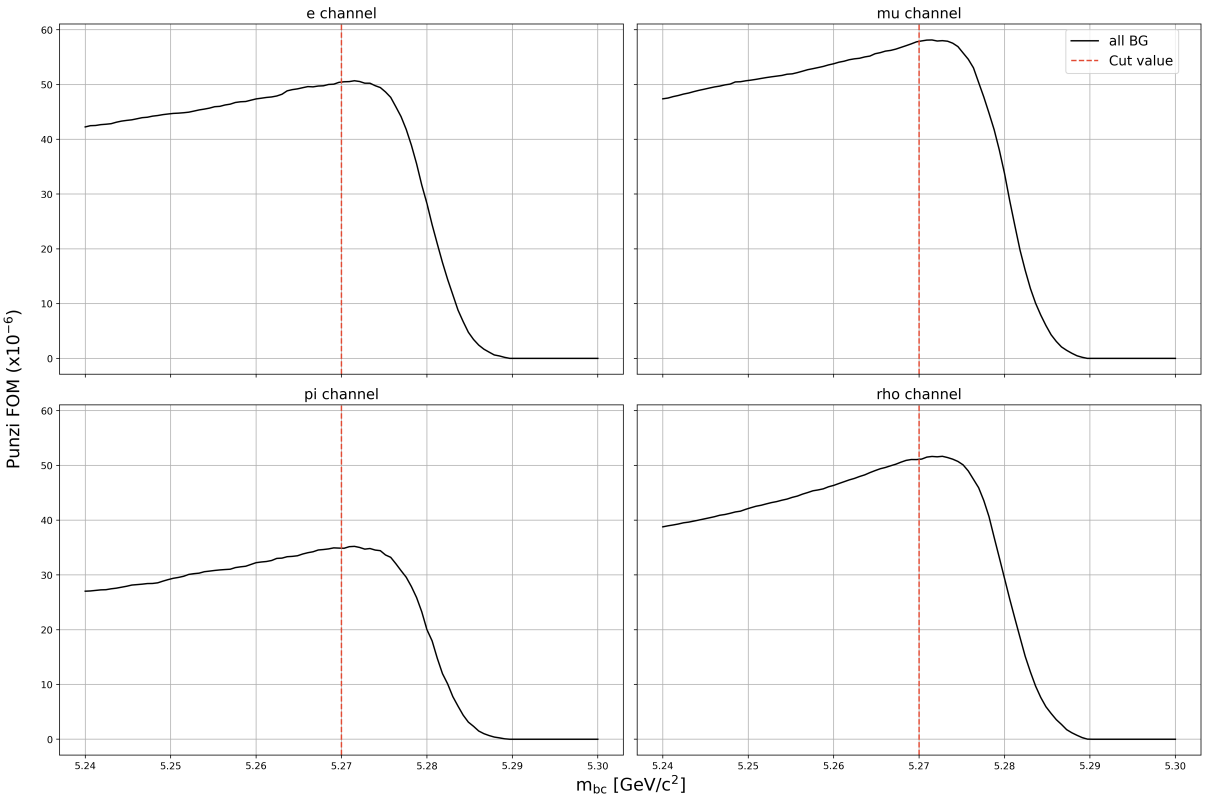

Figure B.1: Plot of the Punzi FOM (defined in Equation 4.4) vs cut value (retaining events above the value) for the variable $B_{\mathrm{tag}}\ m_{\mathrm{bc}}$, with only events from within the tau mass signal region ($\pm 60 MeV/c^2$ the nominal tau mass). The dashed line represents the optimized cut value for this variable. The generic MC is scaled to 200 fb^{-1}.

