

Università degli Studi di Napoli “Federico II”

Scuola Politecnica e delle Scienze di Base  
Area Didattica di Scienze Matematiche Fisiche e Naturali

**Dipartimento di Fisica “Ettore Pancini”**



*Laurea Magistrale in Fisica*

## **Search for $B \rightarrow \tau\nu$ decay at the Belle II experiment**

**Relatori:**

Prof. Guglielmo De Nardo  
Dott. Mario Merola

**Candidato:**

Michele Aversano  
Matricola N94000528

A.A. 2019/2020

“天の光は全て星”



# Contents

<b>Introduction</b>	<b>9</b>
<b>1 Flavour Physics in the Standard Model</b>	<b>11</b>
1.1 Introduction . . . . .	11
1.2 Gauge-invariant Yukawa Couplings and CKM Matrix . . . . .	12
1.3 Standard Parametrization of the CKM Matrix . . . . .	16
1.4 The Unitarity Triangle . . . . .	16
<b>2 The Belle II Experiment</b>	<b>19</b>
2.1 B-Factory and Physics Motivations . . . . .	19
2.2 SuperKEKB Collider . . . . .	21
2.2.1 Nano-Beams Scheme . . . . .	23
2.2.2 Beam-Induced Background . . . . .	24
2.3 Belle II Detector . . . . .	25
2.3.1 Vertex Detector(VXD) . . . . .	26
2.3.2 Central Drift Chamber (CDC) . . . . .	27
2.3.3 Particle Identification System (TOP and ARICH) . . . . .	28
2.3.4 Electromagnetic Calorimeter (ECL) . . . . .	29
2.3.5 Superconducting Coil . . . . .	30
2.3.6 $K_L$ - Muon Detector (KLM) . . . . .	30
2.3.7 Trigger System . . . . .	31
<b>3 B Meson Physics: Leptonic Decays</b>	<b>33</b>
3.1 Theory of B Meson Leptonic Decays . . . . .	33
3.2 $B \rightarrow \tau\nu$ Decay in NP Models . . . . .	34
3.3 $\tau$ Decays . . . . .	35
<b>4 Study of <math>B \rightarrow \tau\nu</math> Decay</b>	<b>37</b>
4.1 Computing Environment . . . . .	37
4.2 Dataset . . . . .	38
4.3 Full Event Interpretation and Tagging . . . . .	39
4.4 Signal Events Selection . . . . .	42
4.4.1 Signal Selection Optimization . . . . .	49
4.5 Signal Cross Feeds . . . . .	60
4.6 $B \rightarrow \tau\nu$ Branching Ratio Measurement . . . . .	61
4.6.1 Single Fits . . . . .	61
4.6.2 Simultaneous Fits . . . . .	67

<b>5</b>	<b>Look at Data Collected in 2019</b>	<b>69</b>
5.1	Leptons . . . . .	70
5.2	Hadrons . . . . .	72
5.3	Data-MC Discrepancy Considerations . . . . .	74
	<b>Conclusions</b>	<b>75</b>
	<b>Bibliography</b>	<b>77</b>

# List of Figures

1.1	Unitary Triangle . . . . .	17
2.1	Integrated luminosity estimation of the SuperKEKB accelerator in the next years. . . . .	20
2.2	SuperKEKB collider structure. . . . .	22
2.3	Geometry of the interaction point in the nano-beams scheme. . . . .	24
2.4	Belle II detector top view. . . . .	25
2.5	Belle II detector inner structure. . . . .	26
2.6	View of the Silicon Pixel Detector. . . . .	27
2.7	View of the Silicon Vertex Detector. . . . .	27
2.8	View of the Central Drift Chamber. . . . .	28
2.9	Belle-II PID systems: one of the modules of the TOP counter (left), principle of operation of the proximity focusing RICH with a non-homogeneous aerogel radiator in the focusing configuration (right). . . . .	29
3.1	Feynman diagram of a leptonic B decay. . . . .	34
3.2	Feynman diagram of $B \rightarrow \tau\nu$ decay through a charged Higgs . . . . .	34
3.3	Feynman diagram of a $\tau$ leptonic decay. . . . .	35
3.4	Feynman diagram of $\tau \rightarrow \pi\nu_\tau$ decay . . . . .	36
3.5	Feynman diagram of $\tau \rightarrow 3\pi\nu_\tau$ decay through $a_1$ and $\rho$ resonances. . . . .	36
4.1	Decay of $Y(4S)$ into a charged B meson pair. The signal-side is shown on the right and the tag-side on the left. . . . .	39
4.2	Hierarchical approach of the Full Event Interpretation (FEI). . . . .	40
4.3	Plot of $M_{bc}$ for $e^\pm$ channel with <i>loose</i> cuts; for $\mu^\pm$ , single $\pi^\pm$ and $\rho$ channels the shapes are similar. . . . .	41
4.4	Plot of <i>sigProb</i> with <i>loose</i> cuts; the shape is the same for all PID signal hypotheses since it depends only on the tag side. . . . .	43
4.5	Plot of <i>sigProb</i> with <i>loose</i> cuts; signal and background are normalized to the same area; the shape is the same for all PID signal hypotheses since it depends only on the tag side. . . . .	44
4.6	Plot of $E_{extra}$ for $e^\pm$ channel with <i>loose</i> cuts; for $\mu^\pm$ , single $\pi^\pm$ and $\rho$ channels the shapes are similar. . . . .	44
4.7	Plot of $E_{extra}$ for $e^\pm$ channel between 0 and 1 GeV with <i>loose</i> cuts; for $\mu^\pm$ , single $\pi^\pm$ and $\rho$ channels the shapes are similar. . . . .	45
4.8	Plot of $M_{miss}^2$ for $e^\pm$ channel with cuts on $M_{bc}$ , <i>sigProb</i> and $E_{extra}$ ; for $\mu^\pm$ channel the shapes is similar. . . . .	45
4.9	Plot of $M_{miss}^2$ for single $\pi^\pm$ channel with cuts on $M_{bc}$ , <i>sigProb</i> and $E_{extra}$ ; for $\rho$ channel the shapes is similar. . . . .	46

4.10	Plot of $p$ for $e^\pm$ channel with cuts on $M_{bc}$ , $sigProb$ and $E_{extra}$ ; only for $\mu^\pm$ channel the shapes is similar. . . . .	46
4.11	Plot of $p$ for single $\pi^\pm$ channel with cuts on $M_{bc}$ , $sigProb$ and $E_{extra}$ . . .	47
4.12	Plot of $p$ for $\pi^\pm$ from $\rho$ channel with cuts on $M_{bc}$ , $sigProb$ and $E_{extra}$ . . .	47
4.13	Plot of $P_{miss}$ for $\rho$ channel with cuts on $M_{bc}$ , $sigProb$ and $E_{extra}$ . . . . .	48
4.14	Plot of $p^*$ for $\pi^\pm$ from $\rho$ channel with cuts on $M_{bc}$ , $sigProb$ and $E_{extra}$ . . .	48
4.15	Plot of $p_0^*$ for $\pi^0$ from $\rho$ channel with cuts on $M_{bc}$ , $sigProb$ and $E_{extra}$ . . .	49
4.16	Plot of the FOM as function of cuts on $electronID$ for the $e^\pm$ channel. . .	50
4.17	Plot of the FOM as function of cuts on $sigProb$ for the $e^\pm$ channel. . . . .	50
4.18	Plot of the FOM as function of cuts on $M_{miss}^2$ for the $e^\pm$ channel. . . . .	51
4.19	Stacked plot of $E_{extra}$ applying the best cuts for the $e^\pm$ channel. . . . .	51
4.20	Plot of the FOM as function of cuts on $muonID$ for the $\mu^\pm$ channel. . . . .	52
4.21	Plot of the FOM as function of cuts on $sigProb$ for the $\mu^\pm$ channel. . . . .	52
4.22	Plot of the FOM as function of cuts on $M_{miss}^2$ for the $\mu^\pm$ channel. . . . .	53
4.23	Stacked plot of $E_{extra}$ applying the best cuts for the $\mu^\pm$ channel. . . . .	53
4.24	Plot of the FOM as function of cuts on $pionID$ for the $\pi^\pm$ channel. . . . .	54
4.25	Plot of the FOM as function of cuts on $sigProb$ for the $\pi^\pm$ channel. . . . .	54
4.26	Plot of the FOM as function of cuts on $M_{miss}^2$ for the $\pi^\pm$ channel. . . . .	55
4.27	Plot of the FOM as function of cuts on $p$ for the $\pi^\pm$ channel. . . . .	55
4.28	Stacked plot of $E_{extra}$ applying the best cuts for the $\pi^\pm$ channel. . . . .	55
4.29	Plot of the FOM as function of cuts on $pionID$ for the $\pi^\pm$ from $\rho$ channel. . .	56
4.30	Plot of the FOM as function of cuts on $sigProb$ for $\rho$ channel. . . . .	56
4.31	Plot of the FOM as function of cuts on $P_{mis}$ for the $\rho$ channel. . . . .	57
4.32	Plot of $\mathcal{LR}_{(p^*, p_0^*)}$ for single $\pi^\pm$ from $\rho$ channel with the best cuts for $electronID$ , $muonID$ , $pionID$ , $sigProb$ and $P_{miss}$ with signal and background normalised to the same area. $\mathcal{LR}_{(p^*, p_0^*)}$ is evaluated on signal events (red) and on background ones (blue). . . . .	57
4.33	2D plots of $p^*$ vs $p_0^*$ for the $\rho$ channel with the best cuts for $electronID$ , $muonID$ , $pionID$ , $sigProb$ and $P_{miss}$ for signal (a) and for background (b). . . . .	58
4.34	Plot of the FOM as function of cuts on $\mathcal{LR}_{(p^*, p_0^*)}$ for the $\rho$ channel. . . . .	59
4.35	Stacked plot of $E_{extra}$ applying the best cuts for the $\rho$ channel. . . . .	59
4.36	1 of the 100000 extML fits for the $e^\pm$ channel. . . . .	62
4.37	BR results of the Toy-MC study for the $e^\pm$ channel. . . . .	63
4.38	BR error results of the Toy-MC study for the $e^\pm$ channel. . . . .	63
4.39	1 of the 100000 extML fits for the $\mu^\pm$ channel. . . . .	63
4.40	BR results of the Toy-MC study for the $\mu^\pm$ channel. . . . .	64
4.41	BR error results of the Toy-MC study for the $\mu^\pm$ channel. . . . .	64
4.42	1 of the 100000 extML fits for the $\pi^\pm$ channel. . . . .	64
4.43	BR results of the Toy-MC study for the $\pi^\pm$ channel. . . . .	65
4.44	BR error results of the Toy-MC study for the $\pi^\pm$ channel. . . . .	65
4.45	1 of the 100000 extML fits for the $\rho$ channel. . . . .	65
4.46	BR results of the Toy-MC study for the $\rho$ channel. . . . .	66
4.47	BR error results of the Toy-MC study for the $\rho$ channel. . . . .	66
4.48	BR results of the Toy-MC study for the simultaneous fit. . . . .	67
4.49	BR error results of the Toy-MC study for the simultaneous fit. . . . .	67
5.1	Plot of $ \cos \theta^{TAG, ROE} $ with <i>loose</i> cuts. . . . .	69

5.2	Plots with progressive cuts (each variable is cut only on the previous ones) of $sigProb$ (a), $ \cos\theta^{TAG,ROE} $ (b), $M_{bc}$ (c) and $M_{miss}^2$ (d) for the leptonic channels. . . . .	70
5.3	Plots with the loose cuts of $E_{extra}$ (a), $E_{miss}$ (b), $P_{miss}$ (c), $\theta_{P_{miss}}$ (d), $\cos\theta_{P_{miss}}$ (e) and $p$ (f) for the leptonic channels. . . . .	71
5.4	Plots with progressive cuts (each variable is cut only on the previous ones) of $sigProb$ (a), $ \cos\theta^{TAG,ROE} $ (b) and $M_{bc}$ (c) for the hadronic channels. . . . .	72
5.5	Plots with the loose cuts of $E_{extra}$ (a), $E_{miss}$ (b), $M_{miss}^2$ (c), $P_{miss}$ (d), $\theta_{P_{miss}}$ (e), $\cos\theta_{P_{miss}}$ (f), $p$ (g), $p^*$ (h) and $p_0^*$ (i) for the hadronic channels. . . . .	74

# List of Tables

2.1	SuperKEKB: design parameters of the low energy (LER) and high energy (HER) accelerator rings [15]. . . . .	21
3.1	Branching fractions of the main $\tau$ -decay modes [19]. . . . .	36
4.1	Categories of simulated events, integrated luminosity and split in signal and background categories. In the $e^+e^- \rightarrow Y(4S) \rightarrow B^+B^-$ background simulation the signal contribution is removed. . . . .	38
4.2	Best cuts of each channel for signal events selection. . . . .	59
4.3	Summary of correctly-reconstructed efficiencies and cross feeds. . . . .	60
4.4	Summary of efficiencies $\epsilon_k$ and fit results for each $\tau$ decay channel considered. . . . .	62

# Introduction

Belle II is a Flavour Physics experiment that aims to search for New Physics beyond the Standard Model (SM) of elementary particles and improve the precision of some fundamental parameters of this model. The experiment is the successor of the Belle experiment and its general-purpose detector is installed at the  $e^+ - e^-$  SuperKEKB collider, at the KEK laboratories in Tsukuba, Japan. The accelerator collides electrons and positrons at an energy of the center-of-mass of  $10.58 \text{ GeV}$ , in correspondence with the  $Y(4S)$  resonance, which decays mainly in B mesons pairs. The energy of the two beams is different in order to have a boost of the center-of-mass and to be able to measure the time-dependent CP asymmetry.

This thesis project is divided into two parts. In the first part, the sensitivity of the Belle II experiment to the search for  $B \rightarrow \tau\nu$  decays and its Branching Ratio (BR) estimation are described, analyzing Monte Carlo (MC) simulated samples with an equivalent integrated luminosity of  $1.0 \text{ ab}^{-1}$  of data. The  $B \rightarrow \tau\nu$  signal is searched using the 1-prong decays of the  $\tau$  lepton:  $\tau \rightarrow e\nu_e\nu_\tau$ ,  $\tau \rightarrow \mu\nu_\mu\nu_\tau$ ,  $\tau \rightarrow \pi\nu_\tau$ ,  $\tau \rightarrow \rho\nu_\tau$  with  $\rho \rightarrow \pi^\pm\pi^0$ . In order to optimize the selection and to study the expected sensitivity to the  $B \rightarrow \tau\nu$  signal with  $1 \text{ ab}^{-1}$  of data, an extended maximum likelihood fit for each of the four  $\tau$  decay modes and a simultaneous fit are performed on pseudo-datasets generated using MC PDFs to extract two BR measurements with a ToyMC study.

In the second part, it is studied the agreement of the MC simulation with data collected in 2019 and corresponding to  $L = 8.86 \text{ fb}^{-1}$ . Detecting any discrepancies between predictions of the MC simulation and the experimental data is extremely useful to identify modeling defects in simulations and study systematic effects and their possible corrections.



# Chapter 1

## Flavour Physics in the Standard Model

In this chapter, we provide the theoretical framework relevant to the study of the leptonic decays of B mesons in the Standard Model and also in New Physics models. We will start introducing the  $SU(2)_L$  group for electroweak interactions and the Gauge-invariant Yukawa couplings in order to describe the Cabibbo-Kobayashi-Maskawa matrix, the most used parameterizations, and the unitarity triangle.

### 1.1 Introduction

Flavour Physics, the research area that deals with the properties of quarks and leptons and their interactions, has played a central role in the development of the Standard Model (SM). In nature, all fundamental matter fields – quarks, charged leptons, and neutrinos – come in three copies, the so-called *flavours*. They can be collected in three fermion generations, with increasing masses, but otherwise identical quantum numbers. Born with the discovery of the building blocks of the SM, Flavour Physics' role has become the measurement of its parameters. The majority of the SM parameters are related to the flavour sector and can thus be determined in flavour-violating decays: not least, these measurements can be also used to put constraints on models of New Physics (NP). With increasing experimental and theoretical accuracy, their determination has by now reached impressive precision.

Grouped according to their QCD and QED quantum numbers,  $SU(3) \otimes U(1)_{em}$ , the SM fermions are,

- $3_{2/3}$ : up type quarks  $u, c, t$
- $3_{-1/3}$ : down type quarks  $d, s, b$ ,
- $1_{-1}$ : charged leptons  $e, \mu, \tau$
- $1_0$ : neutrinos  $\nu_e, \nu_\mu, \nu_\tau$

where, in  $X_Y$ ,  $X$  is the color quantum number and  $Y$  is the electric charge.

Each fermion type comes in three copies, i.e., the SM fermions group into three generations.

## 1.2 Gauge-invariant Yukawa Couplings and CKM Matrix

Before the advent of the gauge models in the late 1960s, weak transitions have been described by local two-current interaction originally due to Fermi and generalized to their present form by Feynman and Gell-Mann,

$$\mathcal{H}_{weak} = \frac{G_F}{\sqrt{2}} J'^{\mu} J_{\mu}^{\dagger} \quad (1.1)$$

Here the Lorentz vector  $J'_{\mu}$  is a *charged current*, so called because the charge of the particle entering the interaction vertex differs by one from that of the particle leaving the vertex. With a current that incorporates both leptons and hadrons,

$$J'_{\mu} = L_{\mu} + H_{\mu} \quad (1.2)$$

the interaction Hamiltonian in Eq. 1.1 provides a complete description of the weak processes at low energies, the only energy region where it is regarded as applicable. The coupling constant  $G_F$  is not dimensionless, being given by

$$G_F = 1.16639 \times 10^{-5} GeV^{-2} \quad (1.3)$$

Numerically,  $G_F$  is very small, but having the dimension of the inverse squared mass, it leads to a non-renormalizable interaction. In fact, corrections beyond the tree-diagram level, which are given by loop graphs with internal particle lines, involve higher powers of  $G_F$  and hence higher powers of momentum in the numerator. This leads to increasingly divergent terms in successive orders of the perturbation theory, which cannot be rearranged so as to be absorbed: the theory is not renormalizable.

But if on dimensional grounds we pose

$$4\sqrt{2}G_F = \frac{e^2}{M_W^2} \quad (1.4)$$

the resulting mass  $M_W$  may be viewed as indicating that the weak interaction might just come from the presence of a very massive quantum exchange between interacting particles. Like the electromagnetic current  $j_{\mu}^{em}$ , the weak charged current  $J'_{\mu}$  is a Lorentz four-vector, and we may use the more familiar form of the electromagnetic interaction,  $-ej_{\mu}^{em}A^{\mu}$ , coupling  $J'_{\mu}$  to a new *massive charged* field  $W_{\mu}$  of mass  $M_W$ , in the form

$$\mathcal{L}_{weak} = -\frac{g}{2\sqrt{2}}(J'^{\mu}W_{\mu}^{\dagger} + J'^{\mu\dagger}W_{\mu}) \quad (1.5)$$

Then to the second order, Eq. 1.5 will generate Eq. 1.1 as an effective low-energy weak interaction with coupling constant  $G_F/\sqrt{2} = g^2/8M_W^2$ , where  $M_W^2$  comes from  $W_{\mu}$  field propagator in the limit of *small momentum*. But the theory is not manifestly renormalizable because at *large momentum* we will have divergent integrals in higher-order diagrams. Nevertheless, such theories can be renormalized provided that the gauge invariance holds.

Any Dirac spinor field can be decomposed into left and right-handed components

$$\chi(x) = \chi_L(x) + \chi_R(x) \quad (1.6)$$

where one defines

$$\chi_L(x) = a_L \chi(x) \quad (1.7)$$

$$\chi_R(x) = a_R \chi(x) \quad (1.8)$$

in terms of  $\chi$  by application of the left and right chiral projection operators

$$a_L \equiv \frac{1}{2}(1 - \gamma_5) \quad (1.9)$$

$$a_R \equiv \frac{1}{2}(1 + \gamma_5) \quad (1.10)$$

The expressions of their adjoint conjugates are easy to obtain

$$\bar{\chi}_L = \chi_L^\dagger \gamma_0 = \chi^\dagger a_L \gamma_0 = \bar{\chi} a_R \quad (1.11)$$

$$\bar{\chi}_R = \chi_R^\dagger \gamma_0 = \chi^\dagger a_R \gamma_0 = \bar{\chi} a_L \quad (1.12)$$

If the field considered is massive, the breakup in Eq. 1.6 has no Lorentz-invariant meaning. But, if on the contrary, the mass of the field is zero, either of the two chiral components, which then coincides with a helicity eigenstate, may provide a complete representation of the Lorentz group.

The key experimental fact is that in the spectra of weak decays, such as  $n \rightarrow p + e^- + \bar{\nu}_e$ ,  $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$  and  $\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$ , only left-handed leptons and right-handed antileptons show up so that the decay amplitudes can be described in terms of a charged current that involves only the left chiral component of the fields. This holds for all the leptons families, in general for the other fermions, but now we focus on the electron and its neutrino,

$$\mathcal{L}_\mu^e = 2\bar{e}_L(x)\gamma_\mu\nu_{eL}(x) = \bar{e}(x)\gamma_\mu(1 - \gamma_5)\nu_e(x) \quad (1.13)$$

This expression suggests that  $\nu_{eL}$  and  $e_L$  should be gathered into a two-component vector which can be associated with an  $SU(2)$  group, the simplest group having a complex doublet representation. On the other hand, the right chiral components  $\nu_{eR}$  and  $e_R$ , which do not interact with any other particles, should be left in one-dimensional representations. But while  $e_R$  should certainly be kept because it has the same nonvanishing charge and mass as  $e_L$ , the right chiral component of the neutrino  $\nu_{eR}$  may be immediately dropped because the neutrino is observed left-handed and electrically neutral, and is assumed to be exactly massless.

Therefore, in this model of the electron family, we have as matter fields a doublet,  $\psi_L$ , and a singlet,  $\psi_R$ , of an  $SU(2)$  group,

$$\psi_L = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, \quad \psi_R = e_R \quad (1.14)$$

As this group acts nontrivially just on the left chiral fermions, it is denoted by  $SU(2)_L$  and referred to as the *weak – isospin* group.

The free Lagrangian for the (massless) field in Eq. 1.14 is [1][2]

$$\mathcal{L}_0 = \bar{\psi}_L i \gamma^\lambda \partial_\lambda \psi_L + \bar{\psi}_R i \gamma^\lambda \partial_\lambda \psi_R = \bar{\nu}_L i \gamma^\lambda \partial_\lambda \nu_L + \bar{e} i \gamma^\lambda \partial_\lambda e \quad (1.15)$$

This Lagrangian has two symmetries. The first is the  $U(1)$  hypercharge symmetry  $Y$ :

$$\psi_L \rightarrow \psi'_L = e^{g' \frac{i}{2} Y \alpha} \psi_L, \quad \psi_R \rightarrow \psi'_R = e^{g' \frac{i}{2} Y \alpha} \psi_R \quad (1.16)$$

The second is the isotopic  $SU(2)$  symmetry with the Pauli's matrices  $\tau_i$  being the generators:

$$\psi_L \rightarrow \psi'_L = e^{g \frac{i}{2} \vec{\tau} \vec{\omega}} \psi_L, \quad \psi_R \rightarrow \psi'_R = \psi_R \quad (1.17)$$

The theory is invariant under these global transformations. The extension of the invariance to local transformations is obtained by substituting  $\alpha \rightarrow \alpha(x)$  and  $\vec{\omega} \rightarrow \vec{\omega}(x)$  which allows to build of a gauge theory. The interactions are derived by introducing a covariant derivative for each symmetry of the Lagrangian.

As for the electron, the left-handed quarks are arranged in doublets of the  $SU(2)_L$ :

$$Q_j = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \quad (1.18)$$

while the right-handed quarks are introduced as  $SU(2)_L$  singlets:

$$U_j = u_R, c_R, t_R \quad D_j = d_R, s_R, b_R \quad (1.19)$$

The quarks' couplings to the gluons, weak gauge bosons  $W^+$ ,  $W^-$  and  $Z$ , and the photon is described by the kinetic term in the Lagrangian

$$\mathcal{L}_{fermion} = \sum_{j=1}^3 \bar{Q}_j i \not{D}_Q Q_j + \bar{U}_j i \not{D}_U U_j + \bar{D}_j i \not{D}_D D_j \quad (1.20)$$

where  $\not{D} = \gamma^\mu D_\mu$ . The covariant derivatives are

$$D_{Q,\mu} = \partial_\mu + ig_s T^a G_\mu^a + ig' Q_Q^Y B_\mu + ig \tau^a W_\mu^a \quad (1.21)$$

$$D_{U,\mu} = \partial_\mu + ig_s T^a G_\mu^a + ig' Q_U^Y B_\mu \quad (1.22)$$

$$D_{D,\mu} = \partial_\mu + ig_s T^a G_\mu^a + ig' Q_D^Y B_\mu \quad (1.23)$$

and the hypercharges assigned as  $Q_Q^Y = 1/6$ ,  $Q_U^Y = 2/3$ ,  $Q_D^Y = 1/3$ .  $T^a (a = 1, \dots, 8)$  and  $\tau^a (a = 1, 2, 3)$  are the generators of  $SU(3)_C$  and  $SU(2)_L$ , respectively, and the index  $j$  runs over the three generations of quark fields. It is evident that the gauge couplings are universal for all three generations.

Flavour non-universality, on the other hand, is introduced by the quark Yukawa couplings to the Higgs field, responsible for the generation of non-zero quark masses:

$$\mathcal{L}_{Yuk} = \sum_{i,j=1}^3 (-Y_{U,ij} \bar{Q}_{Li} \tilde{H} U_{Rj} - Y_{D,ij} \bar{Q}_{Li} H D_{Rj} + h.c.) \quad (1.24)$$

where *h.c.* abbreviates the hermitian conjugate term. The subscripts  $i, j$  are generation indices, and the dual field  $\tilde{H}$  is given as  $\tilde{H} = \epsilon H^* = (H^{0*}, -H^-)^T$ . Replacing the Higgs field  $H$  by its vacuum expectation value  $\langle H \rangle = (0, v)^T$ , we obtain the quark mass terms

$$\sum_{i,j=1}^3 (-m_{U,ij} \bar{u}_{Li} u_{Rj} - m_{D,ij} \bar{d}_{Li} d_{Rj} + h.c.) \quad (1.25)$$

with the quark mass matrices given by  $m_{U,D} = v Y_{U,D}$ .