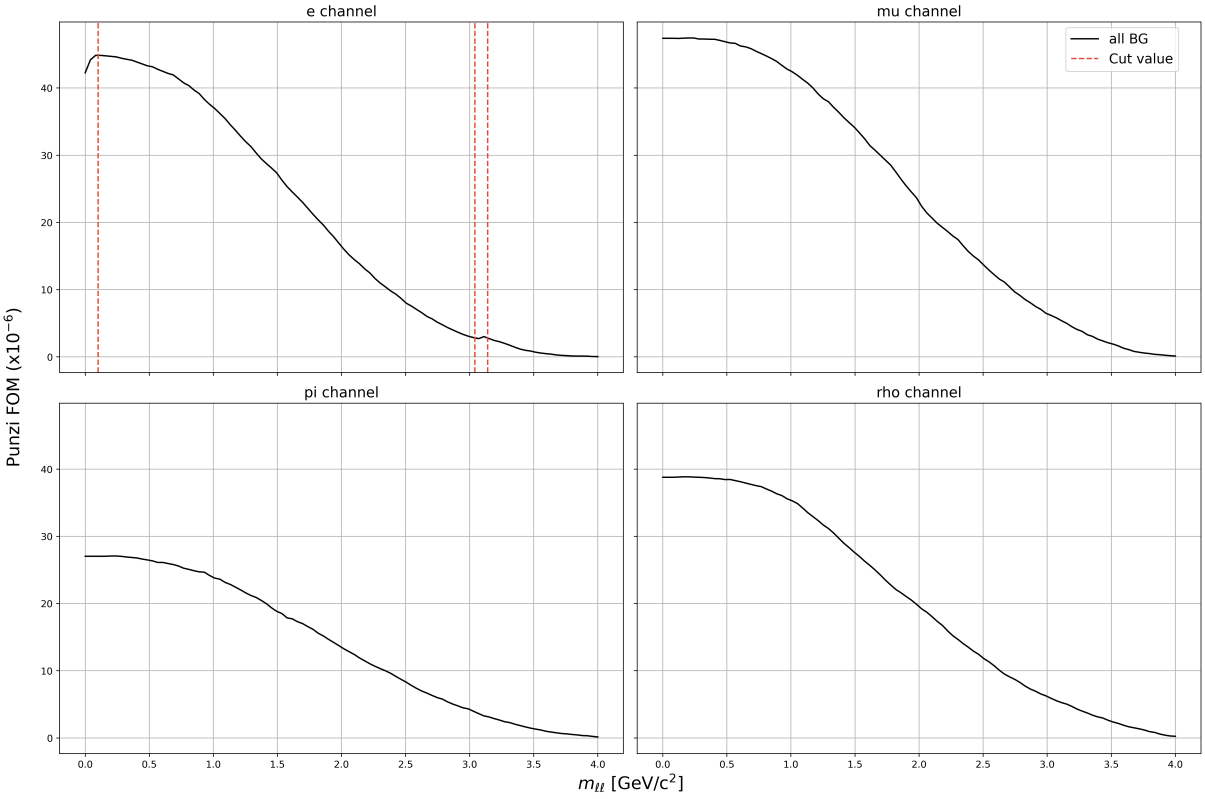

Figure B.2: Plot of the Punzi FOM (defined in Equation 4.4) vs cut value (retaining events above the value) for the variable $m_{\ell\ell}$, with only events from within the tau mass signal region ($\pm 60 MeV/c^2$ the nominal tau mass). The dashed lines represent the optimized cut values for this variable, which only apply to the electron channel. The events below the first line are cut, and the events in between the second and third line are also cut. The generic MC is scaled to 200 fb^{-1}.

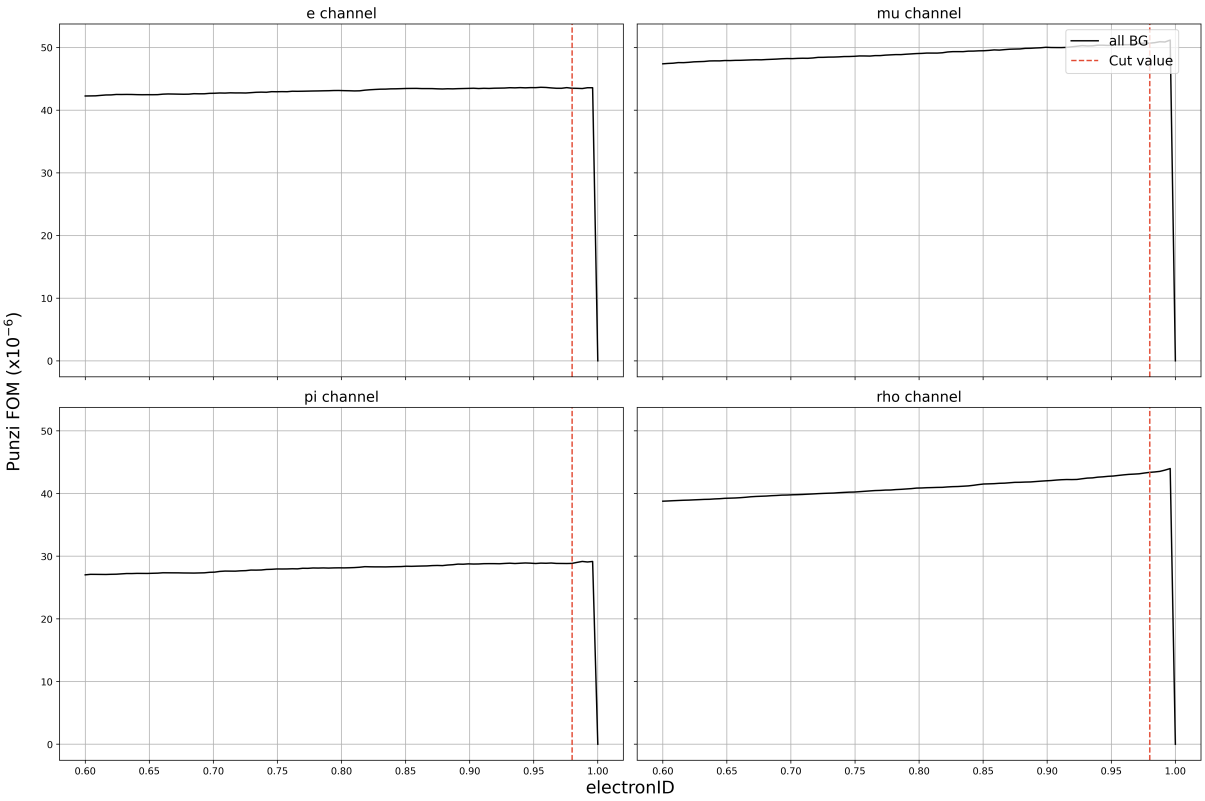

Figure B.3: Plot of the Punzi FOM (defined in Equation 4.4) vs cut value (retaining events above the value) for the primary electron electronID, with only events from within the tau mass signal region ($\pm 60 MeV/c^2$ the nominal tau mass). The dashed line represents the optimized cut value for this variable. The generic MC is scaled to 200 fb^{-1}.