The quark mass matrices  $m_U$  and  $m_D$  are  $3 \times 3$  complex matrices in flavour space with a priori arbitrary entries. They can be diagonalized by making appropriate bi-unitary field redefinitions:

$$u_L = \hat{U}_L u_L^m, \quad u_R = \hat{U}_R u_R^m, \quad d_L = \hat{D}_L d_L^m, \quad d_R = \hat{D}_R d_R^m \quad (1.26)$$

with the superscript  $m$  denoting quarks in their mass eigenstate basis. Unitary transformations of the right-handed quark sector are indeed non-physical, as they drop out from the rest of the Lagrangian. However,  $u_{Li}$  and  $d_{Li}$  form the  $SU(2)_L$  doublets  $Q_i$  (with  $i = 1, 2, 3$ ). Their kinetic term gives rise to the interaction

$$\frac{g}{\sqrt{2}} \bar{u}_{Li} \gamma_\mu W^{\mu+} d_{Lk} \quad (1.27)$$

Transformation of Eq. 1.27 to the mass eigenstate basis yields

$$\frac{g}{\sqrt{2}} \bar{u}_{Li} \hat{U}_{L,ij}^\dagger \hat{D}_{L,jk} \gamma_\mu W^{\mu+} d_{Lk} \quad (1.28)$$

We conclude that the combination

$$\hat{V}_{CKM} = \hat{U}_L^\dagger \hat{D}_L \quad (1.29)$$

is physical and it is called the CKM matrix [3][4]. It describes the misalignment between left-handed up- and down-type quark mass eigenstates, which leads to flavour violating charged current interactions, mediated by the  $W^\pm$  bosons. It is convenient to label the elements of  $\hat{V}_{CKM}$  by the quark flavours involved in the respective charged current interaction:

$$\hat{V}_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad (1.30)$$

For example, the element  $V_{ud}$  appears in the coupling of an up and a down quark to the  $W$  boson.

### 1.3 Standard Parametrization of the CKM Matrix

Let us now determine the number of physical parameters in the CKM matrix. Being a unitary  $3 \times 3$  matrix, it can be parametrized by three mixing angles and six complex phases in general. However, five of these phases are non-physical, as they can be absorbed as unobservable parameters into the up-type and down-type quarks, respectively. Note that an overall phase rotation of all quarks does not affect the CKM matrix. We are then left with three mixing angles  $\theta_{12}$ ,  $\theta_{23}$ ,  $\theta_{13}$  and one complex phase  $\delta$  (the CP-violating phase) as the physical parameters of the CKM matrix. Introducing the short-hand notation  $s_{ij} = \sin \theta_{ij}$  and  $c_{ij} = \cos \theta_{ij}$ , the standard parametrization of the CKM matrix reads [5]

$$\hat{V}_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \quad (1.31)$$

Experimentally, it has been found that the CKM matrix exhibits a rather strong hierarchy, with [6]

$$s_{12} \sim 0.2, \quad s_{23} \sim 0.04, \quad s_{13} \sim 4 \times 10^{-3} \quad (1.32)$$

The CKM matrix hence is close to the unit matrix, with hierarchical off-diagonal elements. Flavour-changing transitions are therefore strongly suppressed in the SM. Similarly, also the quark masses are found to follow a hierarchical pattern, spanning five orders of magnitude in size.

### 1.4 The Unitarity Triangle

The hierarchical structure of the CKM matrix can be used to derive an alternative parametrization, which turns out to be very useful for estimating the size of flavour-violating transitions. In the Wolfenstein parameterization [7]

$$\hat{V}_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \quad (1.33)$$

$\lambda = |V_{us}| \sim 0.2$  is the only small parameter, while  $A$ ,  $\rho$ , and  $\eta$  are  $\mathcal{O}(1)$ . It is therefore convenient to estimate the size of flavour violating decays by making an expansion in powers of  $\lambda$ . The accuracy of this expansion can be improved by changing the parameters of the Wolfenstein parametrization to [8][9]

$$\lambda, \quad A, \quad \bar{\rho} = \left(1 - \frac{\lambda^2}{2}\right)\rho, \quad \bar{\eta} = \left(1 - \frac{\lambda^2}{2}\right)\eta \quad (1.34)$$

As discussed before, the CKM matrix is a unitary matrix, and not all of its elements are independent parameters. Various relations hold among them, which can be tested experimentally. One of the most popular ones,

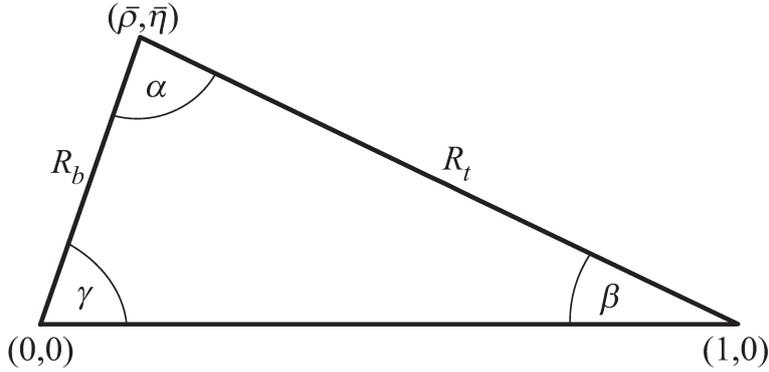


Figure 1.1: Unitary Triangle

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad (1.35)$$

can be displayed as a triangle in the complex plane, the so-called *unitarity triangle (UT)* [10]. With the base of the UT normalized to unity, the apex vertex is simply given by  $(\bar{\rho}, \bar{\eta})$ . The sides  $R_b$  and  $R_t$ , as shown in Fig. 1.1, are given by

$$R_b = \left| \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right| = \left( 1 - \frac{\lambda^2}{2} \right) \frac{1}{\lambda} \frac{|V_{ub}|}{|V_{cb}|} \quad (1.36)$$

$$R_t = \left| \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} \right| = \frac{1}{\lambda} \frac{|V_{tb}|}{|V_{cb}|} \quad (1.37)$$

For the UT angles, two notations are commonly used in the literature. They are related to each other as follows:

$$\alpha \equiv \phi_2, \quad \beta \equiv \phi_1, \quad \gamma \equiv \phi_3 \quad (1.38)$$

The UT can be determined experimentally from various measurements of flavour violating decays of  $K$  and  $B$  mesons. A special role in this determination is played by the length of the side  $R_b$  and the angle  $\gamma$ : being sensitive to the absolute values and CP-violating phases of the elements  $V_{ub}$  and  $V_{cb}$ , they can be determined from  $B$  decays governed by tree level charged current interactions. It is therefore a good approximation to assume that NP contributions to these measurements are negligible. The measurement of  $|V_{ub}|$ ,  $|V_{cb}|$  and  $\gamma = \arg(-V_{ud}V_{ub}^*/V_{cd}V_{cb}^*)$  then leaves us with the *reference unitarity triangle* [11], which determines the CKM matrix independently of potential NP contributions to rare flavour violating decays.

The length of the side  $R_t$  and the angle  $\beta$ , on the other hand, depend on CKM elements involving the top quark. Hence, they can only be measured in loop-induced flavour-changing neutral current (FCNC) processes. Due to their strong suppression in the SM, these observables are sensitive to NP contributions. A model-independent determination of the CKM matrix using these quantities is therefore not possible. NP contributions to the loop-induced processes used in the determination of the UT.

The strategy to hunt for NP contributions to flavour-violating observables is then as follows. First, the CKM matrix and the UT have to be determined from tree-level charged current decays as accurately as possible. As this determination is independent of potential NP contributions, the result can be used as input for precise SM predictions of rare, loop-induced FCNC processes. These predictions are then to be compared with the data, which – in case of a discrepancy – would yield an unambiguous sign of an NP contribution to the decay in question. Clearly, in order to be able to claim an NP discovery in flavour-violating observables, a solid understanding of the SM contribution and its uncertainties is mandatory.

# Chapter 2

## The Belle II Experiment

The Belle II experiment is a Particle Physics experiment meant primarily to search for New Physics (NP) in the flavour sector and to improve the precision of measurements of Standard Model (SM) parameters studying the properties of B mesons (heavy particles containing a bottom quark). Belle II is the upgrade of the Belle experiment, located in the collision point of the beams circulating in the SuperKEKB accelerator, at the KEK laboratories in Tsukuba, in the prefecture of Ibaraki, Japan. The SuperKEKB facility is designed to collide electrons and positrons at center-of-mass energies in the regions of the  $Y$  resonances. Most of the data are collected at the  $Y(4S)$  resonance ( $10.58 \text{ GeV}$ ), which is just above the threshold for B meson pair production. The accelerator is also designed with asymmetric beam energies to provide a boost to the center-of-mass system and thereby allow for time-dependent charge-parity (CP) symmetry violation measurements. The first collision data-taking run was undertaken in 2018, used primarily for accelerator and detector commissioning, and the accelerator will progressively increase the intensity of collisions reaching the design integrated luminosity of  $50 \text{ ab}^{-1}$  between 2028 and 2029, as reported in Fig. 2.1.

The figures and the technical specifications are taken from the Belle II Technical Design Report [12] and from the Belle II Physics Book [13], and updated to the most recent estimations for the luminosity projections.

### 2.1 B-Factory and Physics Motivations

As said before, in the  $e^+ - e^- B - \text{Factory}$  the B mesons are produced in pairs from the decay of the  $Y(4S)$ , a resonance decaying through strong interaction with mass  $m_{Y(4S)} = 10.58 \text{ GeV}$ . The quark composition of  $Y(4S)$  is  $b\bar{b}$  and it's the lightest resonance with a mass sufficient to decay in b-flavoured mesons. The branching ratio in  $B\bar{B}$  pair is about 96%. The two B mesons are produced in an entangled quantum state in the decay so that from the knowledge of the flavour of one B it's possible to assign the (opposite) flavour to the second B. The flavour tagging is performed by reconstructing specific B decays correlated to the flavour of the decaying meson  $B_{tag}$ . If the signal decay  $B \rightarrow f$  is reconstructed from the other tracks of the events, the initial flavour of the  $B_{sig}$  is known. The  $Y(4S)$  is produced boosted with respect to the laboratory frame. This boost and the vertex resolution of B mesons provided by the innermost detectors allow to measure the time-dependent CP violation. This measure consists of measuring the decay times difference of the two B mesons as flavor eigenstates: experimentally this is obtained by measuring the positions of the decay vertices. In its frame, the  $B\bar{B}$  pair

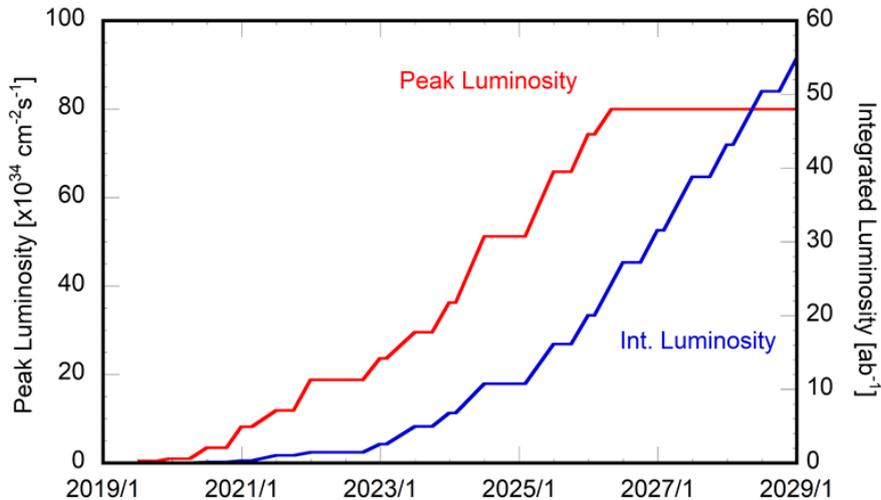


Figure 2.1: Integrated luminosity estimation of the SuperKEKB accelerator in the next years.

is produced almost at rest ( $m_{Y(4S)} - 2m_B = 19 \text{ MeV}$ ), thus in the laboratory frame the flight direction of both B is almost the same of the boost. After the decay and the reconstruction of the first B the decay vertex position  $z_1$  is evaluated. The second B is reconstructed if it decays in the signal  $f$ . From the evaluation of the second vertex position  $z_2$  it's possible to obtain  $\delta z = z_2 - z_1$  and thus  $\Delta t$ . Without the boosted center-of-mass it would not be possible to obtain the time interval in which the two channels can interfere.

The main features of a B-Factory are summarized as follows:

- completely known initial state;
- boosted center of mass, with  $\beta\gamma = 0.28$  at Belle II, necessary to increase the decay length of the  $B\bar{B}$  pair at a measurable level;
- high luminosity, with  $L = 8 \times 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$  at SuperKEKB, in order to produce a large sample of  $B\bar{B}$  pairs;
- high-performance hermetic detector, with good vertex resolution to extrapolate the vertices' position and an excellent particle identification capability of charged and neutral particles.

One of the main questions addressed to the Belle II experiment is the investigation of Beyond Standard Model (BSM) sources of CP violation in the quark sector: the SM CP violation is not sufficient to solve the baryon asymmetry and Belle, BaBar, and LHCb measurements show several tensions with respect to SM expectation. With this purposes, it is interesting to study time-dependent CP violation (TDCPV) in  $b \rightarrow s$  transitions, in which the SM CP violation is expected to be very small, and an observation can be interpreted as a BSM signal. With the same purpose, CP violation in charm mixing can be investigated.

Another interesting sector in Flavour Physics is the search for flavour-changing neutral currents (FCNC) beyond the SM: the FCNC is strongly suppressed by the GIM

mechanism, thus the measurement is very sensitive to NP contributions. Approaches include TDCPV searches in neutral channels or transitions with large missing energy associated with neutrinos escaping the detector, like  $b \rightarrow s\nu\bar{\nu}$ . In this class of decays is crucial the vertex resolution and the neutral reconstruction capability of a B-Factory.

Despite the discovery of the SM Higgs, many extensions of the Higgs sector are not excluded yet, and B decays with  $\tau$  production ( $B \rightarrow \tau\nu$ ,  $B \rightarrow D^*\tau\nu$ ) offer a not trivial way to investigate this sector of possible NP. Currently, this class of decays shows tensions with SM, and the B-Factory environment can exploit its unique features to study the tauonic and semitauonic B decays. More in general, the decays which involve  $\tau$  are challenging because of the large number of neutrinos involved, but accessible with the Belle II experiment. The semileptonic B decays can be useful also to test the lepton universality, and the  $\tau$  decays to study the direct lepton flavour violation

The capability of a new generation of B-Factory to discover NP is not limited to the flavour sector. A B-Factory has a high sensitivity to dark matter via missing energy decays, i.e. via direct detection of new particles. Appropriate specific triggers can be developed for this purpose. In addition, the possibility to tune the center-of-mass of the collider on various strong resonance  $Y(4S)$  allows studying a large family of quarkonia decays, to investigate low energy QCD with a high level of precision.

At last, a B-Factory with increased performance can measure with unprecedented precision the CKM observables increasing the knowledge of the SM.

## 2.2 SuperKEKB Collider

The SuperKEKB is an asymmetric electron-positron collider: it is an upgrade of KEKB that will increase the instantaneous luminosity by about a factor of 40 until it reaches  $8 \times 10^{35} \text{cm}^{-2} \text{s}^{-1}$ . The essential elements in the increase of the luminosity are a reduction in the beam size at the collision point by a factor of 20, from about  $1 \mu\text{m}$  to  $50 \text{nm}$ , and an increase in the currents by a factor of 2 compared to the KEKB values (Tab. 2.1). This is known as a "nano-beam" scheme and was invented by P. Raimondi for the Italian super B-factory [14].

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

Table 2.1: SuperKEKB: design parameters of the low energy (LER) and high energy (HER) accelerator rings [15].

Compared to KEKB, the two beams collide at an even larger angle of  $83 \text{ mrad}$  ( $22 \text{ mrad}$  in KEKB). A somewhat lower beam energy asymmetry of  $7 \text{ GeV}$  (electrons) and  $4 \text{ GeV}$  (positrons), instead of  $8 \text{ GeV}$  and  $3.5 \text{ GeV}$ , was chosen to reduce the beam losses due to Touschek scattering in the lower energy beam. This is expected to

reduce the spatial separation between B-mesons, studied in time-dependent CP violation measurements, but leads to slight improvements in solid angle acceptance for missing energy decays.

The possible energy range of the beams includes the resonant states  $Y(1S)$  (mass  $9.46 \text{ GeV } c^{-2}$ ) and  $Y(6S)$  (mass  $11.24 \text{ GeV } c^{-2}$ ), but most of the data will be taken at the peak of the  $Y(4S)$  resonance. In fact, most of the time, the energy of the center of mass at the collision point will be  $10.58 \text{ GeV}$ , and equal to the mass of the  $Y(4S)$  resonance:

$$E_{CMS} = 2\sqrt{E_{HER(e^-)} \cdot E_{LER(e^+)}} = 10.58 \text{ GeV} = m_{Y(4S)} \quad (2.1)$$

The electrons are produced in a pre-injector by a pulsed laser directed on a cold cathode target, then they are accelerated by a linear accelerator (Linac) to  $7 \text{ GeV}$  and injected in the *High Energy Ring* (HER) of SuperKEKB. The positrons are produced by the collision of electrons with a tungsten target and then they are injected in a damping ring to reduce their emittance. When the positrons reach the required emittance they are accelerated to  $4 \text{ GeV}$  in the Linac and injected in *Low Energy Ring* (LER). The structure of the collider is reported in Fig. 2.2.

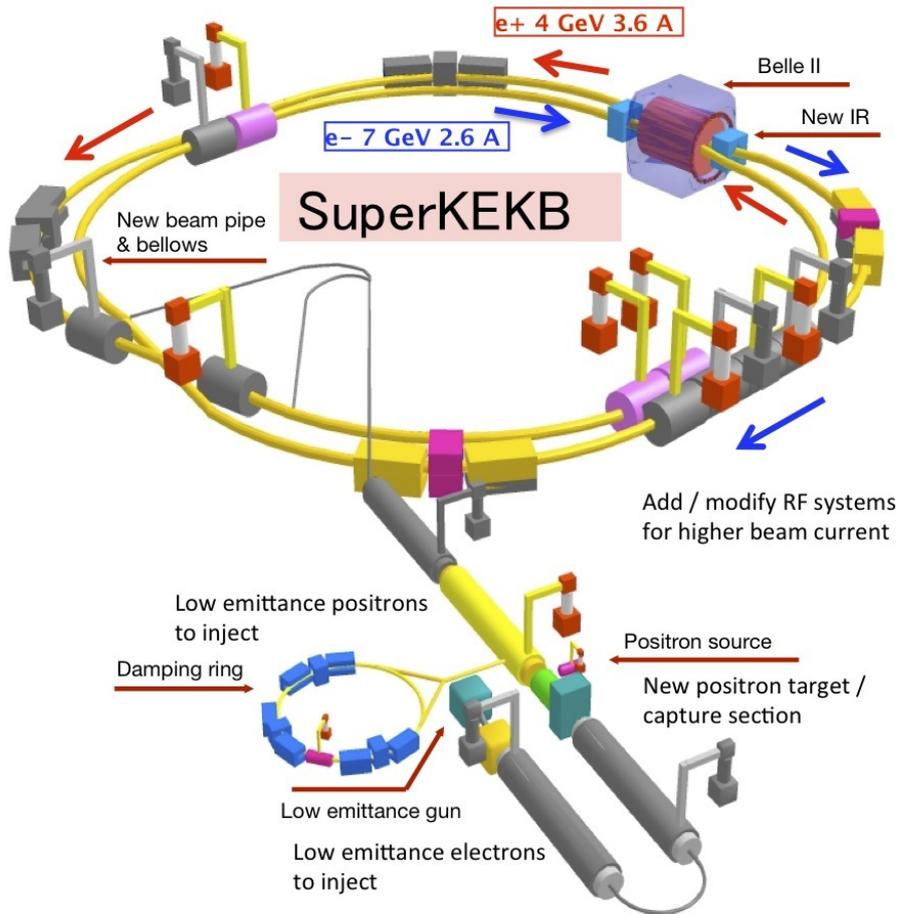


Figure 2.2: SuperKEKB collider structure.

The beam asymmetry produces a Lorentz boost between the center-of-mass frame of the colliding leptons and the detector rest frame (i.e. the laboratory frame) equal to:

$$\beta\gamma = \frac{|\vec{p}_{e^+} + \vec{p}_{e^-}| \cdot c}{\sqrt{s}/c^2} \approx \frac{E_{e^-} - E_{e^+}}{2\sqrt{E_{e^+} \cdot E_{e^-}}} \approx 0.28 \quad (2.2)$$

for a mean flight distance of the B mesons of  $130 \mu m$ . This distance is sufficient to track the vertices of the B mesons, but is reduced with respect to KEKB (it had a  $\beta\gamma = 0.42$ ).

Data taking at SuperKEKB will be performed in two main phases:

- In the first collision data-taking phase (called "phase 2" as "phase 1" denoted the accelerator commissioning phase in 2016 without the final focus and Belle II detector), commencing February 2018 and running until July 2018, SuperKEKB and the interaction region was commissioned before the installation of the sensitive silicon inner detectors. The peak luminosity delivered by SuperKEKB reached  $0.5 \times 10^{34} cm^{-2} s^{-1}$ , and a data set of order  $0.5 fb^{-1}$  ( $472 pb^{-1}$ ) was collected at the  $Y(4S)$  resonance. This small data set may be used for searches of dark sectors that were previously limited by a lack of efficient triggers.
- The second collision phase sees the full detector and allows for the full flavour program to commence, has started in early 2019 and is currently ongoing. The full data-taking program for samples at the different center of mass energies is under development. Most of the data are going to be collected at the  $Y(4S)$  resonance.

## 2.2.1 Nano-Beams Scheme

The nano-beams scheme has been designed to reduce the beam size at IP and so increase the luminosity [14]. The instantaneous luminosity of the collider is given by:

$$L = \frac{\gamma_{\pm}}{2er_e} \left( 1 + \frac{\sigma_y^*}{\sigma_x^*} \right) \frac{I_{\pm} \xi_{y\pm} R_L}{\beta_{y\pm}^* R_{\xi_y}} \quad (2.3)$$

where  $\gamma$  is the relativistic Lorentz factor,  $e$  is the absolute value of the electron charge,  $r_e$  is the classical radius of the electron,  $\sigma_{x,y}^*$  are the widths of the bunch at IP on the transverse plane,  $I$  is the current of the beam,  $\beta_y^*$  is the betatron function at IP (perturbation from the nominal trajectory),  $\xi_y$  is the vertical beam-beam parameter,  $R_L$ , and  $R_{\xi_y}$  are the reduction factors of luminosity and the vertical beam-beam parameter due to not-vanishing crossing angle and the  $\pm$  sign is referred to the charge of the particles in the beam. The idea of the nano-beams is to strongly reduce the  $\beta_y$  function with the minimization of the longitudinal size of the beam overlap at IP (Fig. 2.3).

The dimension of the effective overlap region is  $d \approx \sigma_x^*/\phi$ , where  $2\phi$  is the crossing angle of the beams. For that reason, the crossing angle has been chosen  $2\phi = 83 mrad$  (about four times KEKB crossing angle). In addition  $\sigma_x^*$  is reduced to the size of tens of  $nm$  to reach a beam size at IP of  $50nm$  (from  $\sim 1 \mu m$  of KEKB). In conclusion with this scheme, the  $\beta_y^*$  function is reduced by a factor of 20 with respect to KEKB, and since  $\sigma_y^* \ll \sigma_x^*$ ,  $R_L/R_{\xi_y} \approx 1$ ,  $\xi_y^{SuperKEKB} \approx \xi_y^{KEKB}$ , the current of the beams must be doubled to reach the required luminosity.

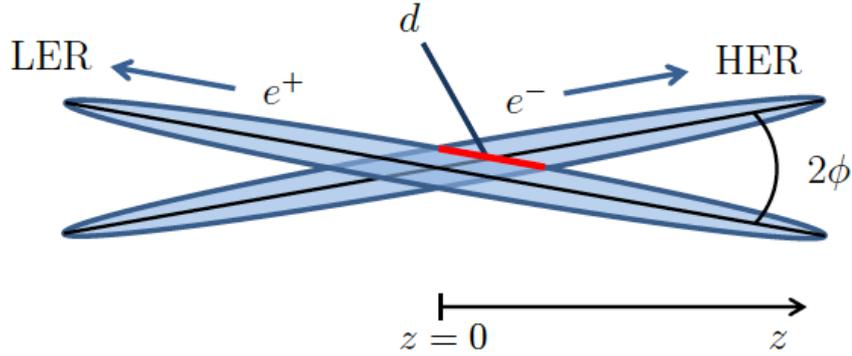


Figure 2.3: Geometry of the interaction point in the nano-beams scheme.

## 2.2.2 Beam-Induced Background

A not negligible background produced by SuperKEKB beams is expected in the Belle II detector. This background is made of particles produced by several physical processes of beam-material or beam-beam interaction. The background yields have been partially measured during the first collision data-taking phase, and precise predictions are also based on simulated SuperKEKB data.

The first background source is the Touschek effect, an intra-bunch Coulomb scattering process that deviates the particle energies from nominal values. The scattered particles propagate around the accelerating ring and finally are lost at the beam-pipe inner wall, producing a shower. That shower might produce signals in the detector. To mitigate this effect, collimators and metal shields are located in the final section of SuperKEKB close to the detector area.

A second background source is the beam-gas scattering, i.e. the interaction between beam particles and residual gas molecules in the beam pipe, with Coulomb scattering or Bremsstrahlung. These interactions might deviate the trajectories and energies from nominal values producing effects similar to Touschek. The countermeasures used for the Touschek background are efficient also for the beam-gas background.

Another source of background is the photons from radiative Bhabha scattering, which interact with the SuperKEKB magnets iron producing neutrons by photo-nuclear resonance mechanism. In addition, a neutron shielding is placed along the ring and close to the detector.

The last background source is the low momentum  $e^+e^-$  pairs produced by two-photon QED process  $e^+e^- \rightarrow e^+e^-e^+e^-$  which might spiral inside the detector.

An additional background source would be the synchrotron radiation of the beam (in particular the HER), which emits photons from a few keV to tens of keV. However, the beam-pipe shape is designed to avoid synchrotron radiation photons pass through the detector. Moreover, the inner surface of the beryllium beam pipe is gold-plated to absorb scattered photons. These precautions should completely suppress the synchrotron radiation background.

## 2.3 Belle II Detector

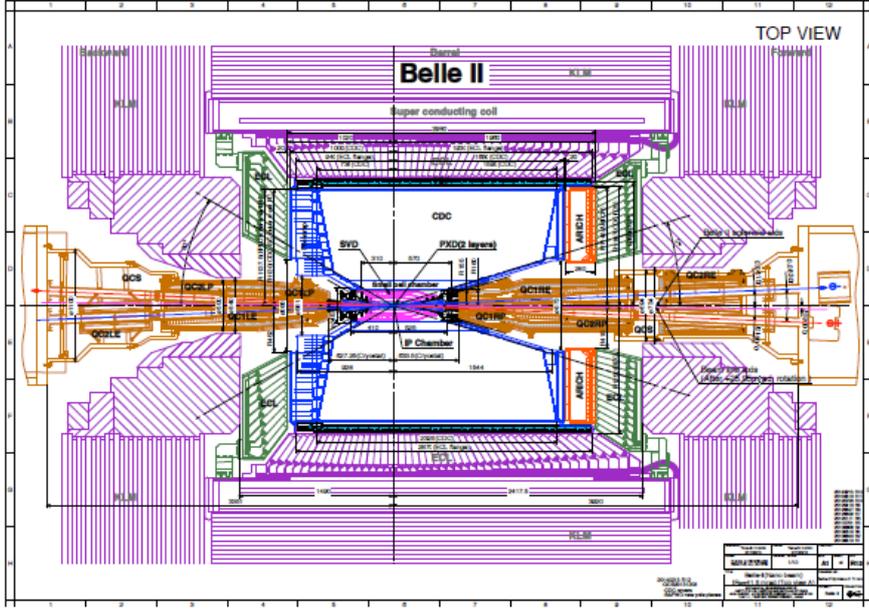


Figure 2.4: Belle II detector top view.

The tool for discoveries at the new generation B-factory will be the Belle II detector (Fig. 2.4). While the new detector clearly fits the same shell as its predecessor, all components are either new or considerably upgraded [12]. Compared to Belle, the Belle II detector will be taking data at an accelerator with a 40 times higher luminosity, and thus has to be able to operate at 40 times higher event rates, as well as with backgrounds rates higher by a factor of 10 to 20 [12].

The following coordinate system is used: the origin is set at IP,  $r \in [0, \infty]$  is the distance from IP on the transverse plane,  $z \in [-\infty, \infty]$  the distance from IP on the longitudinal plane with positive values for the forward region,  $\phi \in [-\pi, \pi]$  is the azimuth angle where the 0 is set in the upward,  $\theta \in [0, 2\pi]$  is the polar angle (0 is set to forward region). In some cases, a Cartesian system is used, with  $z$  set along the electron beam axis in the forward direction,  $y$  upward, and  $x$  in the right direction. The detector has an approximate cylindrical symmetry around the  $z$ -axis, while it has a significant forward-backward asymmetry to improve the solid angle acceptance in the boost direction (the forward direction is the boost direction from IP).

From the innermost to the outermost sub-detector system Belle II is composed by:

- Pixel Detector (PXD): 2 layers of pixel detector (DEPFET technology) which provides 2-dimensional position information;
- Silicon Strip Detector (SVD): 4 layers of Double Sided Silicon Strip Detector which provide 2-dimensional information. It's used to the tracking tasks for online and offline reconstruction. The SVD and the PXD form the Vertex Detector (VXD) of Belle II;
- Central Drift Chamber (CDC): helium-ethane wire drift chamber, composed of 56 layers with stereo and longitudinal geometry, to obtain position information. It is used by the trigger, tracking, and particle ID tasks;

- Particle Identification System (TOP, ARICH): Time-Of-Propagation counter for barrel region with a Cerenkov quartz radiator, and Ring-Imaging Cerenkov Detector, with an aerogel radiator for end-caps regions;
- Electromagnetic Calorimeter (ECL): homogeneous calorimeter composed of CsI(Tl) Crystals that provide  $16.1X_0$ . It's used to measure the energy of photons and energy deposits from charged particles for particle identification;
- Superconducting coil: NbTi/Cu coil that provides a uniform magnetic field of  $1.5 T$  parallel to the beam direction in the internal region. The iron structure of the KLM detector is used to return the yoke of the field;
- $K_L$  and  $\mu$  system (KLM): alternated layers of Resistive Plate Chambers and iron plates in barrel regions and scintillator strips in the end-cap region. It provides a  $3.9$  interaction length in the barrel region and it is used to detect  $K_L$  and  $\mu$  that escape from the internal region.

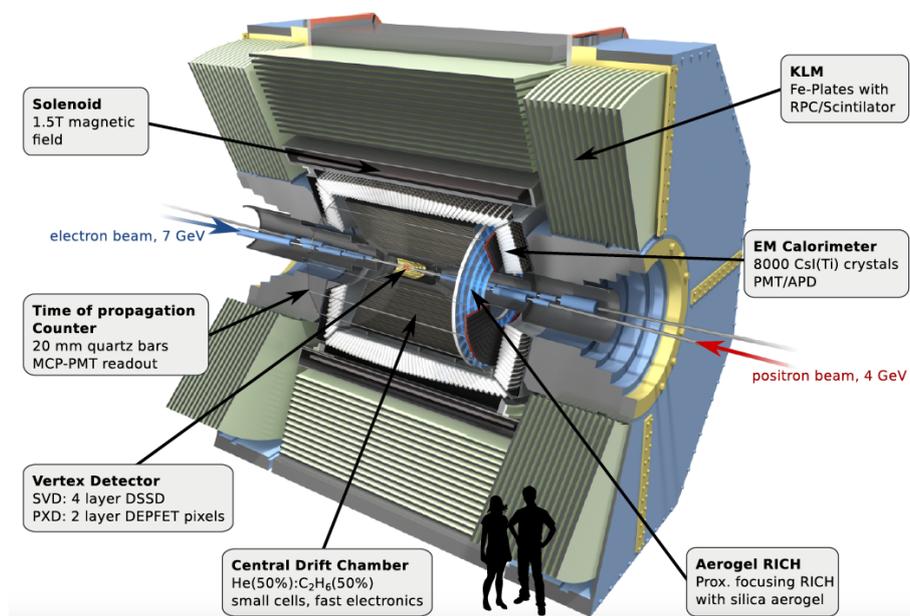


Figure 2.5: Belle II detector inner structure.

### 2.3.1 Vertex Detector(VXD)

The new vertex detector is comprised of two devices, the Silicon Pixel Detector (PXD) and Silicon Vertex Detector (SVD), with altogether six layers around a  $10 \text{ mm}$  radius Be beam pipe. The first layers at  $r = 14 \text{ mm}$  and  $r = 22 \text{ mm}$  compose the PXD (Fig. 2.6) and use pixelated sensors of the DEPFET type. The pixelated sensors have been chosen to sustain the higher hit rate (where a hit is the signal of a detector that returns position information) due to the shorter distance from IP and the higher luminosity with respect to Belle.

The remaining four layers at radii of  $38 \text{ mm}$ ,  $80 \text{ mm}$ ,  $115 \text{ mm}$ , and  $140 \text{ mm}$  compose the SVD and are equipped with double-sided silicon strip sensors. The SVD employs a slanted geometry for the forward sensors to increase the acceptance. In comparison, in

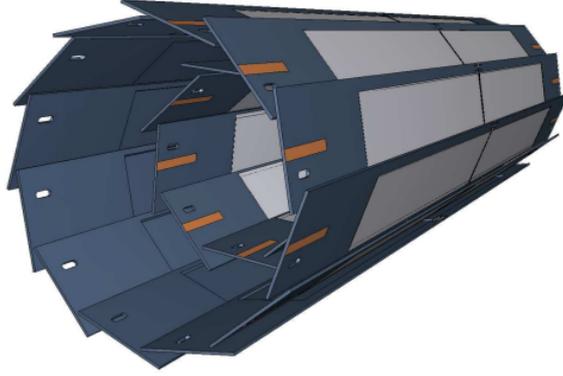


Figure 2.6: View of the Silicon Pixel Detector.

Belle, the outermost vertex detector layer was at a radius of  $88\text{ mm}$ . Compared to the Belle vertex detector, the beam pipe and the first two detector layers are closer to the interaction point, and the outermost layer is at a considerably larger radius. As a result, significant improvement is expected with respect to Belle in the vertex resolution, as well as in the reconstruction efficiency for  $K_0^S \rightarrow \pi^+\pi^-$  decays with hits in the vertex detector.

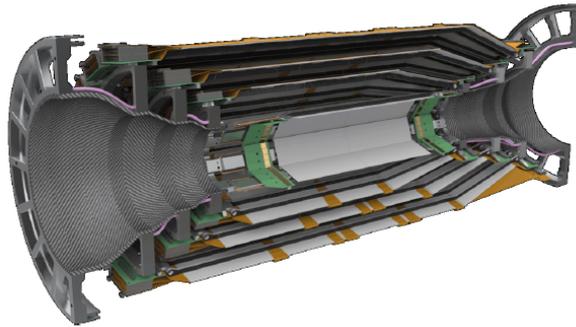


Figure 2.7: View of the Silicon Vertex Detector.

### 2.3.2 Central Drift Chamber (CDC)

In the Belle II detector, the central drift chamber (CDC) plays three important roles (Fig. 2.8). First, it reconstructs charged tracks and measures their momenta precisely. Second, it provides particle identification (PID) information using measurements of energy loss within its gas volume. Low momentum tracks, which do not reach the particle identification device, can be identified using the CDC alone. Finally, it provides efficient and reliable trigger signals for charged particles.

Compared to Belle, it extends to a larger radius ( $1130\text{ mm}$  compared to  $880\text{ mm}$ ) due to the upgrade to a much thinner PID device in the barrel region. To be able to operate at high event rates with increased background levels, the chamber has smaller drift cells than the one used in Belle. In total, the CDC contains 14 336 sense wires arranged in 56 layers, either in *axial* orientation (aligned with the solenoidal magnetic field) or *stereo* (skewed with respect to the axial wires). By combining information from axial and stereo layers it is possible to reconstruct a full 3D helix track. The

chamber gas is comprised of a  $He - C_2H_6$  50 : 50 mixture with an average drift velocity of  $3.3 \text{ cm}/\mu\text{s}$  and a maximum drift time of about  $350 \text{ ns}$  for  $17 \text{ mm}$  cell size.

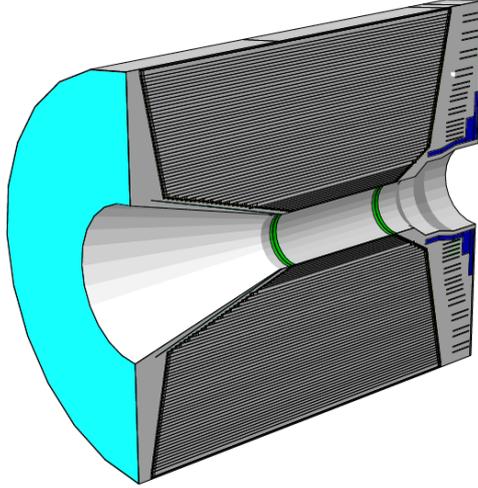


Figure 2.8: View of the Central Drift Chamber.

### 2.3.3 Particle Identification System (TOP and ARICH)

For particle identification in the barrel region, a time-of-propagation (TOP) counter is used. This is a special kind of Cherenkov detector where the two-dimensional information of a Cherenkov ring image is given by the time of arrival and impact position of Cherenkov photons at the photo-detector at one end of a  $2.6 \text{ m}$  long quartz bar (Fig. 2.9). Each detector module (16 in total) consists of a  $45 \text{ cm}$  wide and  $2 \text{ cm}$  thick quartz bar with a small expansion volume (about  $10 \text{ cm}$  long) at the sensor end of the bar. The expansion wedge introduces some additional pinhole imaging, relaxes slightly the precision timing requirements, and reduces the hit occupancy at the photo-detector. At the exit window of the wedge, two rows of sixteen fast multi-anode photon detectors are mounted. The TOP counter requires photo sensors with a single photon time resolution of about  $100 \text{ ps}$ , which can be achieved with a 16-channel MCP PMT specially developed for this purpose. For precision timing required in this type of counter, custom-made waveform sampling read-out electronics is used. Note that for this identification method, the starting (particle production) time has to be known with a precision of about  $50 \text{ ps}$ ; this is indeed challenging but was already achieved for the time-of-flight (TOF) counter of Belle.

In the forward end-cap region, ARICH, a proximity-focusing Cherenkov ring imaging detector with aerogel as Cherenkov radiator will be employed to identify charged particles. The design requirements include a low momentum threshold for pions and good separation of pions and kaons from  $0.4 \text{ GeV}/c$  up to about  $4 \text{ GeV}/c$ . A key parameter of the RICH, the number of detected Cherenkov photons, is increased by a novel method (Fig. 2.9). Two  $2 \text{ cm}$  thick layers of aerogel with different refractive indices ( $n = 1.045$  upstream,  $n = 1.055$  downstream) are used to increase the yield without degrading the Cherenkov angle resolution. As the single photon sensitive high granularity sensor, the hybrid avalanche photon detector (HAPD) is used, developed in a joined effort with Hamamatsu Photonics. In this  $73 \times 73 \text{ mm}^2$  sensor with 144 channels,

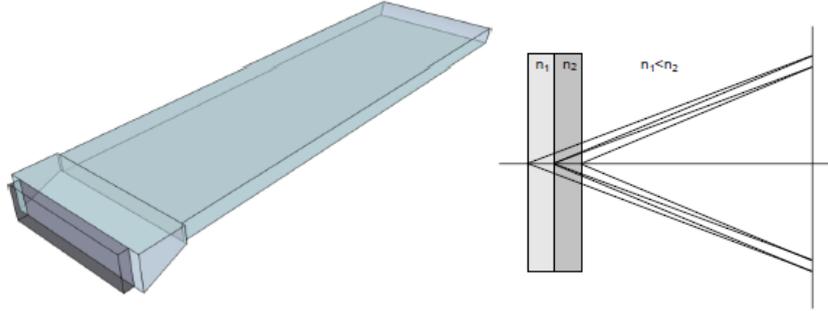


Figure 2.9: Belle-II PID systems: one of the modules of the TOP counter (left), principle of operation of the proximity focusing RICH with a non-homogeneous aerogel radiator in the focusing configuration (right).

photo-electrons are accelerated over a potential difference of  $8\text{ kV}$ , and are detected in avalanche photodiodes (APD). Sensor production was optimized (thicknesses of p and p+ layers, additional intermediate electrode) following radiation tolerance tests with neutrons and gamma rays.

### 2.3.4 Electromagnetic Calorimeter (ECL)

Since one-third of B-decay products are  $\pi^0$ 's or other neutral particles that provide photons in a wide energy range from  $20\text{ MeV}$  to  $4\text{ GeV}$ , a high-resolution electromagnetic calorimeter is a very important part of the Belle II detector. It is a highly-segmented array of thallium-doped cesium iodide CsI(Tl) crystals assembled in a projective geometry (Fig. 2.4). CsI(Tl) was chosen as the scintillation crystal material for the Belle II calorimeter due to its high light output, relatively short radiation length, good mechanical properties, and moderate price. The main tasks of the calorimeter are:

- detection of photons with high efficiency;
- precise determination of the photon energy and angular coordinates;
- electron identification;
- generation of the proper signal for trigger;
- on-line and off-line luminosity measurement;
- $K_L^0$  detection together with the KLM.

All three detector regions, barrel as well as the forward and backward end-caps, are instrumented with a total of 8736 crystals, covering about 90% of the solid angle in the center-of-mass system. The CsI(Tl) crystals, preamplifiers, and support structures have been reused from Belle, whereas the readout electronics and reconstruction software have been upgraded. In the Belle experiment, the energy resolution observed with the same calorimeter was  $\sigma_E/E = 4\%$  at  $100\text{ MeV}$ ,  $1.6\%$  at  $8\text{ GeV}$ , and the angular resolution was  $13\text{ mrad}$  ( $3\text{ mrad}$ ) at low (high) energies;  $\pi^0$  mass resolution was  $4.5\text{ MeV}/c^2$ ; in absence of backgrounds a very similar performance would also be expected in Belle II.

In the presence of considerably elevated background levels as compared to the operation in Belle, the relatively long decay time of scintillations in CsI(Tl) crystals will considerably increase the overlapping of pulses from neighboring (background) events. To mitigate the resulting large pile-up noise, scintillator photo-sensors were equipped with wave-form-sampling read-out electronics. In the forward region of the detector, close to the beam pipe, much higher background rates are expected, such that even with the new wave-form-sampling electronics the pile-up noise will degrade the performance. Some further degradation could come from a reduction of the light yield due to radiation damage, although this effect seems to be less significant than originally anticipated. As a possible solution for this region of the spectrometer, a replacement of CsI(Tl) with considerably faster and radiation-tolerant pure CsI is under study [16].

### 2.3.5 Superconducting Coil

A superconducting coil produces a 1.5 T homogeneous  $\vec{B}$  field parallel to the beam direction. The coil is made of NbTi/Cu, and the internal volume is a cylinder of  $2r = 3.4$  m and a length of 4.4 m. It operates with a 4400 A current and a liquid helium cryogenic system. The iron structure of KLM provides the return yoke of the magnetic field, therefore in the region of KLM outside the coil the direction of  $\vec{B}$  (i.e. the curvature of the tracks) is inverted.

### 2.3.6 $K_L^0$ - Muon Detector (KLM)

The  $K_L^0$  and muon detector (KLM) consists of an alternating sandwich of 4.7 cm thick iron plates and active detector elements located outside the superconducting solenoid. The iron plates serve as the magnetic flux return for the solenoid. They also provide 3.9 interaction lengths or more of material, beyond the 0.8 interaction lengths of the calorimeter, in which  $K_L^0$  mesons can shower hadronically.

The task of the KLM detector is to identify the muon tracks by measuring their penetration depth in the iron and to reconstruct neutral long-lived kaons with the use of the combined information of ECL and the hadronic KLM showers. The barrel detector layers of KLM are Resistive Plate Chambers (RPC): a proportional gas chamber used in streamer mode with a dielectric plate between the electrodes to prevent the propagation of sparks and so increase the spatial resolution. The signal is read with metallic strips on one side of the chamber. Each KLM module is made of two coupled RPC, with an independent power supply and orthogonal strips configuration (this pair of RPC is called superlayer). Both the detector layer and the iron structure of the barrel region are exactly the same ones used in the Belle experiment because the events rate results sustainable despite the increased luminosity. Instead, in the end-caps region and in the innermost barrel super-layer of the KLM, the RPCs have been replaced by two orthogonal layers of scintillator strips coupled with silicon photomultiplier (SiPM), because the RPCs have a too long dead time to sustain the background rate of this region. The muons are identified starting from CDC tracks: each track is extrapolated to KLM region with a  $\pi$  mass hypothesis, and if a KLM hit is present near the extrapolation region it's assigned to that track. The muon detection efficiency plateaus at 89% above 1 GeV/c with a hadron fake rate of about 1.3% mostly due to pions that decay in flight in softer muons. To reconstructs  $K_L^0$  all the KLM hits within a  $5^\circ$  opening angle cone from IP are clustered, then a charged track veto is applied with the use of a CDC track

matching. If the remaining neutral KLM clusters are aligned within a cone of  $15^\circ$  with an ECL cluster the two showers are associated. The  $K_L^0$  detection efficiency rises linearly from 0% at 0  $GeV/c$  to a 80% plateau at 3  $GeV/c$ . The angular resolution is about  $3^\circ$  for KLM-only candidates and  $1.5^\circ$  for KLM+ECL candidates. The SiPMs offer an excellent time resolution of  $\sigma_t \approx 0.7 ns$ , which allows to measure also the time of flight of  $K_L^0$ .

### 2.3.7 Trigger System

The trigger system of Belle II has a nontrivial role to identify events of interest during data taking. The bunch crossing frequency of SuperKEKB is about 250  $MHz$ . Since the bunch crossing time of 4  $ns$  is much faster than the detectors signal collection time for all practical purposes the beam can be considered continuous. Anyway, at full luminosity, the expected event rate is about 50  $kHz$ , and over than 90% of these events are Bhabha scattering or  $2\gamma$  QED processes. A trigger system is therefore required to select events from the beam background and identify interesting ones. Despite  $B\bar{B}$  events are characterized by a higher charged track multiplicity with respect to other events, this variable can not be used in the trigger because  $\tau$  and low multiplicity events would be discarded too. The required trigger must have instead an efficiency close to 100% for  $B\bar{B}$  events and high efficiency for  $\tau$  and low multiplicity events too. Some efficiency degradation is allowed to suppress the Bhabha and  $2\gamma$  QED backgrounds. The trigger rate must be below 30  $kHz$ , the maximum acquisition frequency of DAQ, and the trigger must provide time information with a precision below 10  $ns$  to exploit the potential of the Belle II sub-detectors. To cope with the high background and to the several physics scenarios the trigger system must be robust and flexible. The Belle II trigger is subdivided into two main stages: a hardware trigger or Level 1 trigger (L1) and a software trigger or High-Level Trigger (HLT). The first one removes most of the background events with the use of raw information from the faster sub-detectors with an output trigger rate of 30  $kHz$ , while the second one refines the selection with a more exhaustive analysis and reduces the event rate from L1 trigger to a storable rate of 10  $kHz$ .



# Chapter 3

## B Meson Physics: Leptonic Decays

In this chapter, we consider leptonic B meson decays that proceed in the SM via first-order weak interactions and are mediated by the W boson. B meson decays involving electrons and muons are expected to be dominated by the tree-level W boson decays and any new physics contributions are expected to be highly suppressed with respect to the SM.

The purely leptonic decays  $B \rightarrow l\nu_l$  are of particular interest due to their clear theoretical description. In particular, the  $B \rightarrow \tau\nu_\tau$  decay has an important role in the study of Physics Beyond the Standard Model (BSM).

### 3.1 Theory of B Meson Leptonic Decays

In the SM, the weak decay  $B^+ \rightarrow l^+\nu_l$  of a charged B meson occurs, in the parton model at the lowest perturbative order, through the annihilation of the heavy and light quark inside the meson, and it is therefore mediated by a charged current (Fig. 3.1). The branching ratio is given by

$$BR(B^+ \rightarrow l^+\nu_l) = \frac{G_F^2 m_B m_l^2}{8\pi} \left(1 - \frac{m_l^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B \quad (3.1)$$

where  $m_B$  and  $m_l$  are the B meson and the lepton mass, respectively, and  $\tau_B$  is the B meson lifetime. The decay constant  $f_B$ , which parameterizes the matrix elements of the axial vector current  $p_B^\mu f_B$ , is calculated on the lattice; all the other inputs in Eq. 3.1 are measured experimentally. The branching fraction depends strongly on the mass of the lepton due to helicity suppression, and thus the  $B^+ \rightarrow \tau^+\nu_\tau$  decay is expected to have the largest leptonic branching fraction of the  $B^+$  meson and is the only decay of this kind for which there is experimental evidence. The ratio between the rates for the lepton species  $\tau : \mu : e$  is  $\sim 1 : 10^{-3} : 10^{-7}$ . The predicted values for the SM are found to be [13]:

$$BR_\tau = (7.7 \pm 0.6) \times 10^{-5}, \quad BR_\mu = (3.5 \pm 0.3) \times 10^{-7}, \quad BR_e = (8.1 \pm 0.6) \times 10^{-12} \quad (3.2)$$

Past measurements of  $BR(B^+ \rightarrow \tau^+\nu_\tau)$  by Belle and BaBar were performed with two independent approaches to reconstruct  $B_{tag}$ : using semileptonic and hadronic decays [17] [18]. Combining the measurements by Belle and BaBar, the average is given as

$(1.06 \pm 0.19) \times 10^{-4}$  [17] [18], which has over  $5\sigma$  significance. The last value reported on the PDG review [19] is  $(1.09 \pm 0.24) \times 10^{-4}$ ; both are consistent with the prediction ( $BR_\tau = (7.7 \pm 0.6) \times 10^{-5}$ ) at  $2\sigma$ .

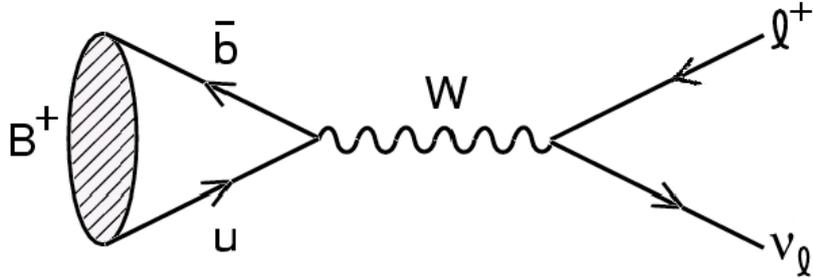


Figure 3.1: Feynman diagram of a leptonic B decay.

In the SM context the observation of  $B^- \rightarrow l^- \bar{\nu}_l$  provides a first direct measurement of  $f_B$  ( $|V_{ub}|$  is measured from semi-leptonic B meson decays). Vice versa if  $f_B$  is calculated precisely from QCD, the branching ratio measurement could infer precise information about  $|V_{ub}|$  value. The ratio between parameters  $|V_{ub}/V_{td}|$  is obtained by comparing  $BR(B^- \rightarrow l^- \bar{\nu}_l)$  with the difference in heavy and light neutral  $B_d$  masses ( $\Delta m_d$ , known from  $B_d$  mixing measurements). Despite the theoretically clean dependence on relevant parameters, the experimental picture is more complicated. The  $\mu$  channel is experimentally simpler but the helicity suppression makes this process quite rare. Regarding the  $\tau$  channel, the necessity to reconstruct the  $\tau$  lepton from its decay products and the presence of two or three undetectable neutrinos in the final state makes the background rejection an experimental challenge.

### 3.2 $B \rightarrow \tau \nu$ Decay in NP Models

The  $B \rightarrow \tau \nu$  decay has an important impact in Physics Beyond the Standard Model (BSM) because it allows to constraint parameters of New Physics (NP). In the two Higgs doublet model (2HDM) [20] the decay involves the contribution of a charged Higgs at tree level, as shown in Fig. 3.2.

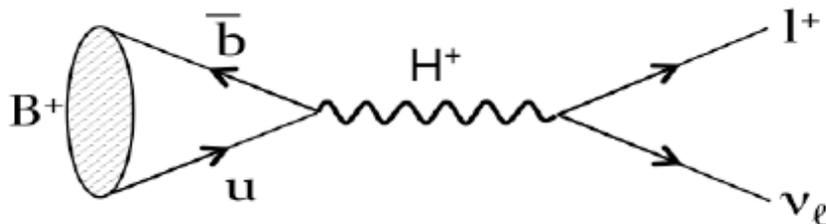


Figure 3.2: Feynman diagram of  $B \rightarrow \tau \nu$  decay through a charged Higgs

Charged Higgs Yukawa couplings are controlled by the parameter  $\tan \beta = v_2/v_1$ , the ratio of vacuum expectation values of the two doublets. The  $W^\pm$  and  $H^\pm$  induce the Fermi interaction

$$\frac{G_F}{\sqrt{2}} V_{ib} [\bar{u}_i \gamma_\mu (1 - \gamma_5) b] [\bar{l} \gamma_\mu (1 - \gamma_5) \nu] - R_l [\bar{u}_i (1 + \gamma_5) b] [\bar{l} (1 - \gamma_5) \nu] \quad (3.3)$$

where

$$R_l = \tan^2 \beta \left( \frac{m_b m_l}{m_{H^-}^2} \right) \quad (3.4)$$

The pseudo-scalar coupling of the  $H^\pm$  boson is given by:

$$- i f_B \left( \frac{m_B^2}{m_b} \right). \quad (3.5)$$

Finally, the branching fraction results to be:

$$BR(B \rightarrow \tau \nu) = BR_{SM} \times \left( 1 - \tan^2 \beta \frac{M_{B^\pm}^2}{m_{H^\pm}^2} \right) \quad (3.6)$$

Comparing the measured value of the branching ratio and the SM prediction, it is possible to exclude regions in the  $(m_{H^\pm}; \tan \beta)$  plane.