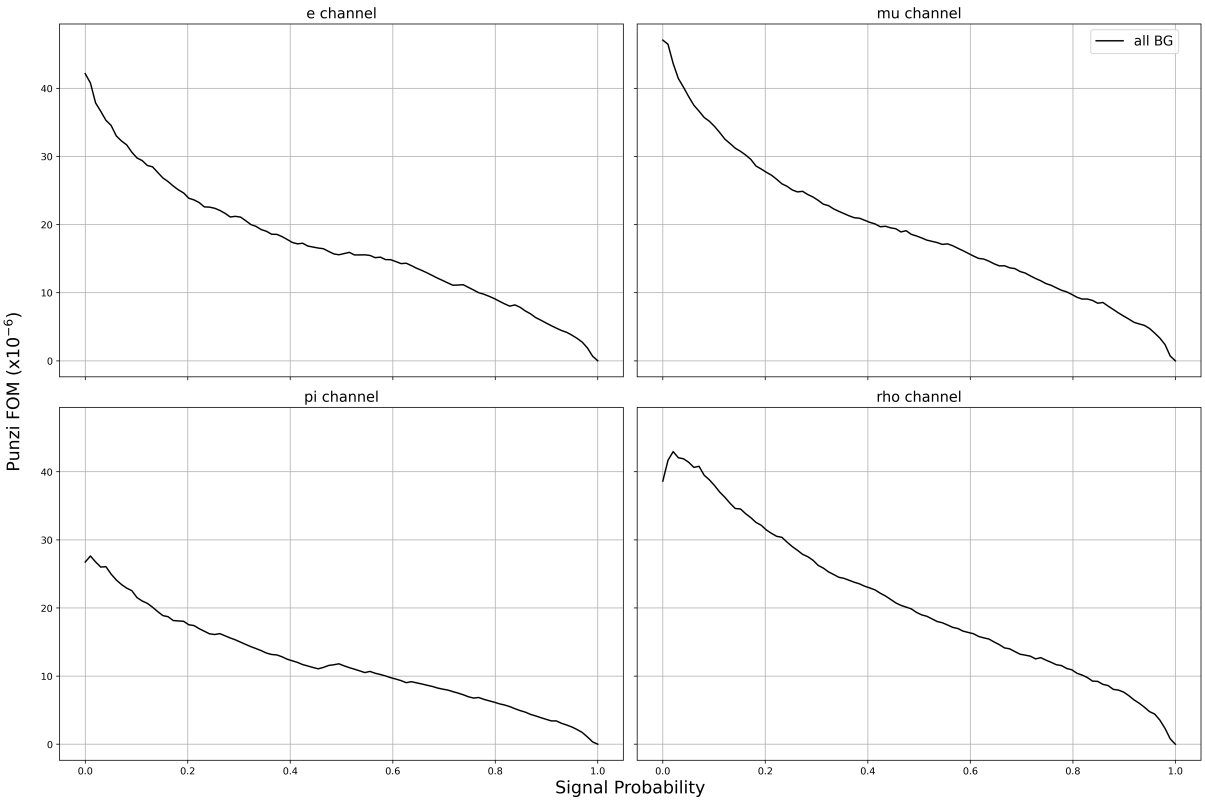

Figure B.4: Plot of the Punzi FOM (defined in Equation 4.4) vs cut value (retaining events above the value) for the Signal Probability, with only events from within the tau mass signal region ($\pm 60 MeV/c^2$ the nominal tau mass). It was optimal to not make a cut on this variable. The generic MC is scaled to 200 fb^{-1}.

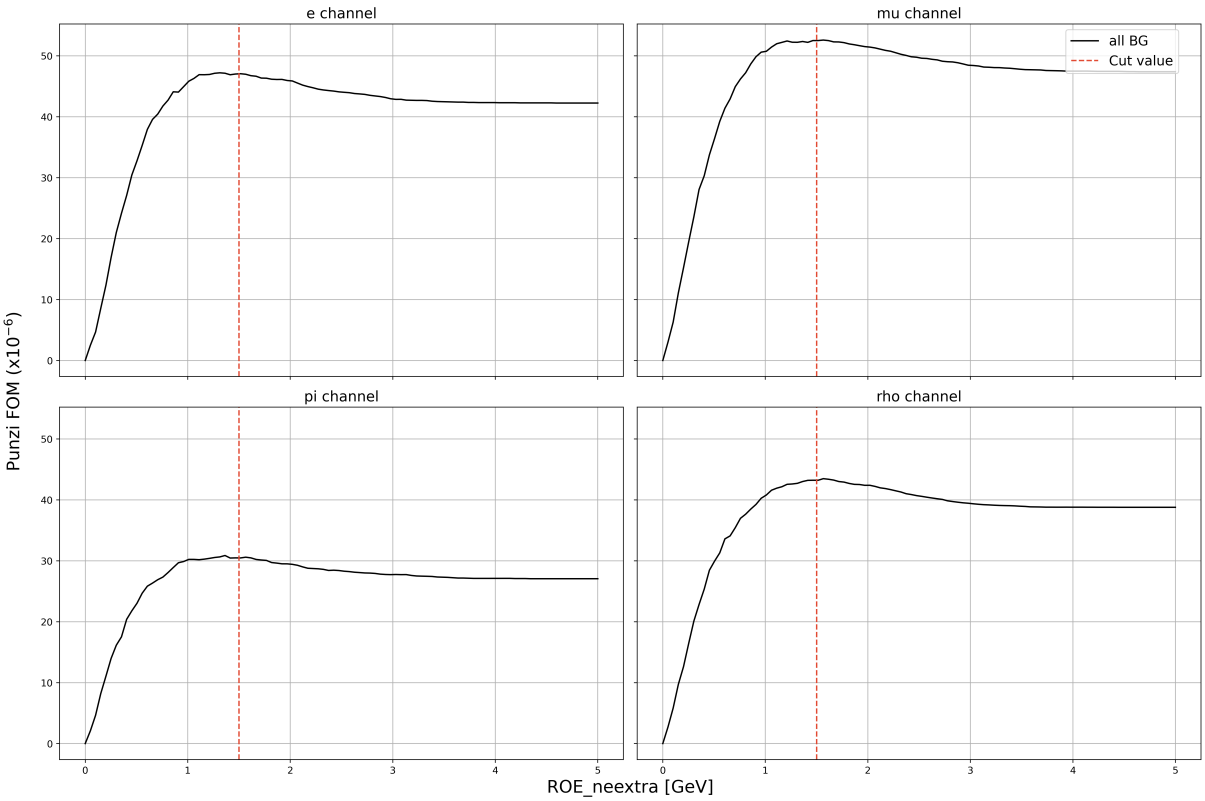

Figure B.5: Plot of the Punzi FOM (defined in Equation 4.4) vs cut value (retaining events below the value) for ROE extra energy, with only events from within the tau mass signal region ($\pm 60 MeV/c^2$ the nominal tau mass). The dashed line represents the optimized cut value for this variable. The generic MC is scaled to 200 fb^{-1}.

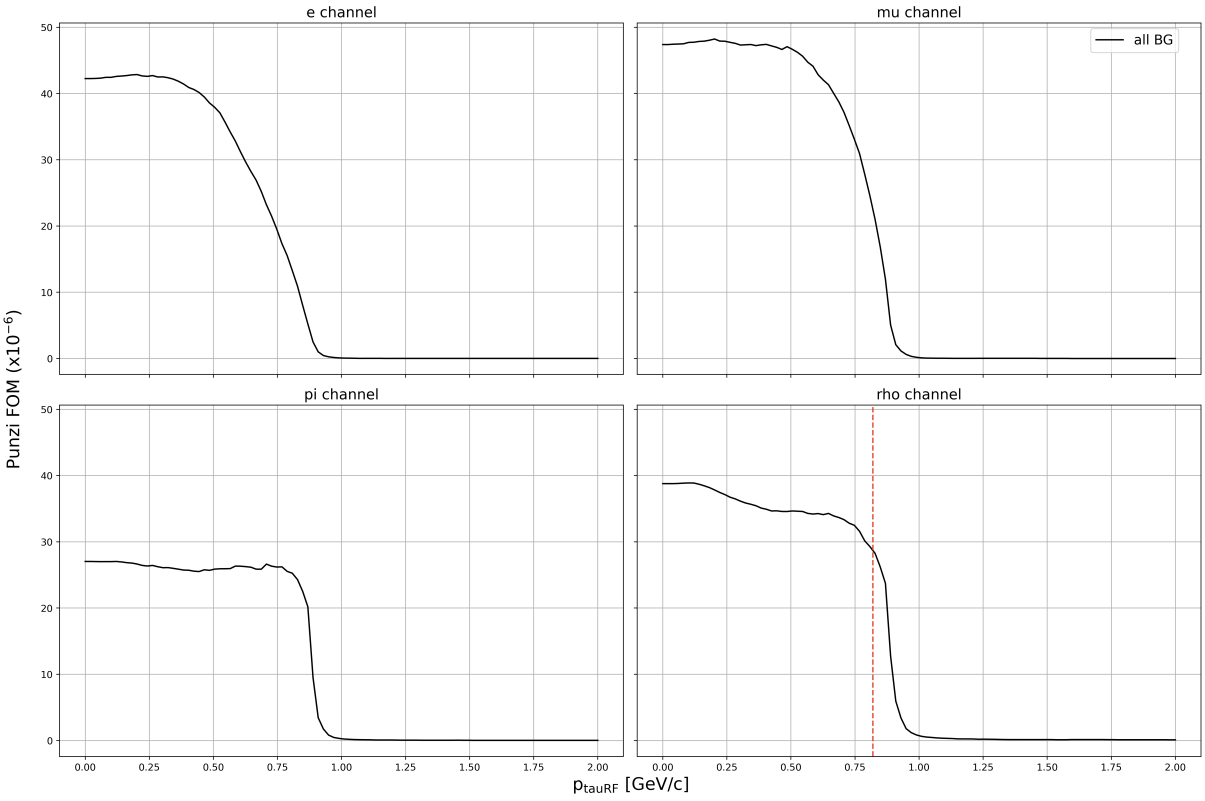

Figure B.6: Plot of the Punzi FOM (defined in Equation 4.4) vs cut value (retaining events above the value) for the tau daughter momentum in tau rest frame, with only events from within the tau mass signal region ($\pm 60 MeV/c^2$ the nominal tau mass). The dashed line represents the optimized cut value for this variable, which only applies to the rho channel. The generic MC is scaled to 200 fb^{-1}.