### 3.3 $\tau$ Decays

Due to its mass ( $m_\tau \approx 1.77 \text{ GeV}/c^2$ ), the  $\tau$  is the only lepton that decays in hadrons. This is a powerful window to study QCD in the energy region less than  $1 \text{ GeV}$ . The main  $\tau$ -decay modes are:

- leptonic decays, i.e.  $\tau \rightarrow l \bar{\nu}_l \nu_\tau$  (Fig. 3.3):

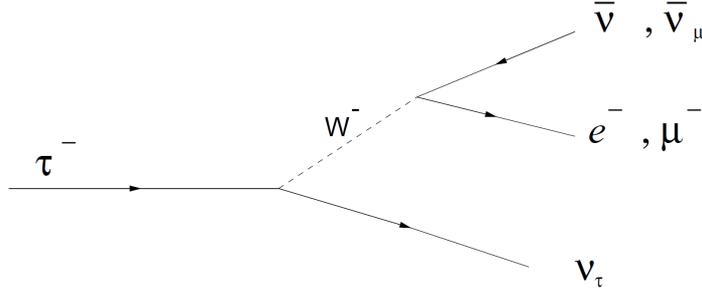


Figure 3.3: Feynman diagram of a  $\tau$  leptonic decay.

- hadronic decays, i.e.  $\tau \rightarrow \pi \nu_\tau$ ,  $\tau \rightarrow \rho \nu_\tau$ ,  $\tau \rightarrow a_1 \nu_\tau$ , where the  $\rho$  and  $a_1$  resonances decays in  $2\pi$  and  $3\pi$  respectively (Fig. 3.4 and 3.5).

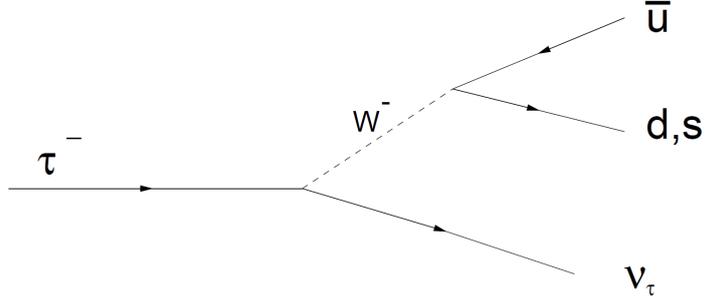


Figure 3.4: Feynman diagram of  $\tau \rightarrow \pi \nu_\tau$  decay

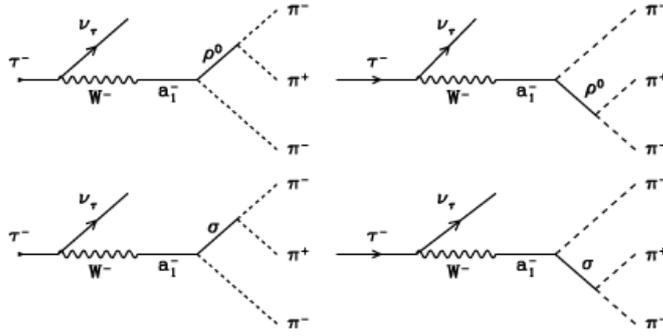


Figure 3.5: Feynman diagram of  $\tau \rightarrow 3\pi \nu_\tau$  decay through  $a_1$  and  $\rho$  resonances.

Decay mode	Branching Fractions
$\mu^- \bar{\nu}_\mu \nu_\tau$	$17.3937 \pm 0.0384$
$e^- \bar{\nu}_e \nu_\tau$	$17.8175 \pm 0.0399$
$\pi^- \nu_\tau$	$10.8164 \pm 0.0512$
$\pi^- \pi^0 \nu_\tau$	$25.4941 \pm 0.0893$
$\pi^- 2\pi^0 \nu_\tau$	$9.2595 \pm 0.0964$
$\pi^- \pi^- \pi^+ \nu_\tau$	$8.9868 \pm 0.0513$

Table 3.1: Branching fractions of the main  $\tau$ -decay modes [19].

# Chapter 4

## Study of $B \rightarrow \tau\nu$ Decay

The purpose of our study is twofold: on one hand to describe the sensitivity of the Belle II experiment to the search for  $B \rightarrow \tau\nu$  decays performed on Monte Carlo simulated samples corresponding to an integrated luminosity of  $1.0 \text{ ab}^{-1}$  and to perform a Branching Ratio estimation; on the other to study the agreement of the MC simulation with data collected in 2019 and corresponding to  $L = 8.86 \text{ fb}^{-1}$ .

In this chapter, a brief technical description of the dataset used is given, followed by the discussion of the strategy implemented in order to reject the most abundant backgrounds. The  $B \rightarrow \tau\nu$  signal is searched using the 1-prong decays of the  $\tau$  lepton:  $\tau \rightarrow e\nu_e\nu_\tau$ ,  $\tau \rightarrow \mu\nu_\mu\nu_\tau$ ,  $\tau \rightarrow \pi\nu_\tau$ ,  $\tau \rightarrow \rho\nu_\tau$  with  $\rho \rightarrow \pi^\pm\pi^0$ . A multivariate technique-based tool, the *Full Event Interpretation*, selects the  $B_{tag}$  in each event and the signal selection and its optimization are described in detail. In order to optimize the selection and to study the expected sensitivity to the  $B \rightarrow \tau\nu$  signal with  $1 \text{ ab}^{-1}$  of data, an extended maximum likelihood fit for each of the four  $\tau$  decay modes and a simultaneous fit are performed on pseudo-datasets generated using MC PDFs and setting the BR to the last reported in the PDG review [19] with a Toy-MC method.

In the next chapter, we will study any discrepancies between predictions of the MC simulation and the experimental data, extremely useful to identify modeling defects in simulations and study systematic effects and their possible corrections.

### 4.1 Computing Environment

The tool used to read and analyze data from the Belle II experiment and from the Monte Carlo simulation is ROOT, a modular scientific software developed by CERN, which provides all the functionality necessary to work with large quantities of data. It mainly uses the C++ language to which are added specific classes of this program, all starting with T, which make it very versatile and suitable for various situations. The main classes used in this thesis work are:

- TTree, to have access to data in .root format.
- TH1D, to create one-dimensional histograms with the data read from the .root files.
- TH2D, with the same function as TH1D, but with two-dimensional histograms.
- THStack, to add multiple histograms and display them on the screen.

- TColor, to use different colors to distinguish the various categories.
- TLegend, to create legends that collected various information.

All data used were manipulated using a specific tool, BASF2, developed by the Belle II collaboration. The raw data coming from the detector are calibrated, reconstructed, and stored on tape using PANTHER based data summary tape (DST) files [21]. PANTHER is a custom serialization format. After each experiment, the calibration constants are recomputed and stored in the Belle Condition Database. Data are reprocessed and stored in a compact form called mDST files, a reduced and compressed form of data summary files in ROOT format. The reconstruction and processing of the mDST files are handled by the Belle II AnalySis Framework (BASF2).

The framework is written in C++ and Phyton plus additional third-party libraries (i.e. EvtGen, GEANT4, ROOT). BASF2 is divided into packages, each of them covering a different aspect of data processing: data acquisition, Monte Carlo event generation, detector and sub-detector simulation, track reconstruction, visualization of individual events inside the detector, and physics analysis. The packages contain libraries, modules, and data-objects. The libraries are implemented in C++ and they provide functionality shared between different modules. The modules are small processing blocks built on top of the libraries and they operate on data event by event performing self-contained tasks. A chain of modules represents a path. The information shared between modules is encoded in data-objects. They are stored in a common DataStore, which every module can read and write. Typical examples for data-objects are *Track*, *ECLCluster*, *Particle* and *ParticleList* objects. To use BASF2 the user has to provide a steering file written in Python.

## 4.2 Dataset

Data were preliminarily organized into 9 files, 8 containing the Monte Carlo Simulation of the most important physics processes and 1 the experimental data of Belle II.

The following categories of events have been simulated:

Category	Int. Lum. ( $fb^{-1}$ )	Sig/Bkg
$e^+e^- \rightarrow Y(4S) \rightarrow B^+B^-(B^\pm \rightarrow \tau^\pm\nu_\tau, B^\mp \rightarrow generic)$	324525	Signal
$e^+e^- \rightarrow Y(4S) \rightarrow B^+B^-$	764	Background
$e^+e^- \rightarrow Y(4S) \rightarrow B^0\bar{B}^0$	763	Background
$e^+e^- \rightarrow u\bar{u}$	1000	Background
$e^+e^- \rightarrow d\bar{d}$	1000	Background
$e^+e^- \rightarrow c\bar{c}$	1000	Background
$e^+e^- \rightarrow s\bar{s}$	1000	Background
$e^+e^- \rightarrow \tau^+\tau^-$	200	Background

Table 4.1: Categories of simulated events, integrated luminosity and split in signal and background categories. In the  $e^+e^- \rightarrow Y(4S) \rightarrow B^+B^-$  background simulation the signal contribution is removed.

To describe the sensitivity of the Belle II experiment to the search for  $B \rightarrow \tau\nu$  decay, each MC sample is preliminary normalized to  $L = 1 ab^{-1}$ . To study the agreement of the

MC simulation with experimental data, each MC sample is normalized to  $L = 8.86 fb^{-1}$ , corresponding to the integrated luminosity of data collected in 2019.

Due to the high level of machine background in Belle II ( $\sim$  a factor 20 more than in Belle), a study has been performed on MC simulated events to optimally select the photon candidates from  $e^+e^-$  collisions (physics photons) and reject beam induced background photon candidates (background photons). Two cluster-related discriminating variables have been exploited, i.e. the cluster energy and the cluster timing. Physics photon candidates are required to satisfy a minimum energy threshold since they have a harder energy spectrum than background photons. Beam-induced photon production is not correlated with bunch crossing, so the cluster time distribution shows a uniform distribution for background photons and a peak near the bunch crossing time for physics photons. These photon candidates are used in  $\pi^0$  reconstruction and for determining the remaining energy deposition in the calorimeter after the  $B_{tag}$  and  $B_{sig}$  candidates reconstruction from physics photons (these extra photons will be important to define a variable introduced in the next sections). The requirements imposed are the following:

- $\gamma$  from  $\pi^0$ :  $E_{cluster} > 50 MeV$  and  $|T_{cluster}| < 2|\Delta T_{cluster}|$
- extra clusters:  $E_{cluster} > 55 MeV$  and  $|T_{cluster}| < |\Delta T_{cluster}|$

where  $E_{cluster}$  is the energy of the cluster,  $T_{cluster}$  is the time measured by the ECL and  $\Delta T_{cluster}$  is the estimated uncertainty of  $T_{cluster}$ .

### 4.3 Full Event Interpretation and Tagging

In order to study signal events, a multivariate tagging method is used (tag-side), so the decay of one of the B mesons in the event is fully reconstructed ( $B_{tag}$ ). Then, on the signal-side, we will study the properties of the remaining particles using MC simulations of the signal and the backgrounds.

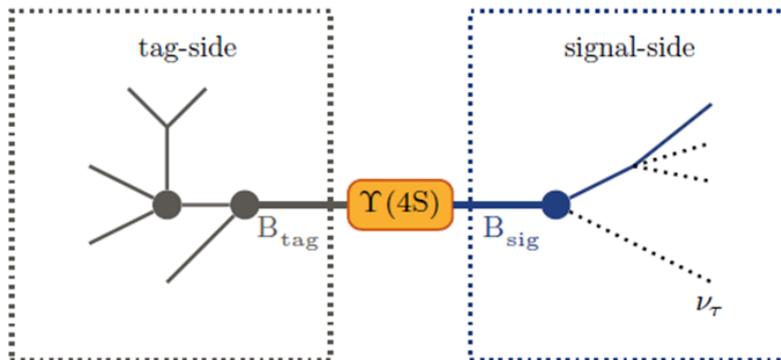


Figure 4.1: Decay of  $Y(4S)$  into a charged B meson pair. The signal-side is shown on the right and the tag-side on the left.

The Full Event Interpretation (FEI) is part of BASF2 software package [22]. The algorithm is implemented purely in Python and takes care of:

- reconstructing a user-defined multi-level decay topology;

- training mutually dependent multivariate classifiers (MVCs) for each decay channel;
- determining sensible channel-specific pre-cuts and particle-specific post cuts to reduce combinatorics;
- generating a document summarizing the key performance indicators and control plots of all decay channels and particles used in the FEI.

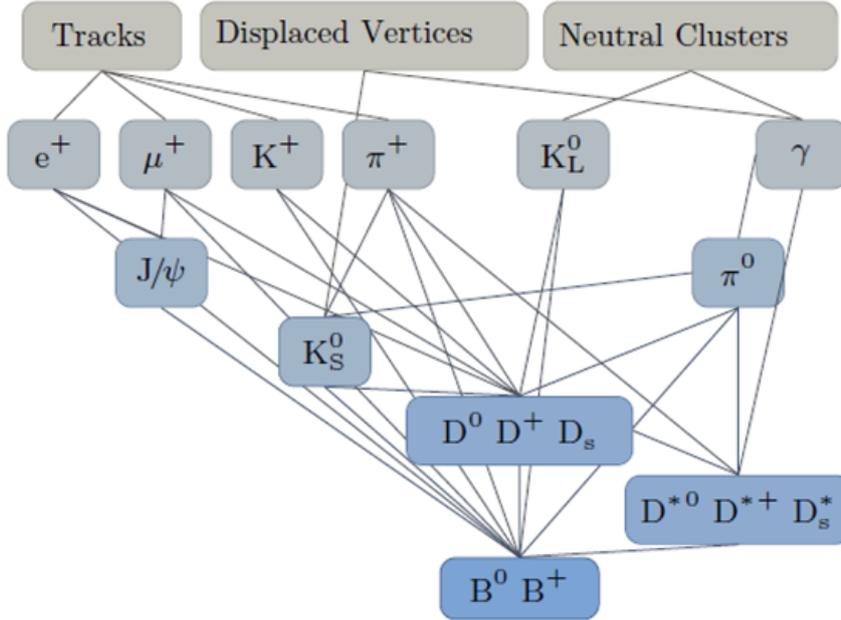


Figure 4.2: Hierarchical approach of the Full Event Interpretation (FEI).

The reconstructed decay topology, methods, and variables used in the multivariate classification are defined by the user. The FEI automatically reconstructs one out of the two B mesons in an  $Y(4S)$  decay to recover information about the remaining B meson. In fact, there is an entire class of analysis methods (tagging-methods) based on this concept. There are three distinct tagging-methods:

- hadronic tagging: uses hadronic decay channels for the reconstruction. Hence, the kinematics of the reconstructed candidates are well known and the tagged sample is the purest. This tagging is only possible for a tiny fraction of the dataset on the order of a few per mille;
- semileptonic tagging: uses semileptonic B decay channels. Due to the high branching ratio of semileptonic decays, this approach usually has a higher tagging efficiency. On the other hand, the semileptonic reconstruction suffers from missing kinematic information due to the neutrino in the final state of the decay. Hence, the sample is not as pure as in the hadronic case;
- inclusive tagging (or no-tagging): combines the four-momenta of all particles in the rest of the event of the signal-side B candidate. The achieved tagging efficiency is usually one order of magnitude above the hadronic and semileptonic tagging. Yet

the decay topology is not explicitly reconstructed and cannot be used to discard wrong candidates. In consequence, this method suffers from a high background and the tagged sample is very impure.

The FEI is an exclusive tagging algorithm and automatically constructs plausible  $B_{tag}$  meson decay chains compatible with the observed tracks and clusters, and calculates for each decay chain the probability of it correctly describing the true process using gradient-boosted decision trees. *Exclusive* refers to the reconstruction of a particle (here the  $B_{tag}$ ) assuming an explicit decay channel. Consequently, exclusive tagging reconstructs the  $B_{tag}$  independently of the  $B_{sig}$  using either hadronic ( $B \rightarrow DX$ ) or semileptonic ( $B \rightarrow D^*l\nu$ ) B meson decay channels.

The basic idea of the Full Event Interpretation is to reconstruct the particles and train the multivariate classifiers in a hierarchical approach, using a high number of combined B decay channels. The approach is depicted in Fig. 4.2. At first, the final-state particle candidates are selected and corresponding classification methods are trained using the detector information. Building on this, intermediate particle candidates are reconstructed and a multivariate classifier is trained for each employed decay channel. The MVC combines all information about a candidate into a single value, i.e. the signal-probability ( $sigProb$ ). In consequence, candidates from different decay channels can be treated equally in the following reconstruction steps. The B candidates are reconstructed and the corresponding classifiers are trained. The final output of the FEI to the user contains four particle lists:  $B^+$ , hadronic;  $B^+$ , semileptonic;  $B^0$ , hadronic;  $B^0$ , semileptonic. In the case that multiple candidates are reconstructed in the event the one with the highest FEI discriminant is chosen.

In this analysis, a hadronic tagging method is exploited.

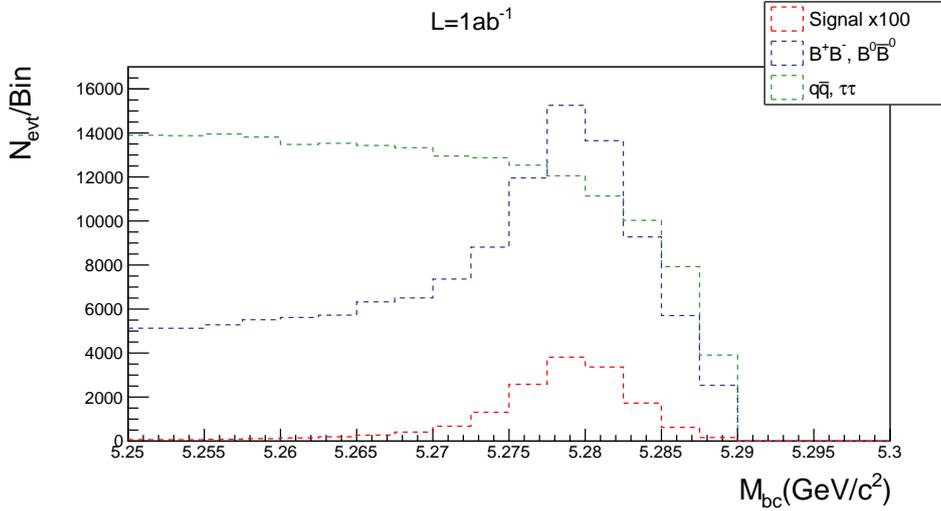


Figure 4.3: Plot of  $M_{bc}$  for  $e^\pm$  channel with *loose* cuts; for  $\mu^\pm$ , single  $\pi^\pm$  and  $\rho$  channels the shapes are similar.

Together with the FEI output, two kinematic variables are used to discriminate between correctly reconstructed  $B_{tag}$  candidates and misreconstructed events: the beam energy-constrained mass  $M_{bc} = \sqrt{s/4 - p_B^2}$  and the energy difference  $\Delta E = E_B - \sqrt{s}/2$ , where  $\sqrt{s}$  is the total energy in the  $Y(4S)$  center-of-mass system (CM) and  $p_B$  and  $E_B$  respectively denote the momentum and the energy of the  $B_{tag}$  candidate in the CM. Events with a  $B_{tag}$  candidate arise from two possible classes with different  $M_{bc}$

distributions. One class includes signal events with a correctly reconstructed  $B_{tag}$ , and background events from  $Y(4S) \rightarrow B^+B^-$  with a correctly reconstructed  $B_{tag}$ . All these events are characterized by an  $M_{bc}$  distribution peaked at the nominal B mass (signal and peaking background) (Fig. 4.3). The other classes of events consist of continuum background,  $e^+e^- \rightarrow q\bar{q}$  ( $q = u, d, c, s$ ) and  $e^+e^- \rightarrow \tau^+\tau^-$ , and combinatorial background,  $Y(4S) \rightarrow B^0\bar{B}^0$  or  $B^+B^-$  in which the  $B_{tag}$  is misreconstructed. These events are characterized by a smooth  $M_{bc}$  distribution.

A *loose* pre-selection is applied in order to reduce the number of  $B_{tag}$  candidates:

- $M_{bc} > 5.24 \text{ GeV}$ ;
- $|\Delta E| < 200 \text{ MeV}$ ;
- $sigProb > 0.0001$ .

## 4.4 Signal Events Selection

The objective of this section is to optimize a series of cuts for each  $\tau$  decay channel considered that maximize the significance  $Z = s/\sqrt{b}$ <sup>1</sup> and minimize a figure of merit (FOM) obtained through a Toy-MC study.

In the events where a  $B_{tag}$  is reconstructed, the presence of only one additional opposite charged track is required, consistent with a 1-prong  $\tau$  decay mode. The charged particle identification (PID) relies on likelihood-based selectors: the information from the detector systems, i.e. specific ionization ( $dE/dx$ ) from the SVD and the CDC,  $E/p$  from ECL and measurements from TOP, ARICH, and KLM are analyzed independently to determine a likelihood for each charged particle hypothesis (electron, muon, pion, kaon, proton, and deuteron). The likelihoods from each detector are used to construct a combined likelihood ratio:

$$\frac{\mathcal{L}(particle)}{\mathcal{L}(e) + \mathcal{L}(\mu) + \mathcal{L}(\pi) + \mathcal{L}(K) + \mathcal{L}(p) + \mathcal{L}(d)} \quad (4.1)$$

Particle identification criteria are applied to select the following  $\tau$  decay modes:

- $\tau \rightarrow e\nu_e\nu_\tau$ ,
- $\tau \rightarrow \mu\nu_\mu\nu_\tau$ ,
- $\tau \rightarrow \pi\nu_\tau$
- $\tau \rightarrow \rho\nu_\tau$  with  $\rho \rightarrow \pi^\pm\pi^0$

The selected categories all together correspond to approximately 71.5% of all  $\tau$  decays.

---

<sup>1</sup>A generic definition of the significance  $Z$  is the number of standard deviations which, in a normal distribution, would give the same p-value. According to Wilk's theorem,  $-\ln \lambda$ , where  $\lambda$  is a likelihood-ratio that meets particular conditions, is distributed as a  $\chi^2$ , then, in general, holds  $Z = \sqrt{-\ln \lambda}$ . The particular expression used is obtained from the general one in the limit case of event counting and  $b \gg s$ . In this expression,  $s$  stands for the signal events number and  $b$  for the background events one.

Before the optimization, tracks satisfying  $electronID > 0.1$  are taken as electron candidates for the  $e\nu_e\nu_\tau$  mode. Tracks not passing the  $electronID$  selection and satisfying  $muonID > 0.1$  are taken as muon candidates for the  $\mu\nu_\mu\nu_\tau$  mode. Tracks not passing the  $electronID$  and the  $muonID$  selections and satisfying  $pionID > 0.1$  are taken as pion candidates for both the  $\pi\nu_\tau$  and  $\pi^\pm\pi^0\nu_\tau$  modes. In addition, for the specific  $\pi^\pm\pi^0\nu_\tau$  mode one  $\pi^0$  candidate is reconstructed and the invariant mass of the state  $\pi^\pm\pi^0$  is required to be within  $625 \text{ MeV}/c^2 < M_{\pi^\pm\pi^0} < 925 \text{ MeV}/c^2$ . In turn, the  $\pi^0$  candidates are reconstructed by pairing two neutral clusters and applying the invariant mass cut on the  $\gamma\gamma$  pair of  $110 \text{ MeV}/c^2 < m_{\gamma\gamma} < 160 \text{ MeV}/c^2$ .

As described in Sec. 4.3,  $M_{bc}$  and  $sigProb$  are two kinematic variables used to discriminate between correctly reconstructed  $B_{tag}$  candidates and misreconstructed events. So, in order to suppress misreconstructed  $B_{tag}$ , before starting the optimization, is required for each  $\tau$  decay channel:

- $M_{bc} > 5.27 \text{ GeV}$  (Fig. 4.3)
- $sigProb > 0.01$  (Fig. 4.4)

The plot with signal and background normalized to the same area is also reported for  $sigProb$  to justify the cut (Fig. 4.5). The cut on  $sigProb$  will be optimized in the following, while the cut on  $M_{bc}$  will be kept the same.

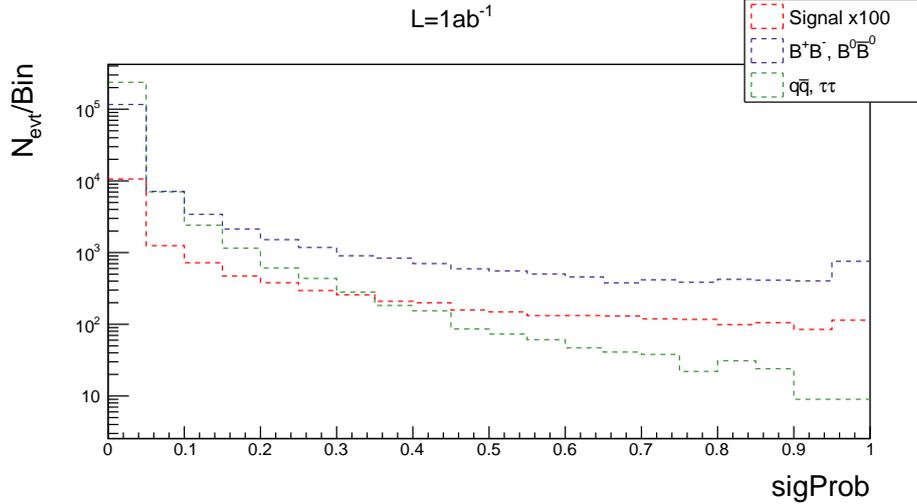


Figure 4.4: Plot of  $sigProb$  with *loose* cuts; the shape is the same for all PID signal hypotheses since it depends only on the tag side.

The best discriminant variable between the  $B \rightarrow \tau\nu$  signal and all of the background contributions is the extra energy in the electromagnetic calorimeter ( $E_{extra}$ ).  $E_{extra}$  (Fig. 4.6) is defined as the sum of the energy deposits in the calorimeter that cannot be directly associated with the reconstructed decay tree of the  $B_{tag}$  or the  $B_{sig}$ .