Appendix C

Data/MC Plots

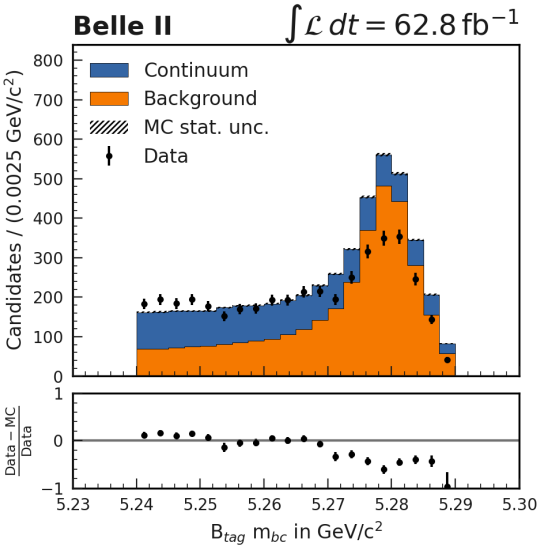

Figure C.1: Data vs generic MC for B_{tag} m_{bc} in the mu channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

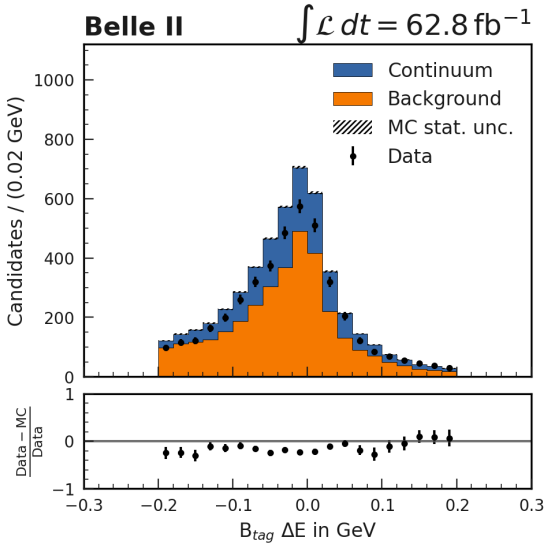

Figure C.2: Data vs generic MC for B_{tag} ΔE in the mu channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

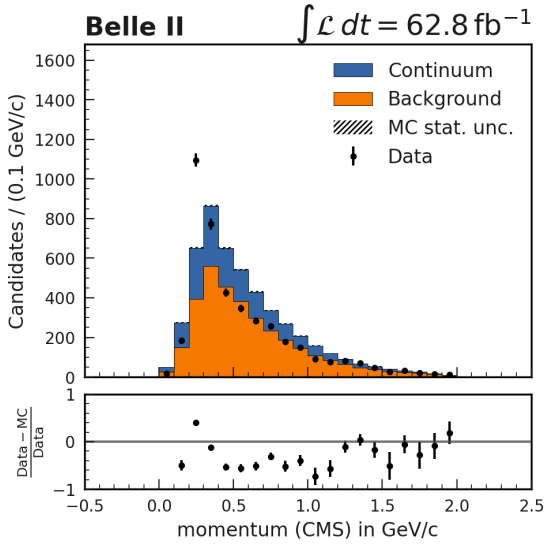

Figure C.3: Data vs generic MC for the primary electron momentum in the CMS frame, in the mu channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

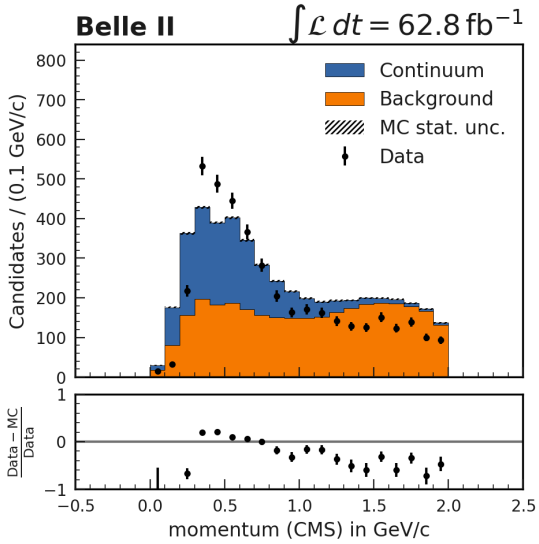

Figure C.4: Data vs generic MC for the charged tau daughter momentum in the CMS frame, in the mu channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

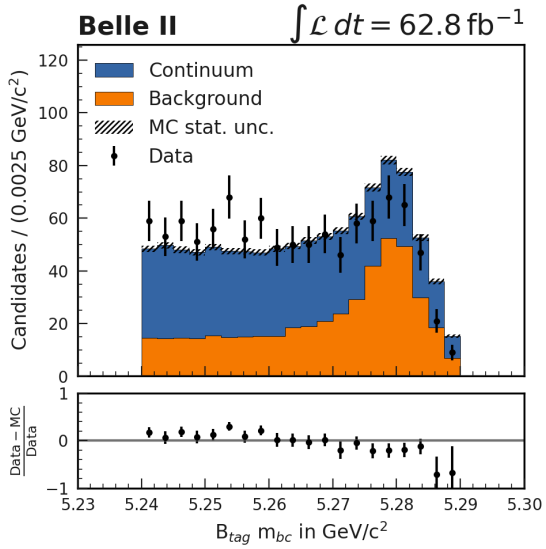

Figure C.5: Data vs generic MC for B_{tag} m_{bc} in the pi channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

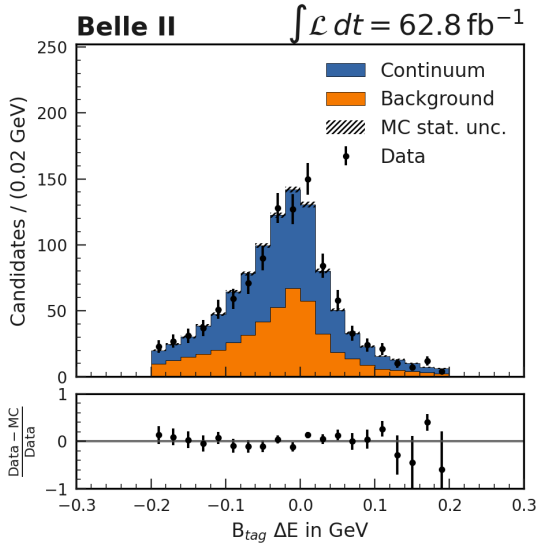

Figure C.6: Data vs generic MC for B_{tag} ΔE in the pi channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

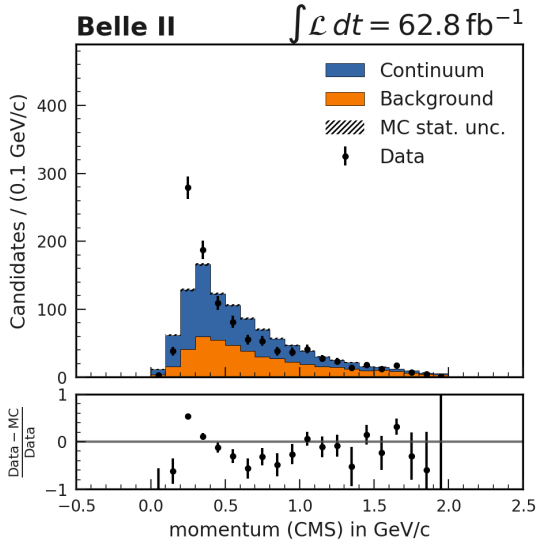

Figure C.7: Data vs generic MC for the primary electron momentum in the CMS frame, in the pi channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

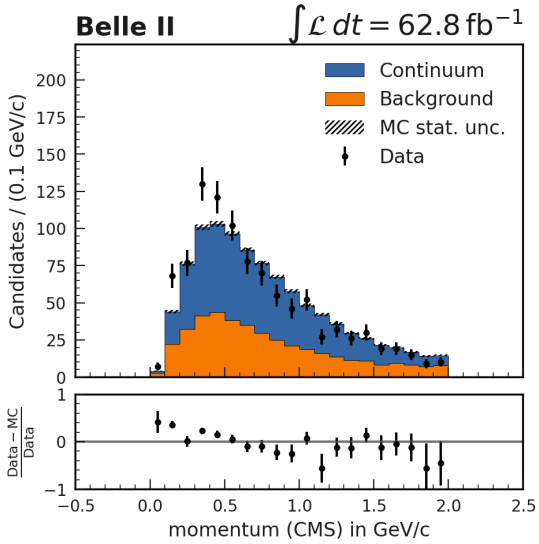

Figure C.8: Data vs generic MC for the charged tau daughter momentum in the CMS frame, in the pi channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