For signal events,  $E_{extra}$  must be either zero or a small value arising from beam background hits and imperfect reconstruction (since neutrinos do not interact in the ECL). Moreover, most background events are distributed toward higher  $E_{extra}$  due to the contribution from additional clusters produced by unassigned tracks and neutrals from the misreconstructed  $B_{tag}$  or  $B_{sig}$  mesons. To study the signal region and fit the BR, this cut on  $E_{extra}$  will not be changed.

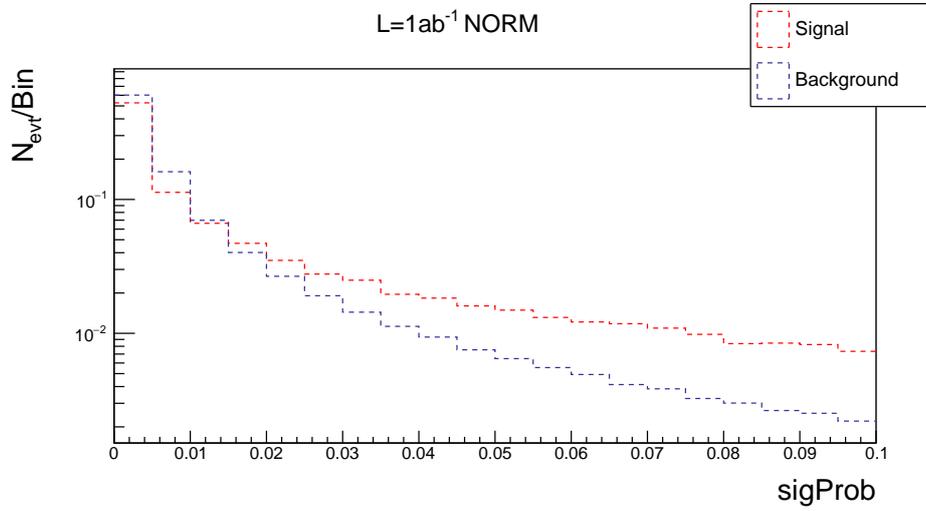


Figure 4.5: Plot of  $sigProb$  with *loose* cuts; signal and background are normalized to the same area; the shape is the same for all PID signal hypotheses since it depends only on the tag side.

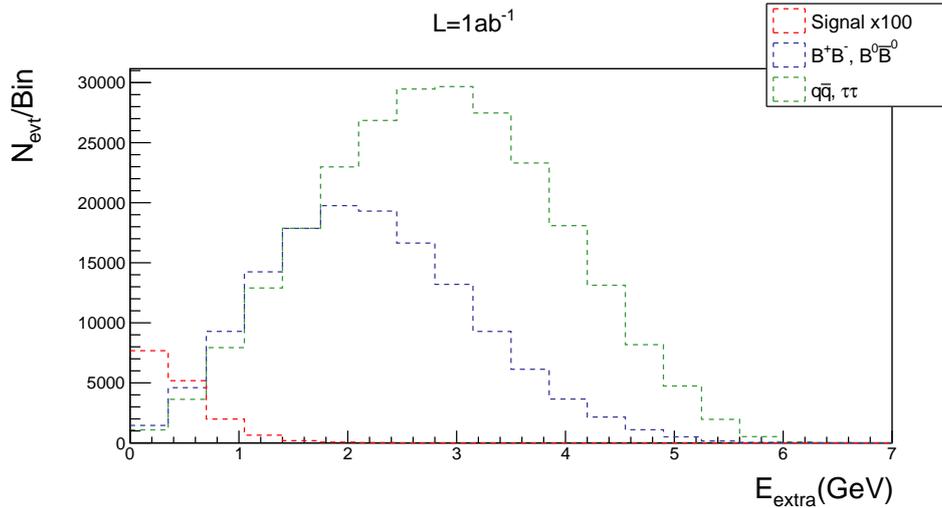


Figure 4.6: Plot of  $E_{extra}$  for  $e^\pm$  channel with *loose* cuts; for  $\mu^\pm$ , single  $\pi^\pm$  and  $\rho$  channels the shapes are similar.

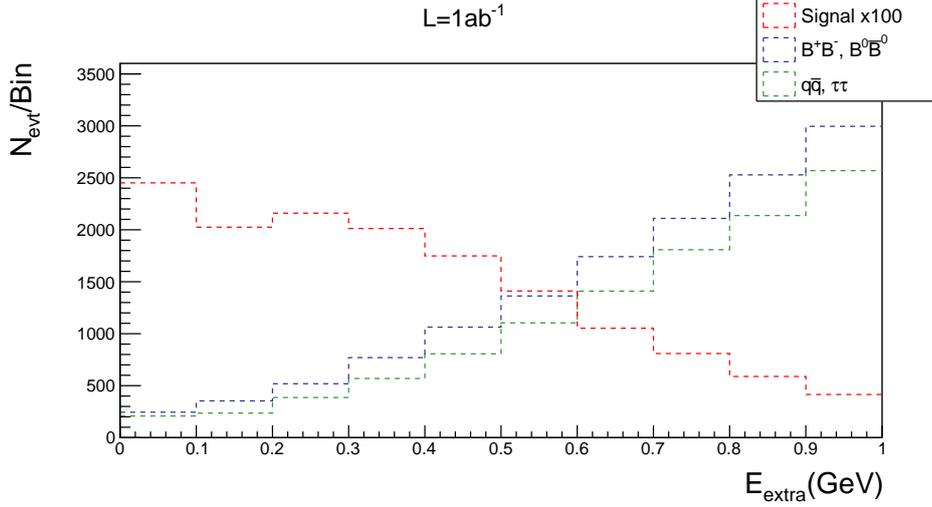


Figure 4.7: Plot of  $E_{extra}$  for  $e^\pm$  channel between 0 and 1 GeV with *loose* cuts; for  $\mu^\pm$ , single  $\pi^\pm$  and  $\rho$  channels the shapes are similar.

A powerful constrain to suppress the background is the missing mass squared  $M_{miss}^2$  defined as:

$$M_{miss}^2 = E_{miss}^2 - P_{miss}^2 = (2E_{beam} - E_{B_{tag}} - E_{B_{sig}})^2 - (\vec{p}_{B_{tag}} - \vec{p}_{B_{sig}})^2 \quad (4.2)$$

where  $E_{miss}$  and  $P_{miss}$  are respectively the missing energy and the missing momentum.

Leptonic  $\tau$  decay channels are characterized by a high  $M_{miss}^2$  (Fig. 4.8), while hadronic decay channels by a low one (Fig. 4.9). Specifically, cutting on this quantity reduces the combinatorial background for leptons and the continuum one for hadrons. As for *sigProb*, the cut on  $M_{miss}^2$  will be optimized for each channel but  $\rho$  one.

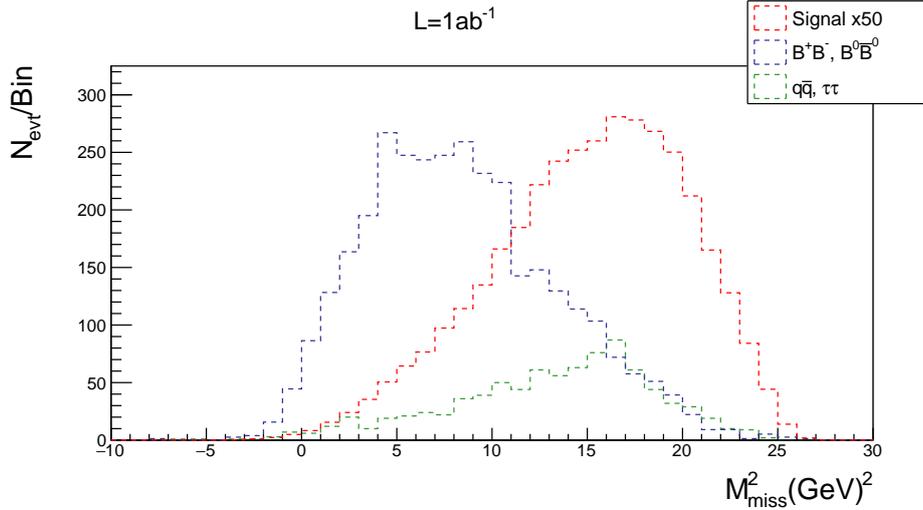


Figure 4.8: Plot of  $M_{miss}^2$  for  $e^\pm$  channel with cuts on  $M_{bc}$ , *sigProb* and  $E_{extra}$ ; for  $\mu^\pm$  channel the shapes is similar.

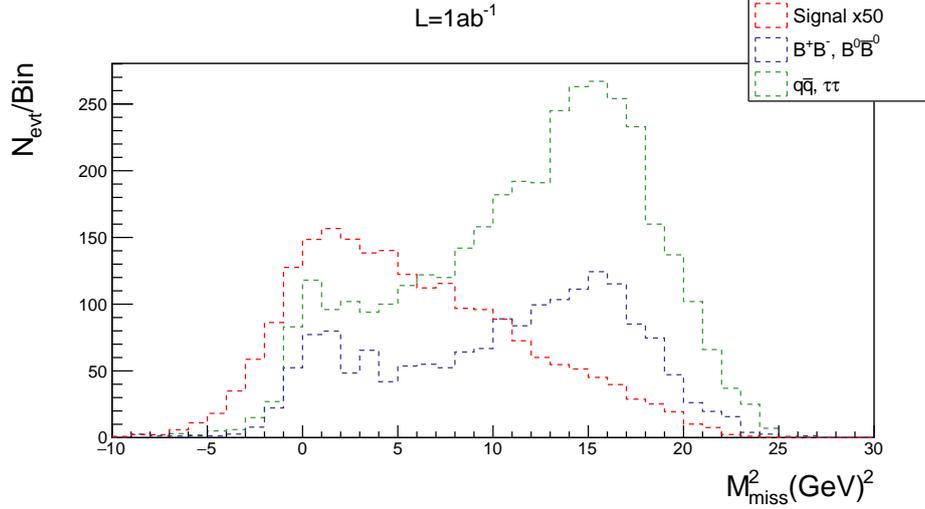


Figure 4.9: Plot of  $M_{miss}^2$  for single  $\pi^\pm$  channel with cuts on  $M_{bc}$ ,  $sigProb$  and  $E_{extra}$ ; for  $\rho$  channel the shapes is similar.

The last variable used is the charged particle momentum  $p$  which, together with the missing mass, is also useful to reduce the signal channels cross feeds (Sec. 4.5).

As for the missing mass, the charged particle momentum exhibits different behavior for leptons and hadron final states. The spectrum is softer for leptons because the lepton originates from a  $\tau$  three-body decay, while the charged particle in hadronic modes originates from a two-body decay ( $\tau \rightarrow \pi\nu$ ) and from two subsequent two-body decays ( $\tau \rightarrow \rho\nu$ ,  $\rho \rightarrow \pi^\pm\pi^0$ ).

Only for  $\rho$  we will consider the missing momentum (Fig. 4.13), the charged particle momentum  $p^*$  (Fig. 4.14) and the neutral particle momentum  $p_0^*$  (Fig. 4.15) in the centre-of-mass reference frame.

We will optimize the selection cuts on  $p$  for the  $\pi^\pm$  channel and on  $P_{miss}$ ,  $p^*$  and  $p_0^*$  for the  $\rho$  one.

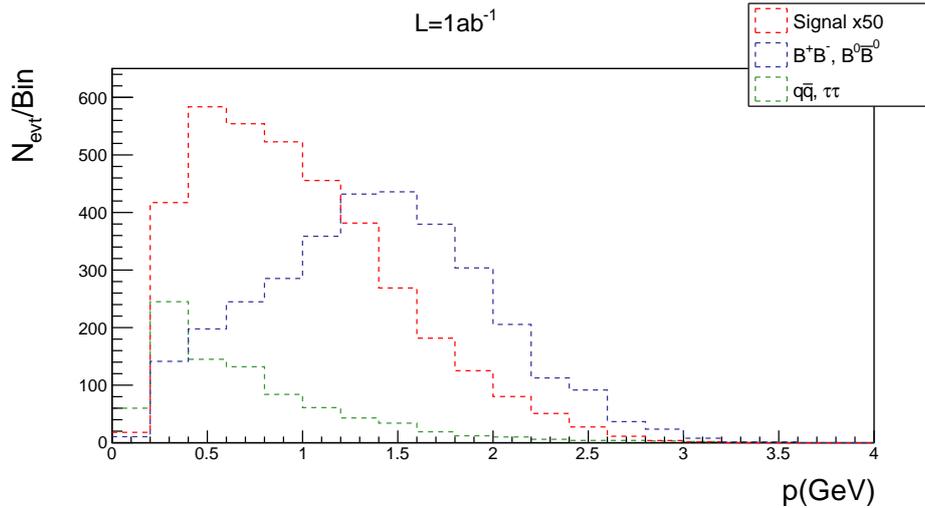


Figure 4.10: Plot of  $p$  for  $e^\pm$  channel with cuts on  $M_{bc}$ ,  $sigProb$  and  $E_{extra}$ ; only for  $\mu^\pm$  channel the shapes is similar.

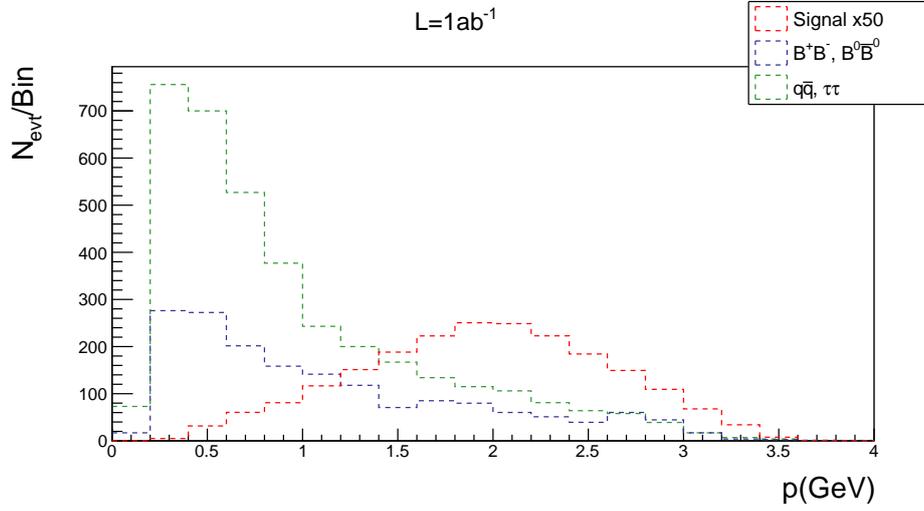


Figure 4.11: Plot of  $p$  for single  $\pi^\pm$  channel with cuts on  $M_{bc}$ ,  $sigProb$  and  $E_{extra}$ .

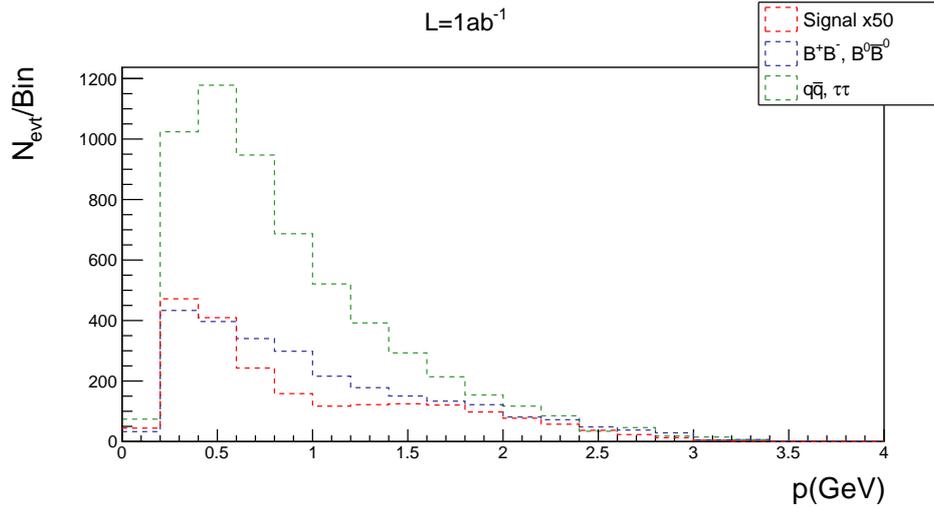


Figure 4.12: Plot of  $p$  for  $\pi^\pm$  from  $\rho$  channel with cuts on  $M_{bc}$ ,  $sigProb$  and  $E_{extra}$ .

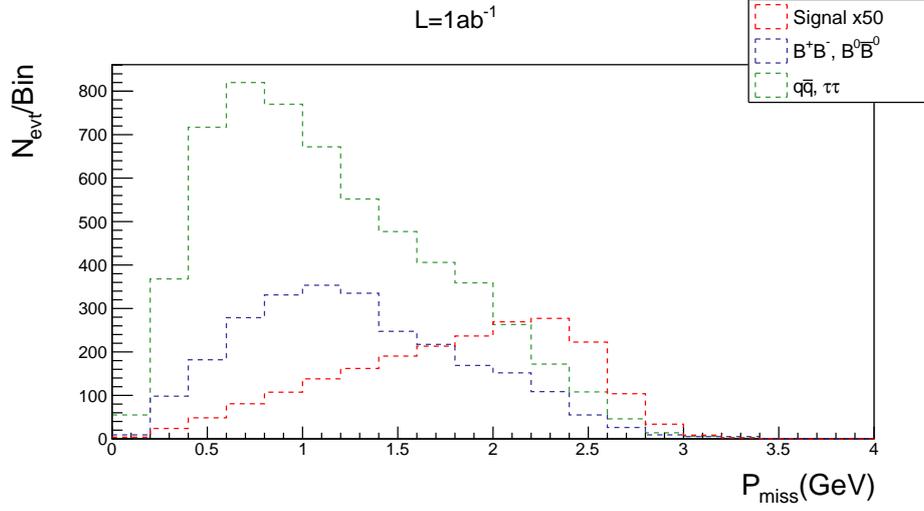


Figure 4.13: Plot of  $P_{miss}$  for  $\rho$  channel with cuts on  $M_{bc}$ ,  $sigProb$  and  $E_{extra}$ .

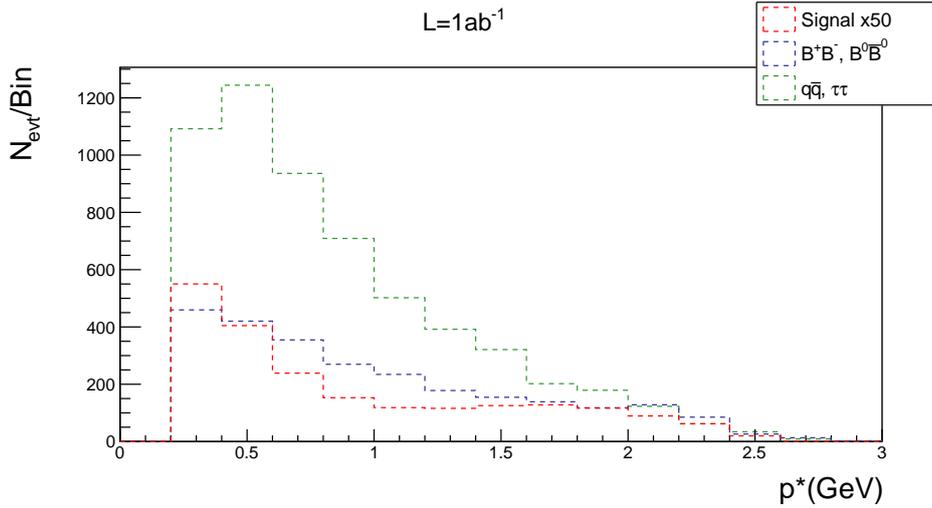


Figure 4.14: Plot of  $p^*$  for  $\pi^\pm$  from  $\rho$  channel with cuts on  $M_{bc}$ ,  $sigProb$  and  $E_{extra}$ .

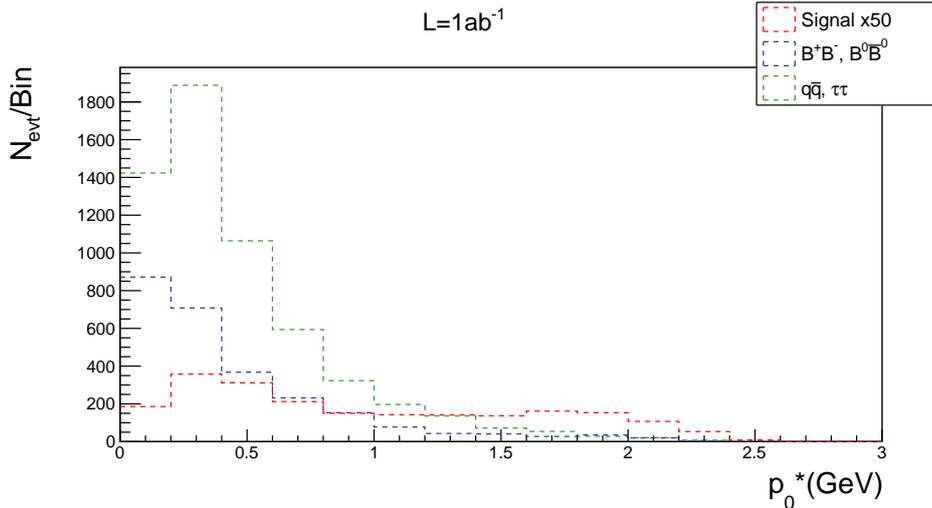


Figure 4.15: Plot of  $p_0^*$  for  $\pi^0$  from  $\rho$  channel with cuts on  $M_{bc}$ ,  $sigProb$  and  $E_{extra}$ .

#### 4.4.1 Signal Selection Optimization

The signal selection optimization is performed by cutting on the variables introduced in the previous section for each of the different  $\tau$  decay channels. To do this, we will choose the best cuts through an iterative procedure: we vary one quantity, fixing the others until a minimum is reached. In particular, we will see on the one hand how leptons ( $e^\pm$  and  $\mu^\pm$ ) share almost the same cuts because of their similar properties, on the other how hadrons (single  $\pi^\pm$  and  $\pi^\pm\pi^0$  pairs) differ for the presence of the neutral pion in  $\rho$  channel.

For each step of the selection optimization, a Toy-MC study is performed: 10000 pseudo-experiments are produced and for each of them a certain pseudo-dataset is generated according to the signal and background MC expectations, with Poisson fluctuations. An extended binned maximum likelihood fit to  $E_{extra}$  is performed in each of these pseudo-experiments using a two-component parameterized function, in which the  $E_{extra}$  distributions for signal and background events are taken from simulation, in order to extract the signal ( $N_S$ ) and background ( $N_B$ ) yields with their respective errors ( $\sigma_S$  and  $\sigma_B$ ). For each step in the optimization, it is evaluated the figure of merit (FOM) below:

$$FOM = \frac{\bar{\sigma}_S}{\bar{N}_S} \quad (4.3)$$

where  $\bar{N}_S$  and  $\bar{\sigma}_S$  are the average values of  $N_S$  and  $\sigma_S$  of these 10000 pseudo-experiments. This FOM is nothing more than the relative statistical uncertainty on the average signal yield.

The optimal cut will be the one with the lowest FOM, paying attention to those points that do not really represent a minimum, but only a statistical fluctuation. If there are several minima with the same FOM value, the one with the highest significance  $Z = s/\sqrt{b}$  is chosen, where  $s$  and  $b$  are respectively the number of signal and background events with  $E_{extra} < 0.5 \text{ GeV}$ .

## Electrons

The quantities optimized and their best cuts are:

$$\begin{array}{l|l} \mathit{electronID} & > 0.6 \\ \mathit{sigProb} & > 0.03 \\ M_{miss}^2(\text{GeV}^2) & > 12 \end{array}$$

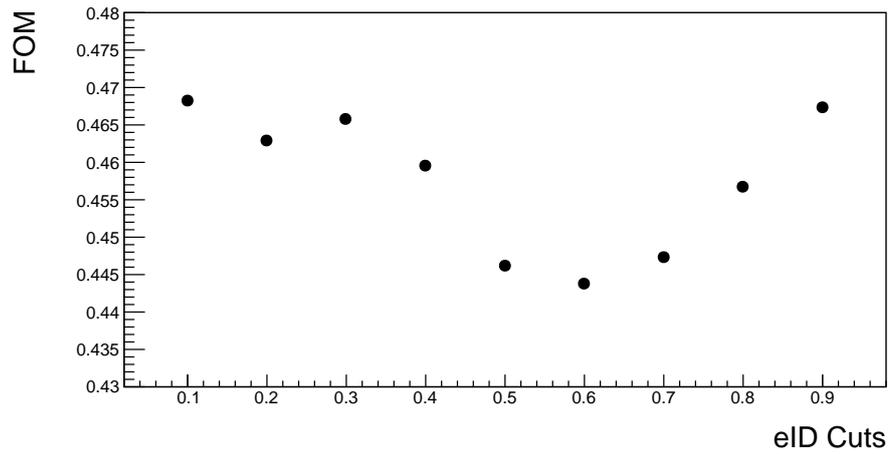


Figure 4.16: Plot of the FOM as function of cuts on  $\mathit{electronID}$  for the  $e^\pm$  channel.

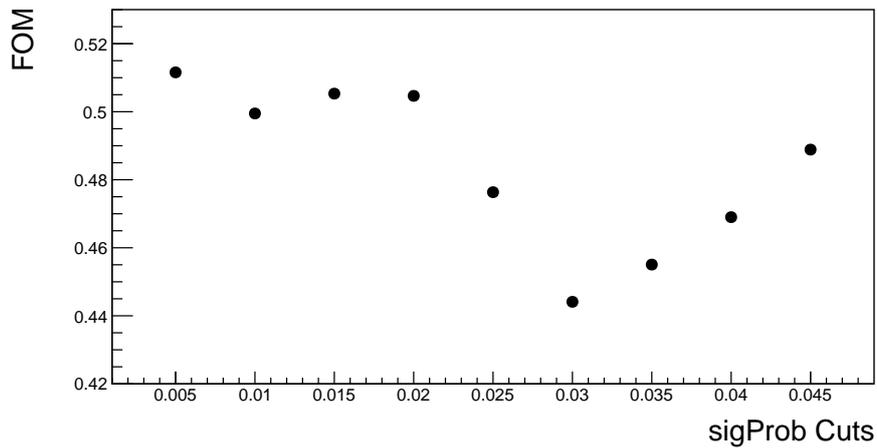


Figure 4.17: Plot of the FOM as function of cuts on  $\mathit{sigProb}$  for the  $e^\pm$  channel.

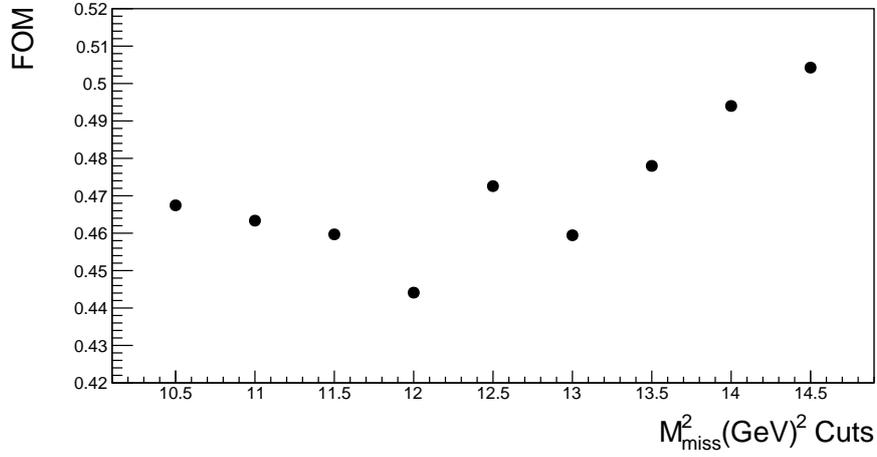


Figure 4.18: Plot of the FOM as function of cuts on  $M_{miss}^2$  for the  $e^\pm$  channel.

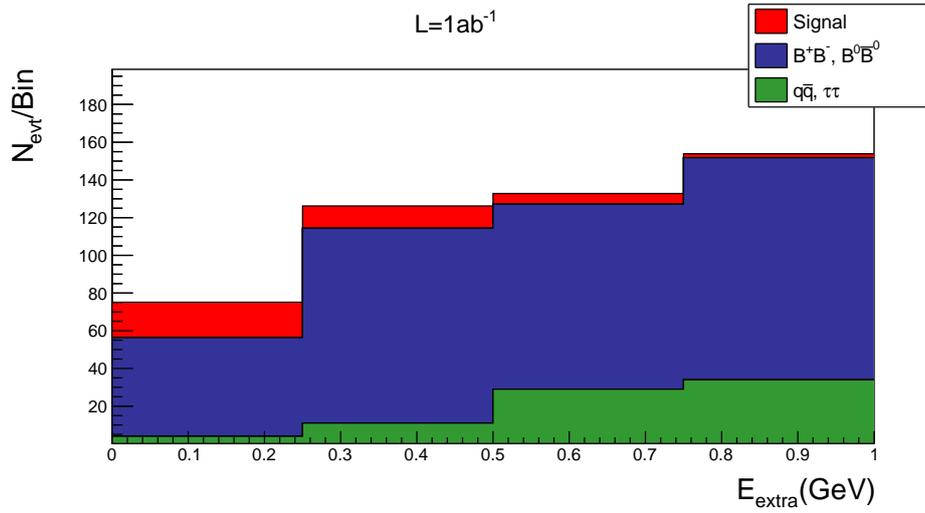


Figure 4.19: Stacked plot of  $E_{extra}$  applying the best cuts for the  $e^\pm$  channel.

## Muons

The quantities optimized and their best cuts are:

$electronID$	$< 0.6$
$muonID$	$> 0.4$
$sigProb$	$> 0.03$
$M_{miss}^2 (GeV^2)$	$> 11.5$

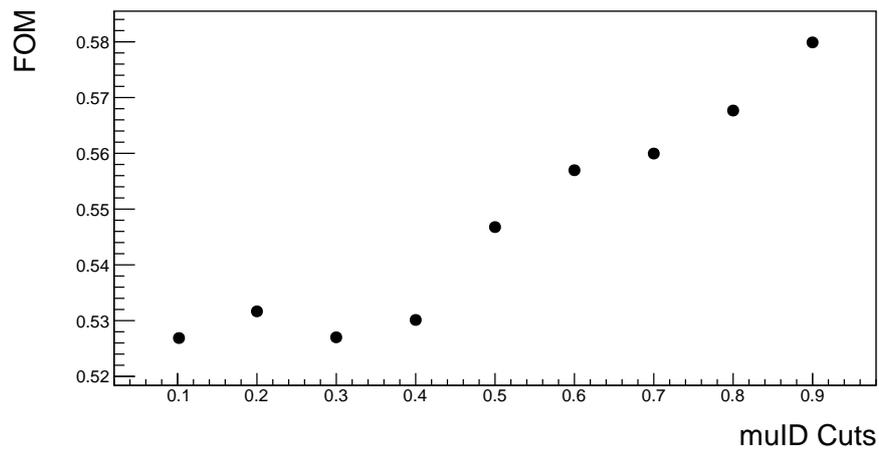


Figure 4.20: Plot of the FOM as function of cuts on  $muonID$  for the  $\mu^\pm$  channel.

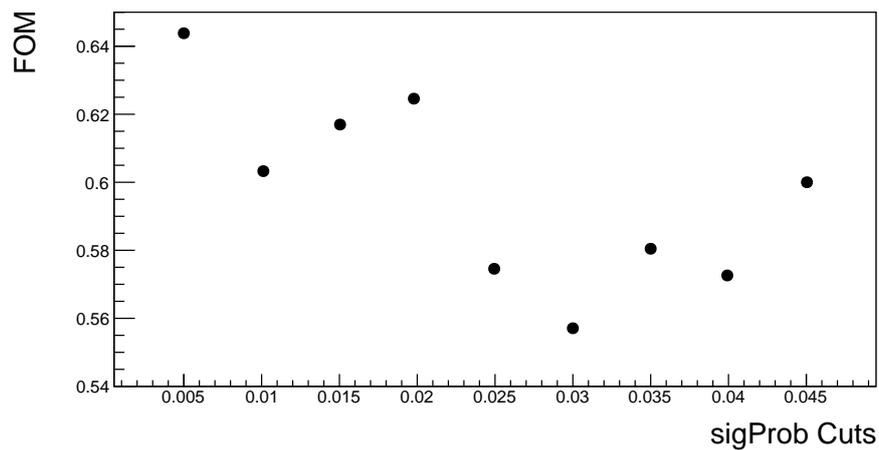


Figure 4.21: Plot of the FOM as function of cuts on  $sigProb$  for the  $\mu^\pm$  channel.

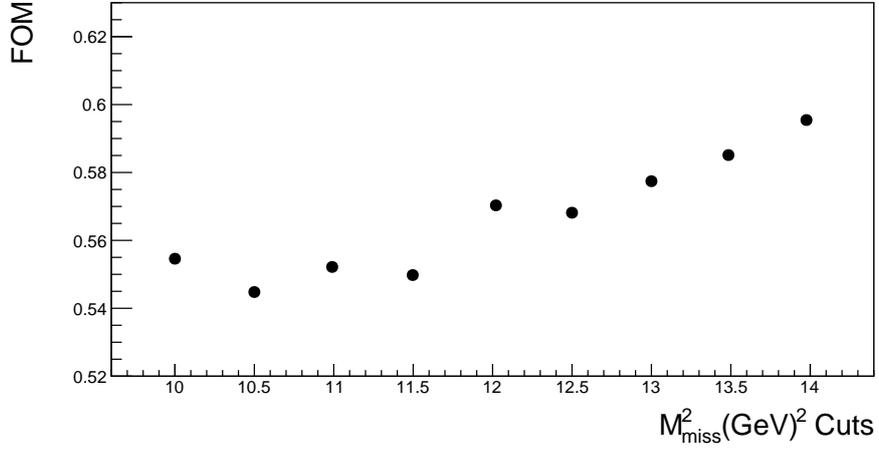


Figure 4.22: Plot of the FOM as function of cuts on  $M_{miss}^2$  for the  $\mu^\pm$  channel.

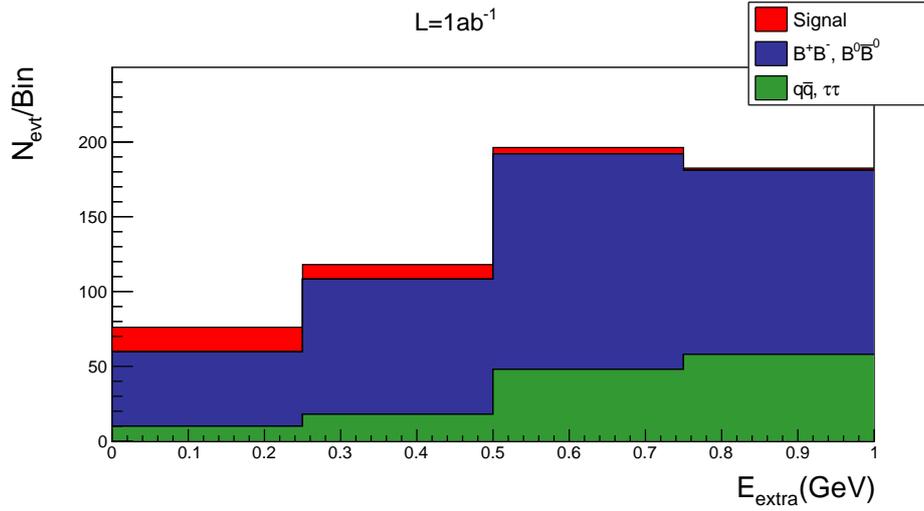


Figure 4.23: Stacked plot of  $E_{\text{extra}}$  applying the best cuts for the  $\mu^\pm$  channel.

## Single Pions

The quantities optimized and their best cuts are:

$electronID$	$< 0.6$
$muonID$	$< 0.4$
$pionID$	$> 0.2$
$sigProb$	$> 0.04$
$M_{miss}^2 (GeV^2)$	$< 10.5$
$p (GeV)$	$> 1.4$

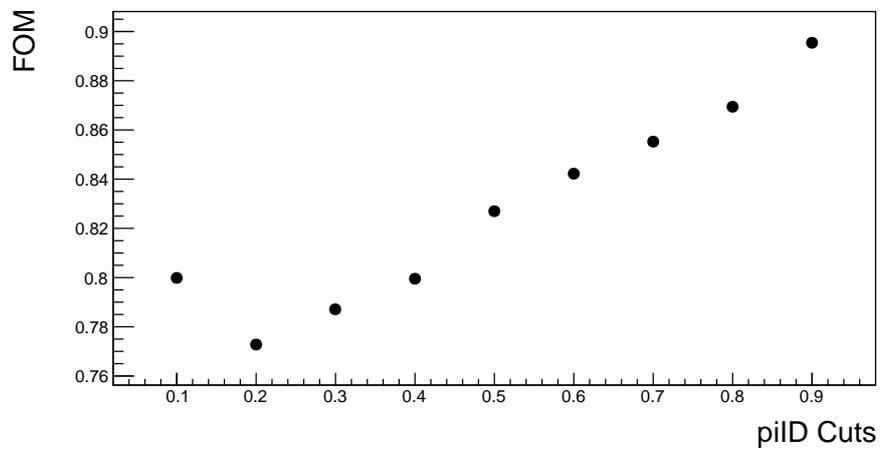


Figure 4.24: Plot of the FOM as function of cuts on  $pionID$  for the  $\pi^\pm$  channel.

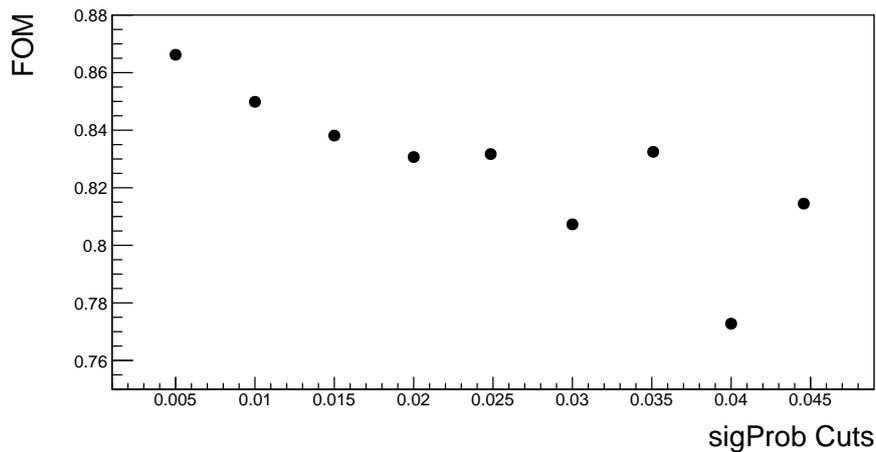


Figure 4.25: Plot of the FOM as function of cuts on  $sigProb$  for the  $\pi^\pm$  channel.

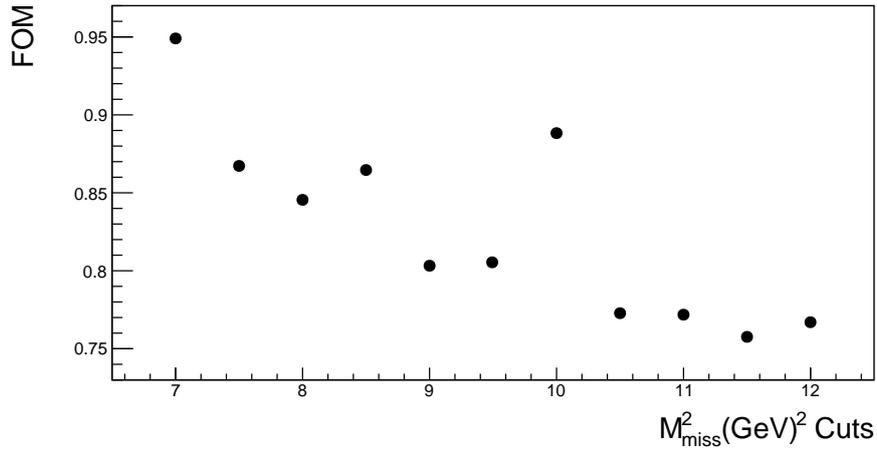


Figure 4.26: Plot of the FOM as function of cuts on  $M_{miss}^2$  for the  $\pi^\pm$  channel.

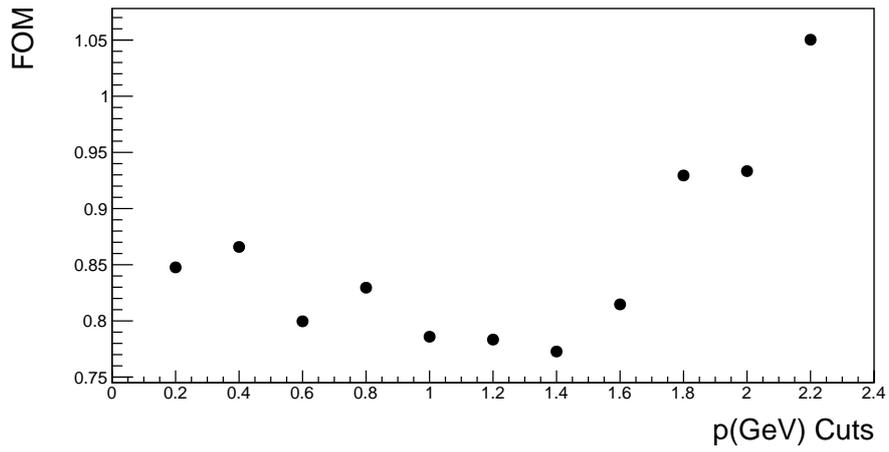


Figure 4.27: Plot of the FOM as function of cuts on  $p$  for the  $\pi^\pm$  channel.

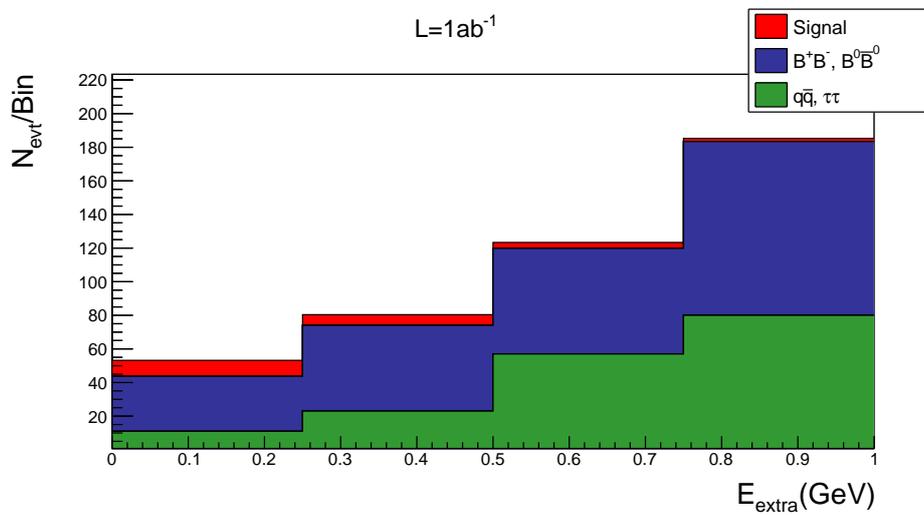


Figure 4.28: Stacked plot of  $E_{\text{extra}}$  applying the best cuts for the  $\pi^\pm$  channel.

## Rho

The quantities optimized and their best cuts are:

$electronID$	$< 0.6$
$muonID$	$< 0.4$
$pionID$	$> 0.2$
$sigProb$	$> 0.035$
$P_{miss}(GeV)$	$> 1.3$

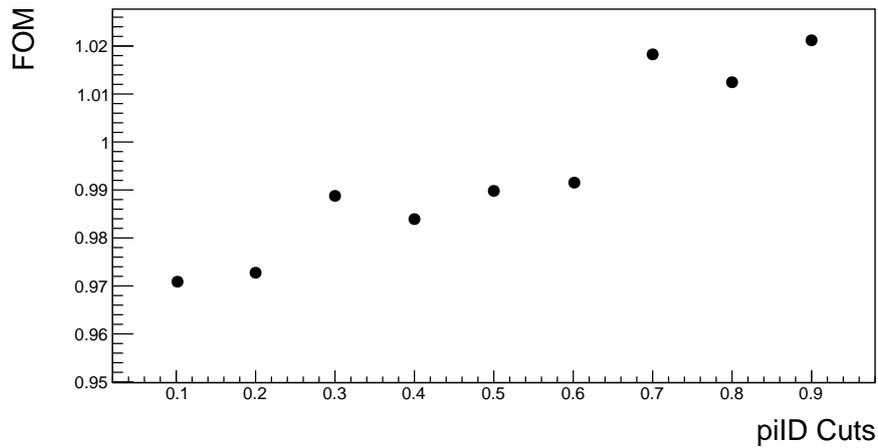


Figure 4.29: Plot of the FOM as function of cuts on  $pionID$  for the  $\pi^\pm$  from  $\rho$  channel.

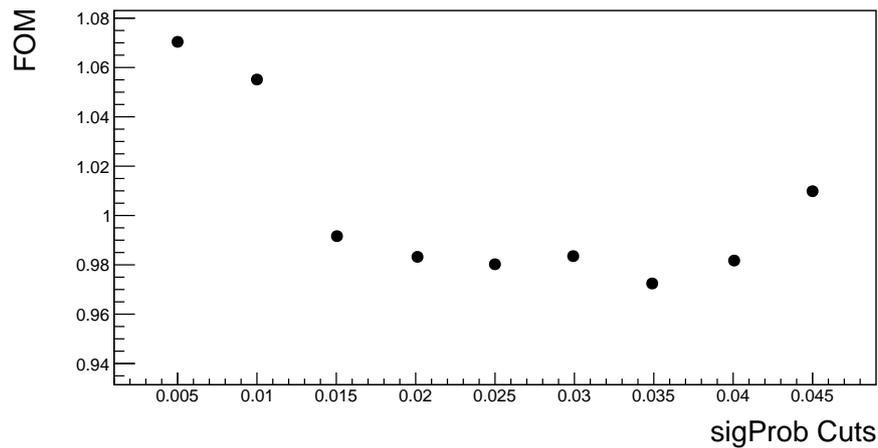


Figure 4.30: Plot of the FOM as function of cuts on  $sigProb$  for  $\rho$  channel.

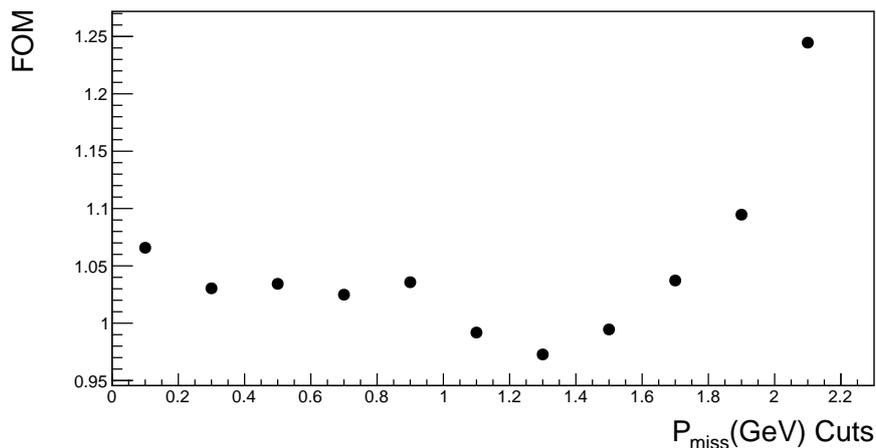


Figure 4.31: Plot of the FOM as function of cuts on  $P_{miss}$  for the  $\rho$  channel.

In order to get a better signal-background ratio, rather than individually considering  $p^*$  and  $p_0^*$  and cutting over them, a different approach has been attempted: building a Likelihood Ratio ( $\mathcal{LR}$ ) (Fig. 4.32) using the 2D PDFs obtained from the 2D histograms of these two variables for signal (Fig. 4.33a) and background (Fig. 4.33b) derived from the MC distributions and cutting over that instead. The  $\mathcal{LR}$  has the following expression:

$$\mathcal{LR}_{(p^*, p_0^*)} = \frac{\mathcal{P}_s(p^*, p_0^*)}{\mathcal{P}_b(p^*, p_0^*)} \quad (4.4)$$

where  $\mathcal{P}_s$  and  $\mathcal{P}_b$  are respectively the signal and background 2D PDFs. With the same iterative optimization procedure used for the other variables (Fig. 4.34), the best cut that results is  $\mathcal{LR}_{(p^*, p_0^*)} > 1$ . This variable occurs to be more discriminating than the individual two.

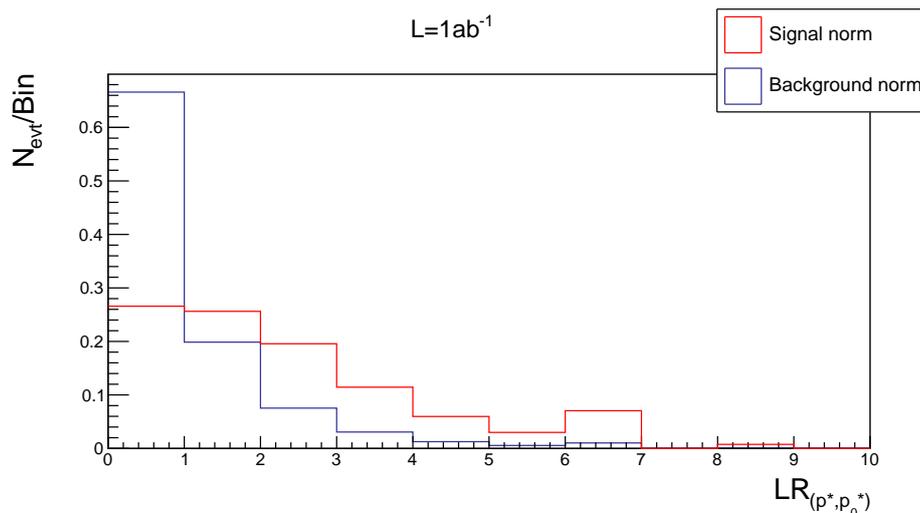


Figure 4.32: Plot of  $\mathcal{LR}_{(p^*, p_0^*)}$  for single  $\pi^\pm$  from  $\rho$  channel with the best cuts for *electronID*, *muonID*, *pionID*, *sigProb* and  $P_{miss}$  with signal and background normalised to the same area.  $\mathcal{LR}_{(p^*, p_0^*)}$  is evaluated on signal events (red) and on background ones (blue).

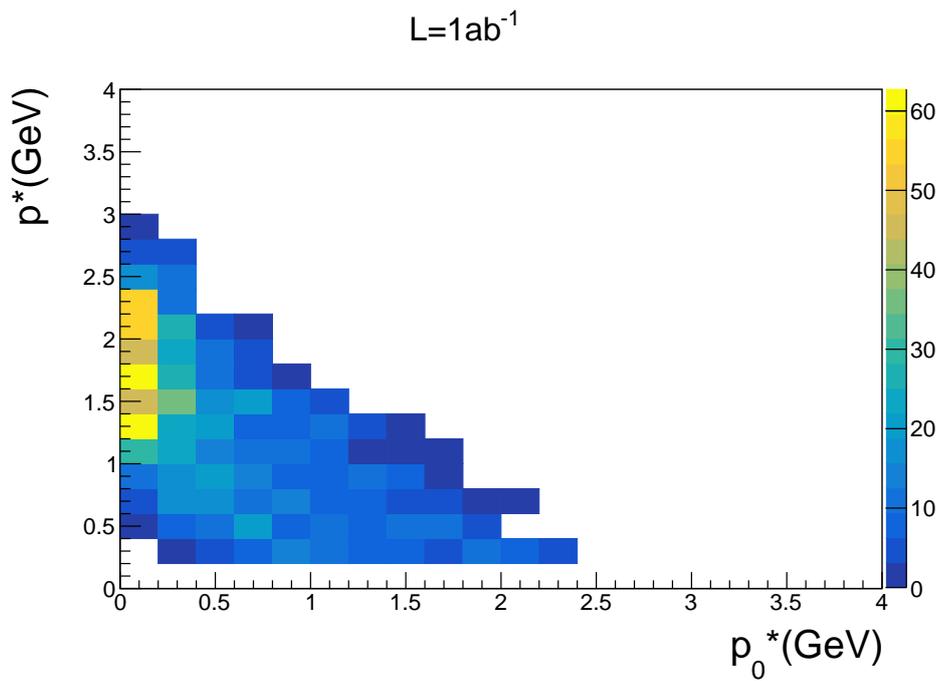
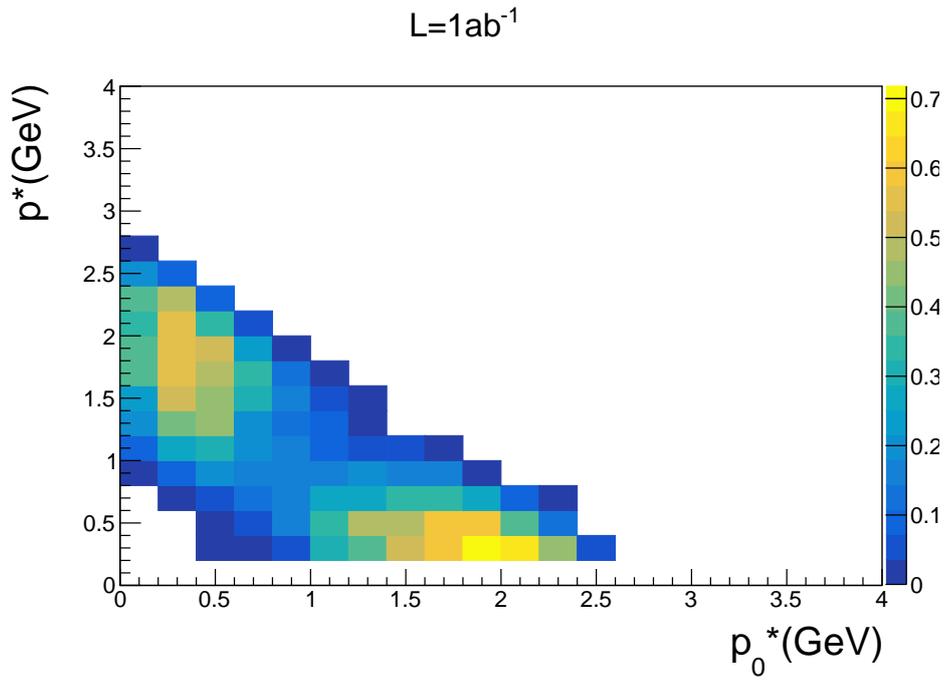


Figure 4.33: 2D plots of  $p^*$  vs  $p_0^*$  for the  $\rho$  channel with the best cuts for  $electronID$ ,  $muonID$ ,  $pionID$ ,  $sigProb$  and  $P_{miss}$  for signal (a) and for background (b).