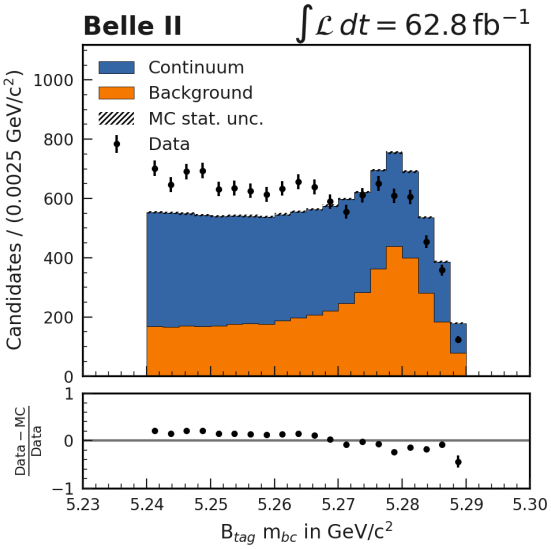

Figure C.9: Data vs generic MC for B_{tag} m_{bc} in the rho channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

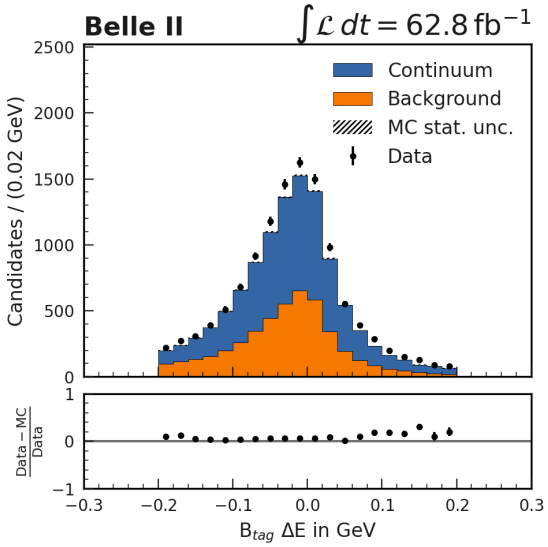

Figure C.10: Data vs generic MC for B_{tag} ΔE in the rho channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

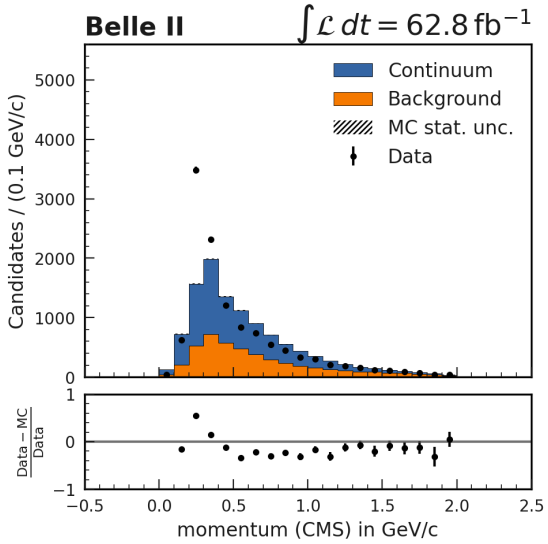

Figure C.11: Data vs generic MC for the primary electron momentum in the CMS frame, in the rho channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.

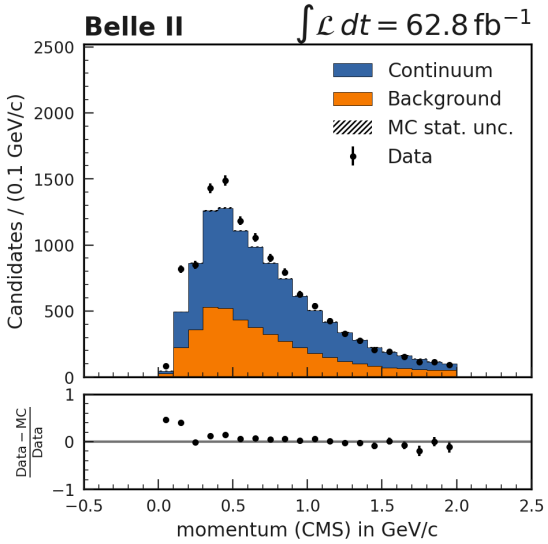

Figure C.12: Data vs generic MC for the charged tau daughter momentum in the CMS frame, in the rho channel. The continuum contains ccbar, ssbar, uubar, and ddbar. The background contains charged and mixed samples. The MC has been scaled by 0.0314.