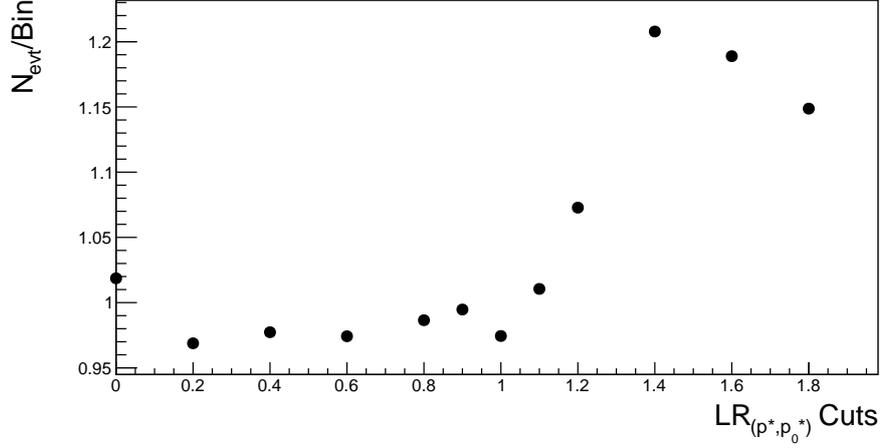


Figure 4.34: Plot of the FOM as function of cuts on  $\mathcal{LR}_{(p^*, p_0^*)}$  for the  $\rho$  channel.

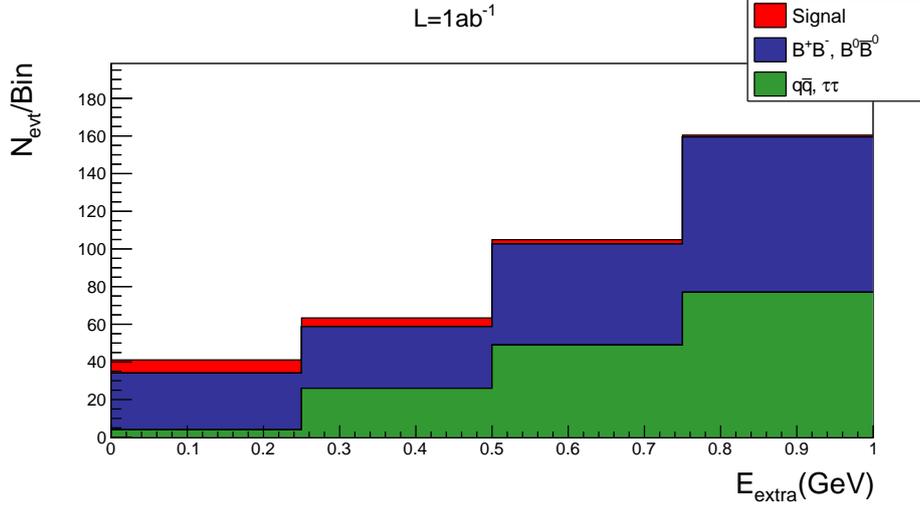


Figure 4.35: Stacked plot of  $E_{extra}$  applying the best cuts for the  $\rho$  channel.

Cuts	$e^\pm$	$\mu^\pm$	$\pi^\pm$	$\rho$
<i>sigProb</i>	> 0.03	> 0.03	> 0.04	> 0.035
$M_{bc}(GeV)$		> 5.27		
$E_{extra}(GeV)$		< 1		
<i>electronID</i>	> 0.6	< 0.6	< 0.6	< 0.6
<i>muonID</i>		> 0.4	< 0.4	< 0.4
<i>pionID</i>			> 0.2	> 0.2
$M_{miss}^2(GeV^2)$	> 12	> 11.5	< 10.5	
$P_{miss}(GeV)$				> 1.3
$p(GeV)$			> 1.4	
$\mathcal{LR}_{(p^*, p_0^*)}$				> 1

Table 4.2: Best cuts of each channel for signal events selection.

## 4.5 Signal Cross Feeds

After the signal selection, events may be found to be reconstructed in a  $\tau$  decay mode different with respect to the mode generated (cross feed), so, e.g. a true pion could pass the muon or electron selection, and so on. Due to the organization of the ROOT ntuple used, it is not possible to calculate the cross feeds of the  $\rho$  channel (with future implementation it will be possible).

Thus, for each decay channel, we will have different fractions:

$$f_{k,i} = \frac{N_{k,i}^{true}}{N_k^{reco}} \quad (4.5)$$

where  $i = (e, \mu, \text{single } \pi)$  denotes the true  $\tau$  decay channel,  $k = (e, \mu, \text{single } \pi, \pi \text{ from } \rho)$  denotes the reconstructed one,  $N_{k,i}^{true}$  is the number of  $i$  – mode decays, satisfying MC truth matching conditions, that pass the  $k$  – mode cut and  $N_k^{reco}$  is the number of all the reconstructed  $k$  – mode decays.

In Tab. 4.3 is reported the percentage of correctly reconstructed events that we expect and the cross feeds; with *other* we indicate the rest of the  $\tau$  decays that we do not consider.

Signal $\tau$ decay mode	Type	Fraction(%)
$e^\pm \nu_e \nu_\tau$	correctly-reconstructed	92.1
	$\mu^\pm \nu_\mu \nu_\tau$	1.2
	$\pi^\pm \nu_\tau$	0.2
	$\rho \nu_\tau$	0.3
	other	6.2
	$\mu^\pm \nu_\mu \nu_\tau$	$e^\pm \nu_e \nu_\tau$
$\mu^\pm \nu_\mu \nu_\tau$	correctly-reconstructed	90.4
	$\pi^\pm \nu_\tau$	1.3
	$\rho \nu_\tau$	1.9
	other	6.2
	$\pi^\pm \nu_\tau$	$e^\pm \nu_e \nu_\tau$
$\pi^\pm \nu_\tau$	$\mu^\pm \nu_\mu \nu_\tau$	1.1
	correctly-reconstructed	67.0
	$\rho \nu_\tau$	24.1
	other	7.7

Table 4.3: Summary of correctly-reconstructed efficiencies and cross feeds.

## 4.6 $B \rightarrow \tau\nu$ Branching Ratio Measurement

In this section, we will estimate the sensitivity to the measurement of the Branching Ratio of  $B \rightarrow \tau\nu$  with  $L = 1 \text{ ab}^{-1}$  of data. Two different approaches are investigated:

- through an extended maximum likelihood fit performed on  $E_{extra}$  with a Toy-MC method for each of the  $\tau$  decay channels considered, we estimate the BR for each channel and then perform a weighted average;
- through a simultaneous fit performed on  $E_{extra}$  with a Toy-MC method considering each of the  $\tau$  decays, we directly estimate the BR.

The BR for a generic decay is the fraction of particles that decay by an individual decay mode with respect to the total number of decaying particles. In our particular case:

$$BR(B \rightarrow \tau\nu) = \frac{N^{true}(B \rightarrow \tau\nu)}{N^{true}(B \rightarrow X)} = \frac{N^{true}(B \rightarrow \tau\nu)}{2L_{int}\sigma_{B^+B^-}} \quad (4.6)$$

where  $N^{true}(B \rightarrow \tau\nu)$  is the true number of B mesons that decay in  $\tau\nu$  and  $N^{true}(B \rightarrow X)$  is the true total number of B mesons that decay,  $L_{int}$  is the integrated luminosity and  $\sigma_{B^+B^-} = 0.5645 \text{ nb}$  is the  $B^+B^-$  cross-section;  $N^{true}(B \rightarrow X) = 2L_{int}\sigma_{B^+B^-}$  because  $L_{int}\sigma_{B^+B^-}$  is the number of total  $B^+B^-$  pairs produced at the superKEKB accelerator IP and we multiply it for 2 to obtain the total number of charged B mesons.

We use the apex <sup>true</sup> to distinguish the true number from the measured one, which will be linked to the true one through an efficiency  $\epsilon$ . Also, we do not directly measure the number of  $\tau$  particles from the B mesons, but we measure its decay products ( $e^\pm$ ,  $\mu^\pm$ ,  $\pi^\pm$  and  $\rho$ ). So we can rewrite our BR:

$$\begin{aligned} BR(B \rightarrow \tau\nu) &= \frac{N^{true}(B \rightarrow \tau\nu)}{2L_{int}\sigma_{B^+B^-}} = \frac{N^{measured}(B \rightarrow \tau\nu)}{2L_{int}\sigma_{B^+B^-} \cdot \epsilon} = \\ &= \frac{N^{measured}(\tau \rightarrow k - mode)}{2L_{int}\sigma_{B^+B^-} \cdot \epsilon_k} \end{aligned} \quad (4.7)$$

where  $k$  is one of the 4 particular decay channels considered,  $N^{measured}(\tau \rightarrow k - mode)$  is the number of measured decays products in the  $k - mode$  channel and  $\epsilon_k = N_k^{sel}/N^{gen}$  (not to be confused with cross feeds) is the efficiency introduced earlier for the particular  $k - mode$  channel;  $N_k^{sel}$  is the number of decays reconstructed in a particular channel and  $N^{gen}$  the one generated.

The extended maximum likelihood fits for each of the four  $\tau$  decay modes and the simultaneous fits are performed on 100000 generated pseudo-data (Toy-MC study) using MC PDFs and setting the BR to  $(1.09 \pm 0.24) \times 10^{-4}$ , the last reported in the PDG review [19].

### 4.6.1 Single Fits

In this section, we will present the results of the extended maximum likelihood fits for each  $\tau$  decay channel and calculate the weighted average BR with  $L = 1 \text{ ab}^{-1}$  of data.

The likelihood function for the  $N_k = n_{s,k} + n_{b,k}$  candidates reconstructed in one of the four  $\tau$  decay modes  $k$  is:

$$\mathcal{L}_k = \frac{e^{-(n_{s,k}+n_{b,k})}}{N_k!} \prod_{i=1}^{N_k} \{n_{s,k} \mathcal{P}_k^s(E_{i,k}) + n_{b,k} \mathcal{P}_k^b(E_{i,k})\} \quad (4.8)$$

where  $n_{s,k}$  is the signal yield,  $n_{b,k}$  is the background yield,  $E_{i,k}$  is the  $E_{extra}$  value of the  $i$ th event,  $\mathcal{P}_k^s$  is the PDF of signal events, and  $\mathcal{P}_k^b$  is the PDF of background events. The background yields in each decay mode are permitted to float independently of each other in the fit, while the signal yields are constrained to a single BR via the relation obtained from Eq. 4.7:

$$n_{s,k} = 2L_{int}\sigma_{B\bar{B}} \cdot \epsilon_k \cdot BR(B \rightarrow \tau\nu) \quad (4.9)$$

The parameters  $L_{int}$ ,  $\sigma_{B\bar{B}}$ , and  $\epsilon_k$  are fixed in the fit, while BR is allowed to vary. Results are summarized in Tab. 4.4.

The weighted average BR is:

$$BR_{singFit} = (1.08 \pm 0.33) \times 10^{-4} \quad (4.10)$$

Decay Channel	$\epsilon_k$	Signal Yield	$BR(B \rightarrow \tau\nu)$
$\tau \rightarrow e\nu_e\nu_\tau$	$3.40 \times 10^{-4}$	$38 \pm 17$	$(1.09 \pm 0.56) \times 10^{-4}$
$\tau \rightarrow \mu\nu_\mu\nu_\tau$	$2.87 \times 10^{-4}$	$33 \pm 19$	$(1.09 \pm 0.61) \times 10^{-4}$
$\tau \rightarrow \pi\nu_\tau$	$2.59 \times 10^{-4}$	$22 \pm 17$	$(1.09 \pm 0.78) \times 10^{-4}$
$\tau \rightarrow \rho\nu_\tau$	$1.85 \times 10^{-4}$	$16 \pm 15$	$(1.05 \pm 0.74) \times 10^{-4}$

Table 4.4: Summary of efficiencies  $\epsilon_k$  and fit results for each  $\tau$  decay channel considered.

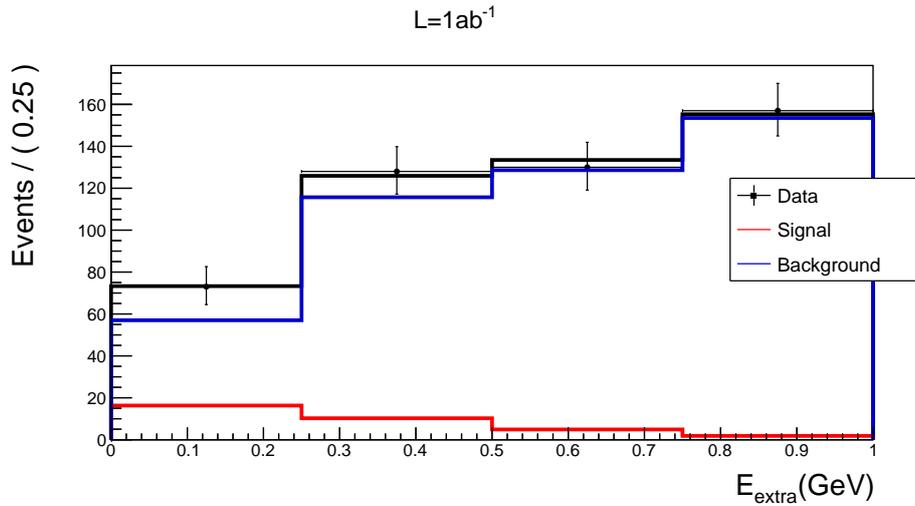


Figure 4.36: 1 of the 100000 extML fits for the  $e^\pm$  channel.

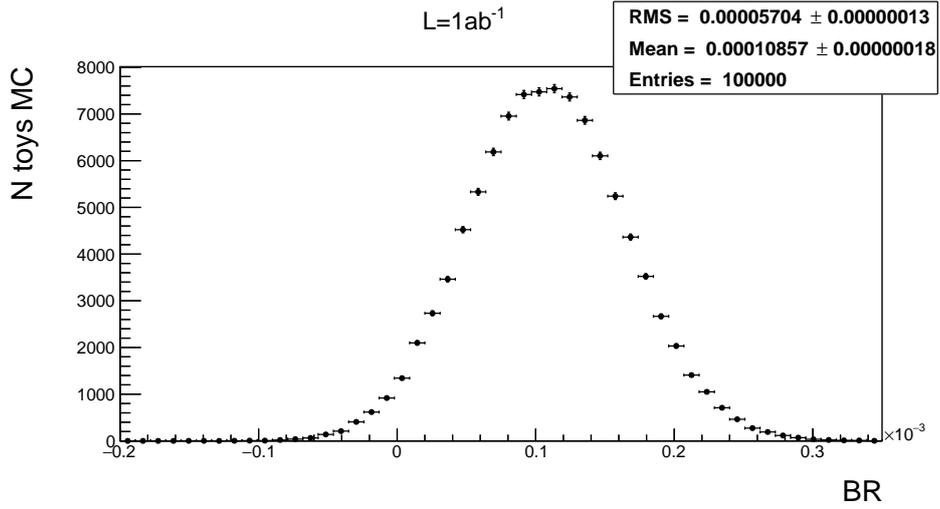


Figure 4.37: BR results of the Toy-MC study for the  $e^\pm$  channel.

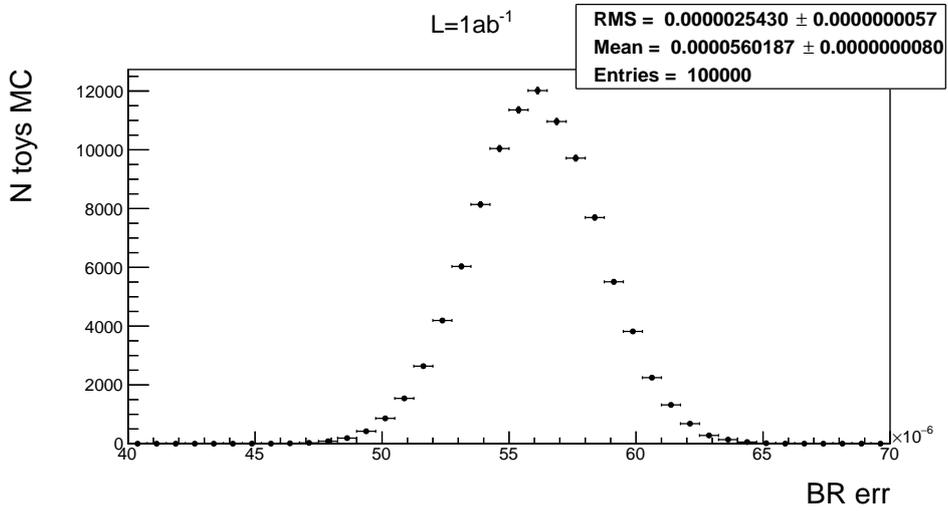


Figure 4.38: BR error results of the Toy-MC study for the  $e^\pm$  channel.

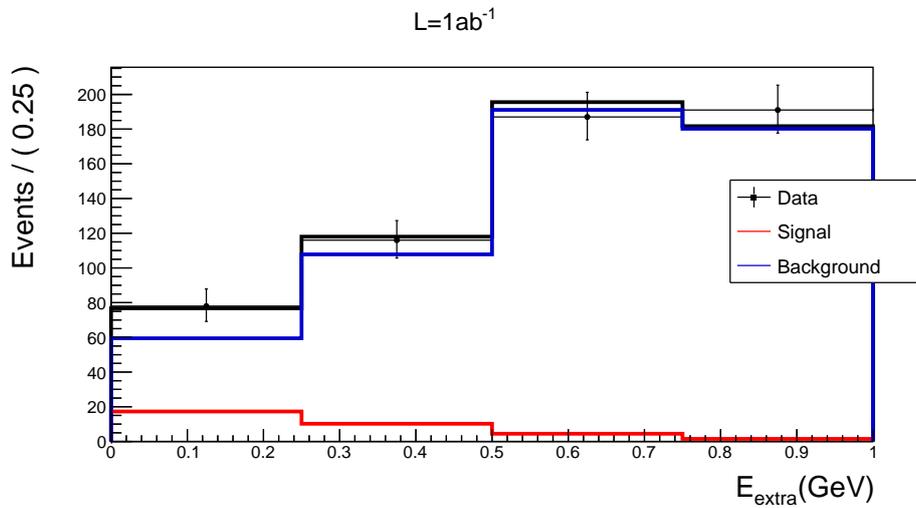


Figure 4.39: 1 of the 100000 extML fits for the  $\mu^\pm$  channel.

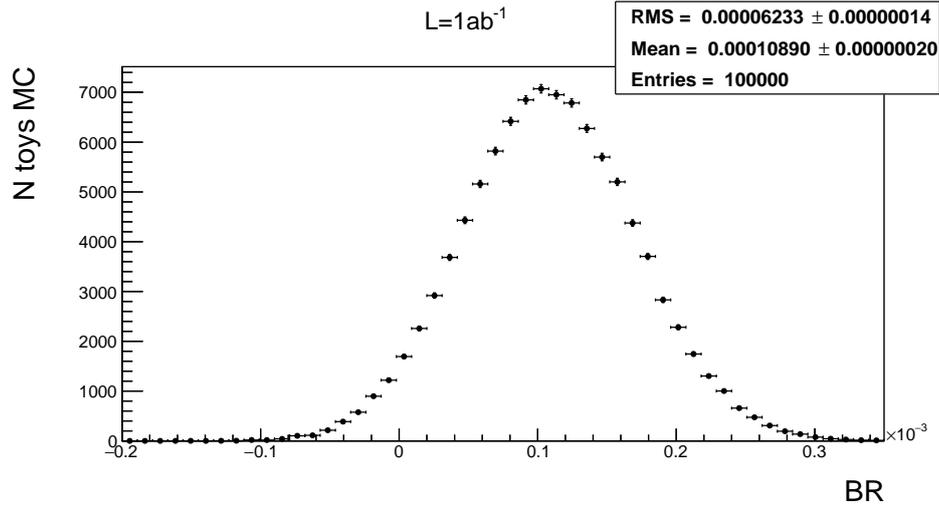


Figure 4.40: BR results of the Toy-MC study for the  $\mu^\pm$  channel.

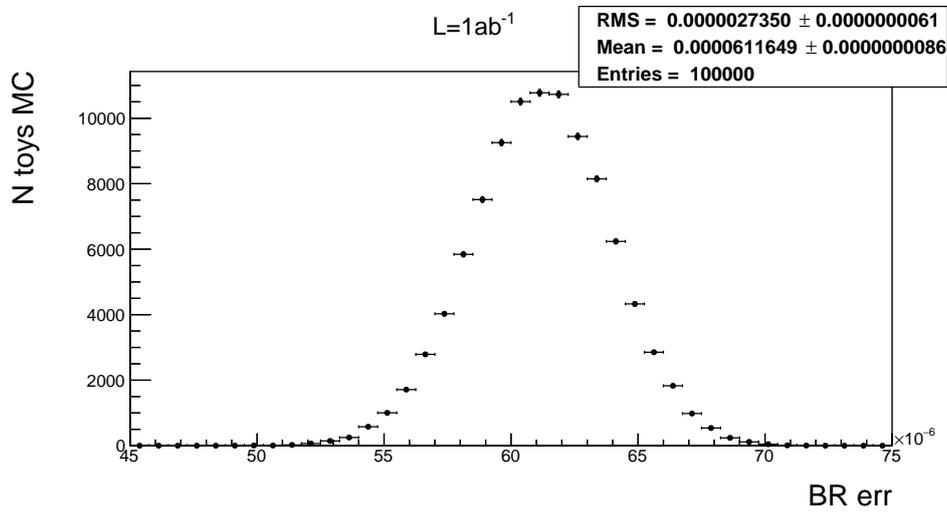


Figure 4.41: BR error results of the Toy-MC study for the  $\mu^\pm$  channel.

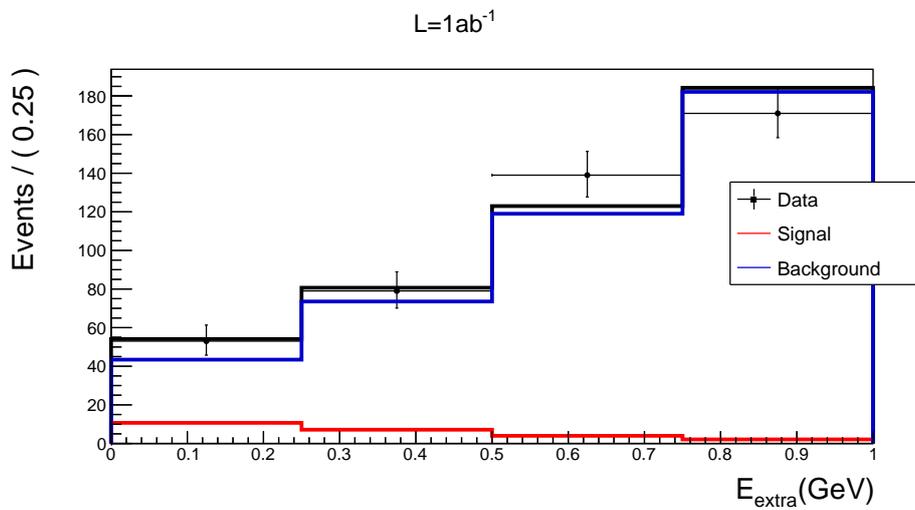


Figure 4.42: 1 of the 100000 extML fits for the  $\pi^\pm$  channel.

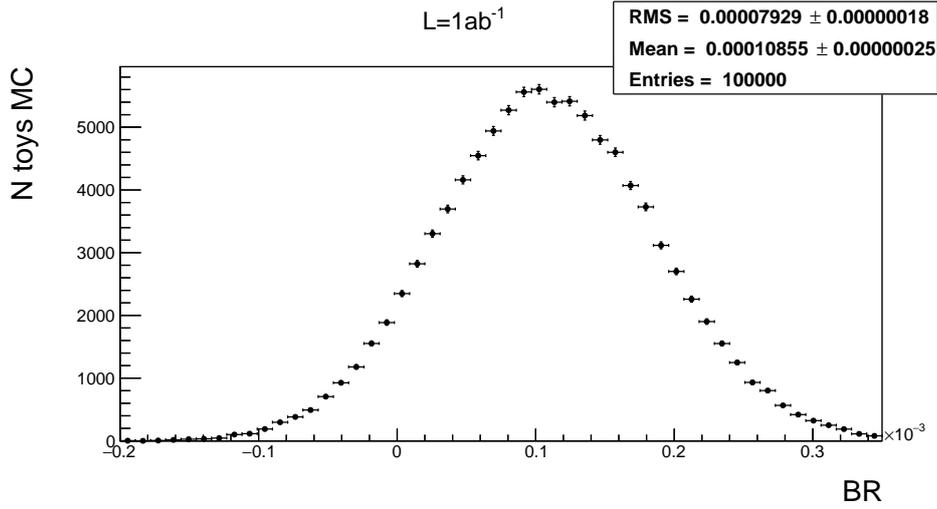


Figure 4.43: BR results of the Toy-MC study for the  $\pi^\pm$  channel.

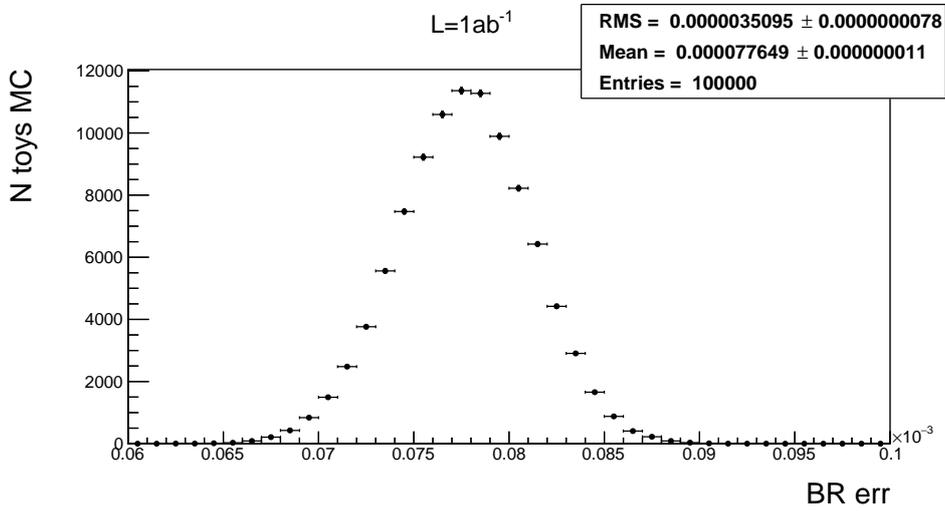


Figure 4.44: BR error results of the Toy-MC study for the  $\pi^\pm$  channel.

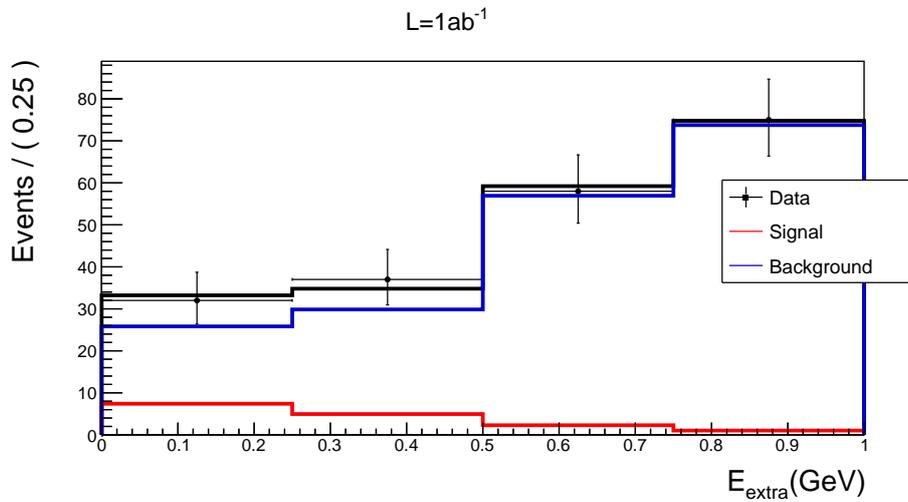


Figure 4.45: 1 of the 100000 extML fits for the  $\rho$  channel.

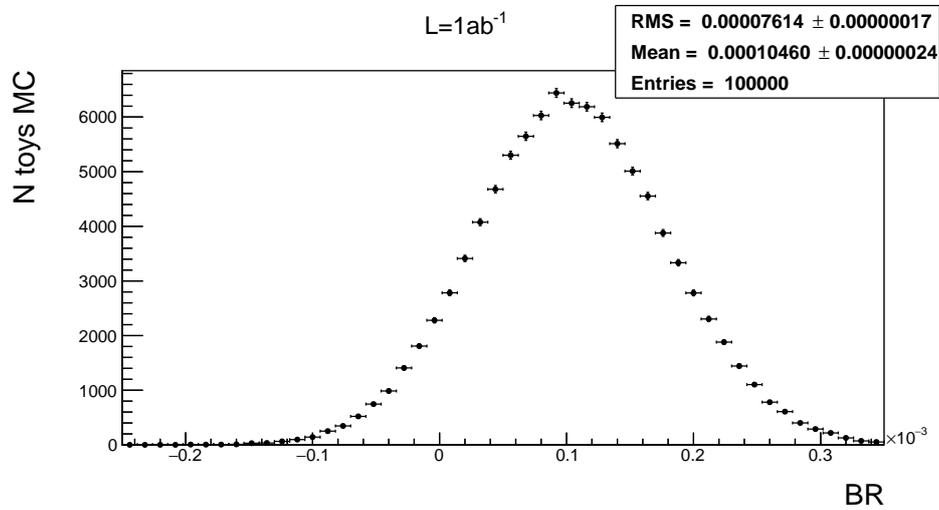


Figure 4.46: BR results of the Toy-MC study for the  $\rho$  channel.

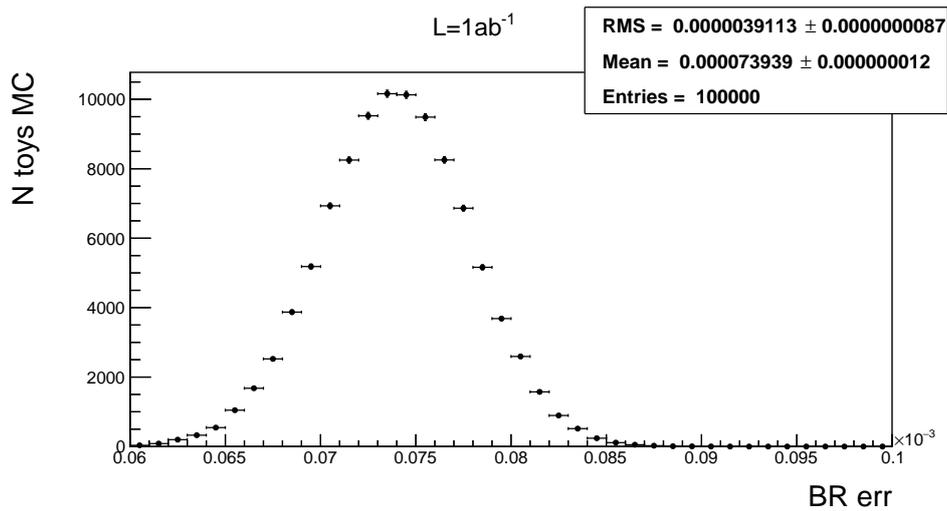


Figure 4.47: BR error results of the Toy-MC study for the  $\rho$  channel.

## 4.6.2 Simultaneous Fits

In this section, we will present the results of the simultaneous fits performed on each  $\tau$  decay channel considered with  $L = 1 \text{ ab}^{-1}$  of data.

The likelihood function has the same expression of Eq. 4.8 and the signal yields are constrained to the common BR via Eq. 4.7:

$$n_{s,k} = 2L_{int}\sigma_{B\bar{B}} \cdot \epsilon_k \cdot BR(B \rightarrow \tau\nu) \quad (4.11)$$

The parameters  $L_{int}$ ,  $\sigma_{B\bar{B}}$  and  $\epsilon_k$  (Tab. 4.4) are fixed in the fit, while BR is allowed to vary.

The BR from the simultaneous fits is:

$$BR_{simFit} = (1.08 \pm 0.33) \times 10^{-4} \quad (4.12)$$

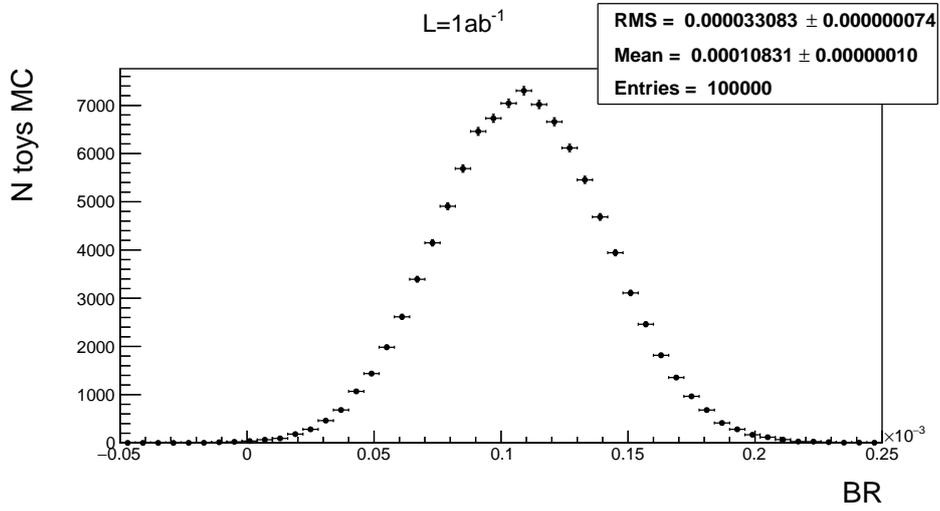


Figure 4.48: BR results of the Toy-MC study for the simultaneous fit.

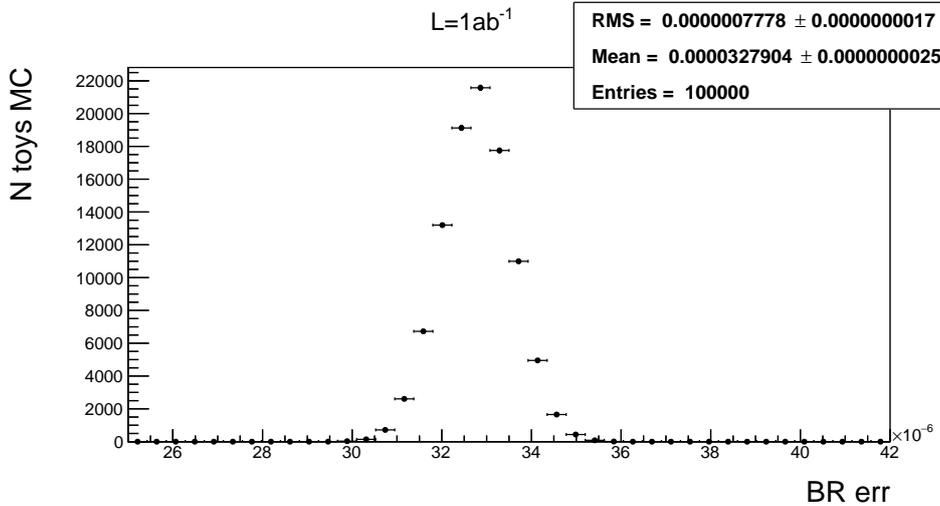


Figure 4.49: BR error results of the Toy-MC study for the simultaneous fit.



# Chapter 5

## Look at Data Collected in 2019

In this chapter we compare the distributions of some variables in the experimental data collected by Belle II in 2019 and in MC simulations, considering together the leptons ( $e$  and  $\mu$ ) and the hadrons ( $\pi$  and  $\rho$ ). For a correct comparison, it is necessary to normalize all the MC distributions to the same integrated luminosity of the experimental data, corresponding to  $L = 8.86 fb^{-1}$ . In particular, the agreement is evaluated in a signal region with a selection looser than the one described before.

One variable not used for the selection of signal events, because it becomes less discriminating by cutting on the missing quantities, but useful for studying the Data-MonteCarlo agreement is the  $|\cos\theta^{TAG,ROE}|$ , where  $\theta^{TAG,ROE}$  is the angle between the *thrust axis* of the  $B_{tag}$  and the *thrust axis* of the Rest Of Event: the *thrust axis*, for a set of momenta, is defined as the direction along which the sum of the momenta projection is maximum, while the Rest Of Event represents everything that is not part of the reconstructed  $B_{tag}$ . For signal and combinatorial background events, we expect a uniform distribution, while for continuum background events a distribution accumulates to 1 (Fig. 5.1).

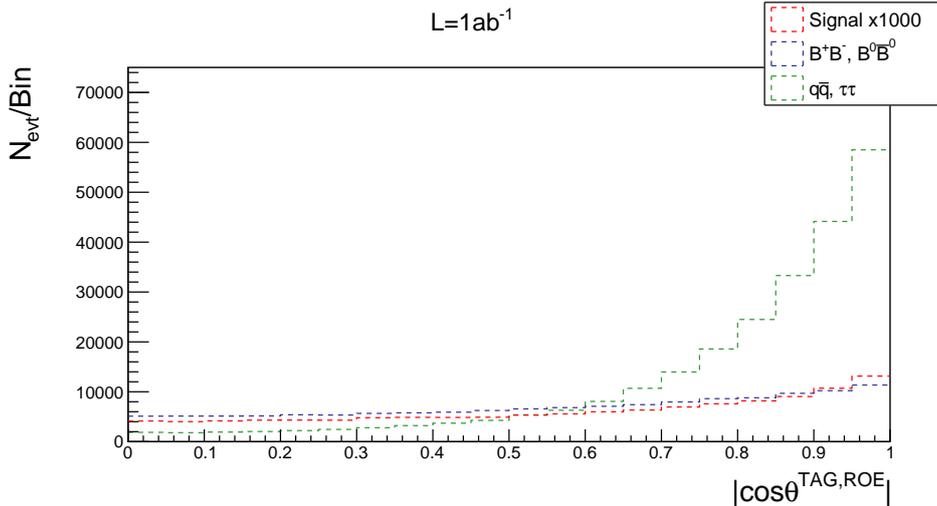


Figure 5.1: Plot of  $|\cos\theta^{TAG,ROE}|$  with *loose* cuts.

## 5.1 Leptons

As seen before,  $e^\pm$  and  $\mu^\pm$  channels share the same kinematic and event properties, so we will study them together, to cope with the small statistics of the data sample.

The loose cuts are:

- $electronID > 0.6$  “OR” ( $muonID > 0.4$  “AND”  $electronID < 0.6$ );
- $M_{bc} > 5.27 \text{ GeV}$ ;
- $sigProb > 0.03$ ;
- $|\cos \theta^{TAG,ROE}| < 0.8$ ;
- $M_{miss}^2 > 6 \text{ GeV}^2$ .

The *tag* cuts are the ones optimized for the MC simulation with an integrated luminosity of  $1 \text{ ab}^{-1}$ ; the cuts on  $|\cos \theta^{TAG,ROE}|$  and  $M_{miss}^2$  are set observing the distributions in Fig. 5.2.

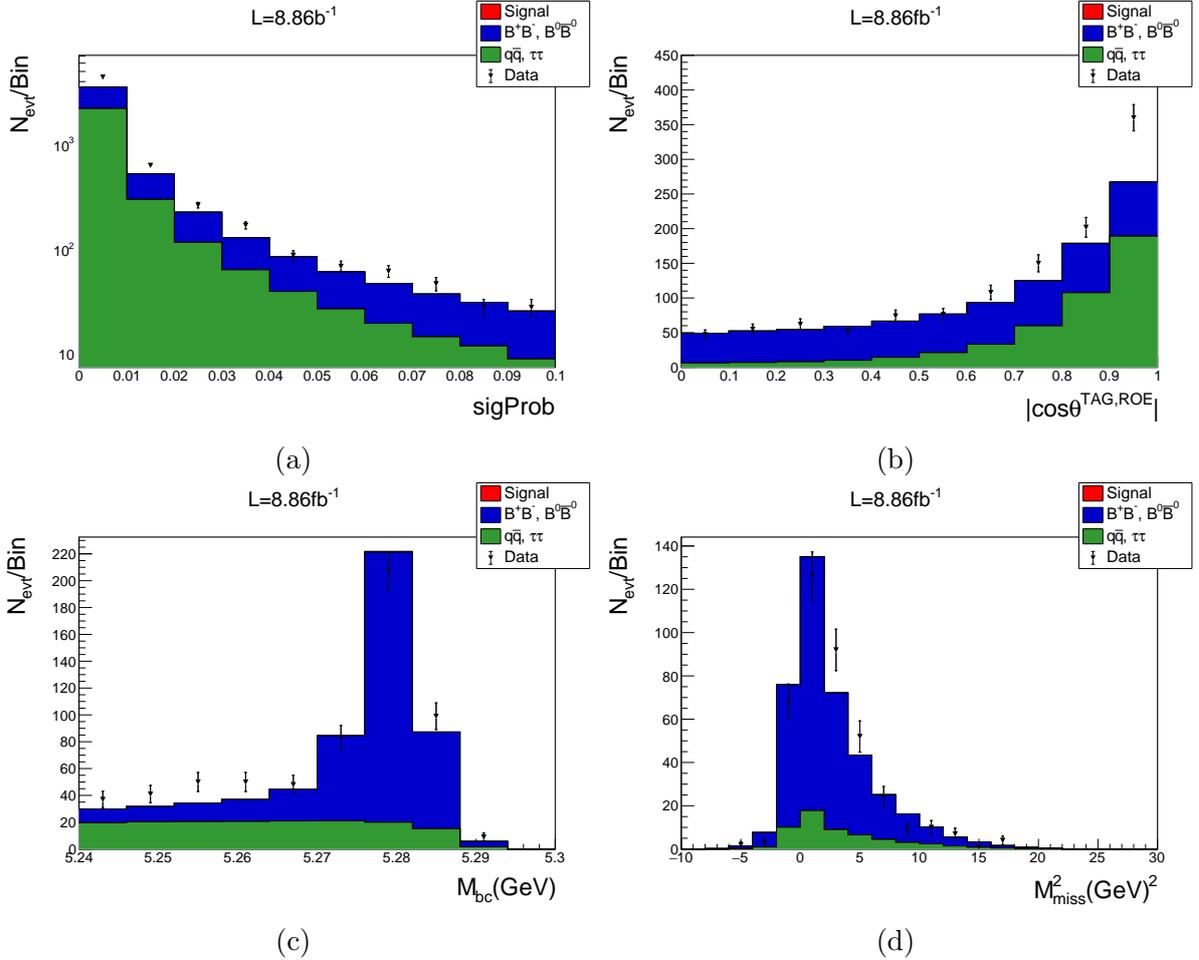


Figure 5.2: Plots with progressive cuts (each variable is cut only on the previous ones) of  $sigProb$  (a),  $|\cos \theta^{TAG,ROE}|$  (b),  $M_{bc}$  (c) and  $M_{miss}^2$  (d) for the leptonic channels.

Two variables not used in the previous sections, but which we now display to study the agreement are  $\theta_{P_{miss}}$  and  $\cos\theta_{P_{miss}}$ , where  $\theta_{miss}$  is the angle between  $P_{miss}$  and the direction of the beams.

The distributions are shown in Fig. 5.3.

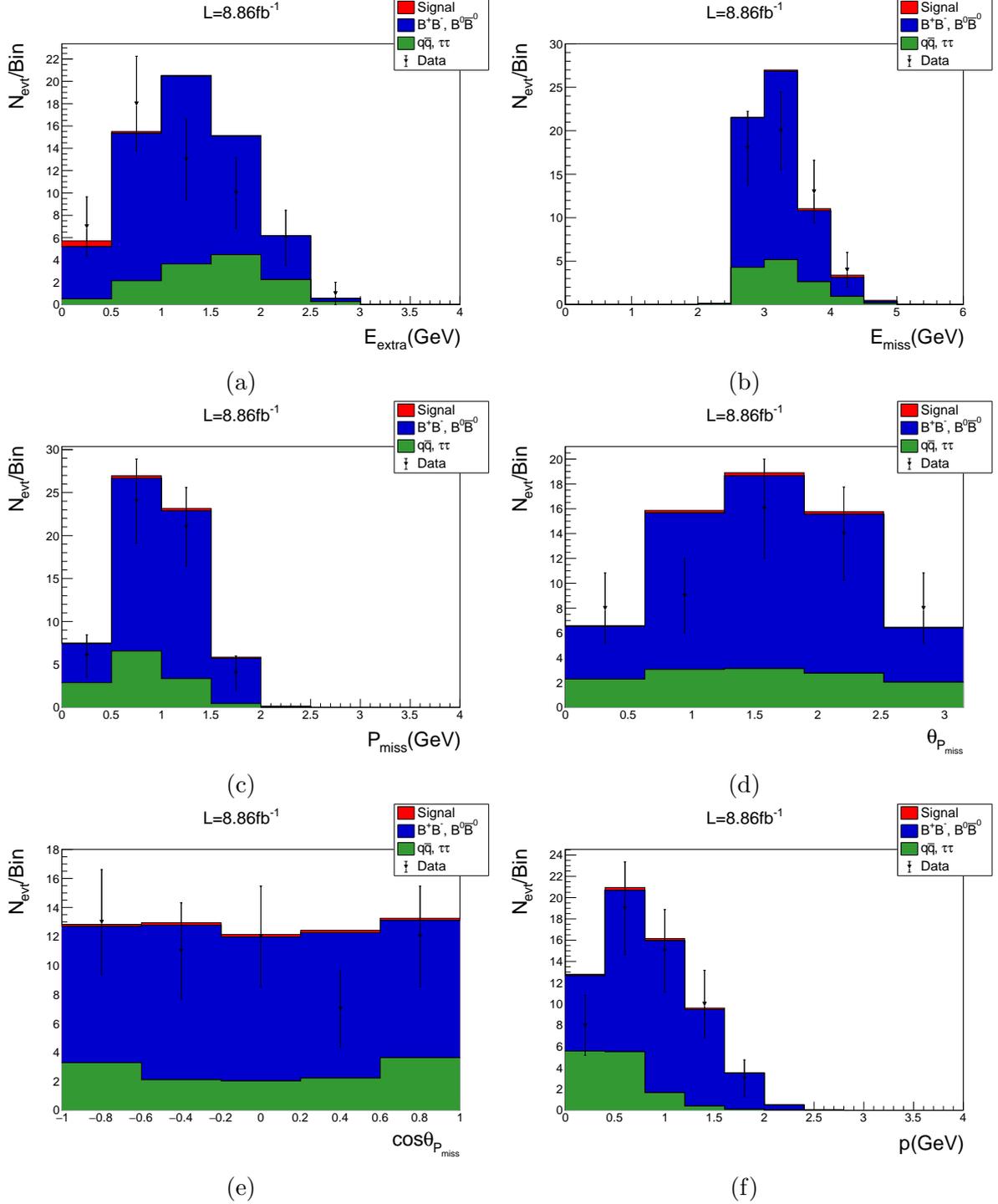


Figure 5.3: Plots with the loose cuts of  $E_{\text{extra}}$  (a),  $E_{\text{miss}}$  (b),  $P_{\text{miss}}$  (c),  $\theta_{P_{\text{miss}}}$  (d),  $\cos\theta_{P_{\text{miss}}}$  (e) and  $p$  (f) for the leptonic channels.

The data and MC comparisons show a good level of agreement in the signal region.

## 5.2 Hadrons

As seen before,  $\pi^\pm$  and  $\rho$  channels share almost the same properties, because of the  $\pi^\pm$  from the  $\rho$ , so we will study them together, to cope with the small statistics of the data sample.

The loose cuts are:

- $pionID > 0.2$  “AND”  $muonID < 0.4$  “AND”  $electronID < 0.6$ ;
- $M_{bc} > 5.27$  GeV;
- $sigProb > 0.04$ ;
- $|\cos\theta^{TAG,ROE}| < 0.8$ ;

The *tag* cuts are the ones optimized for the MC simulation with an integrated luminosity of  $1\text{ ab}^{-1}$ ; the cuts on  $|\cos\theta^{TAG,ROE}|$  is deduced by observing the distributions in Fig. 5.4.

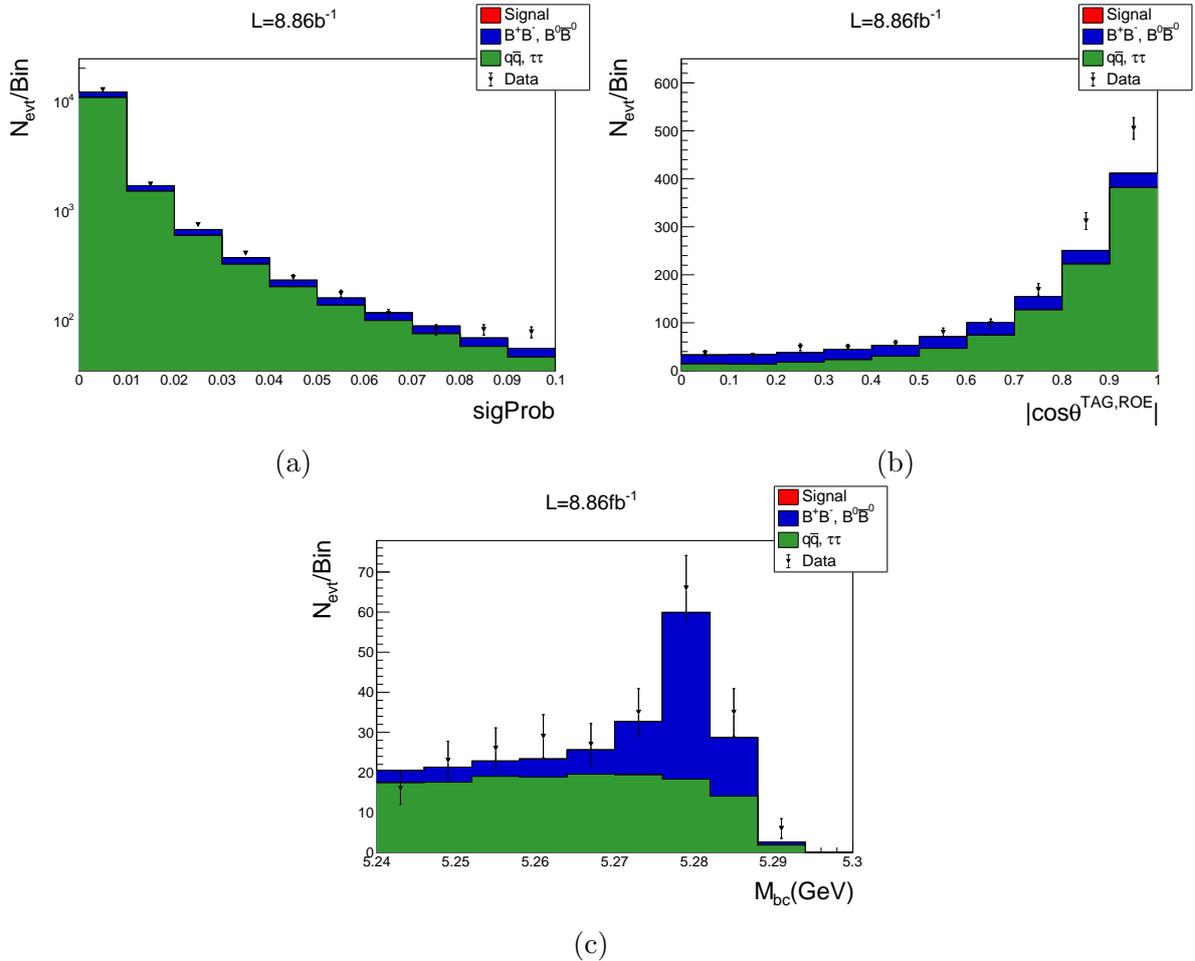
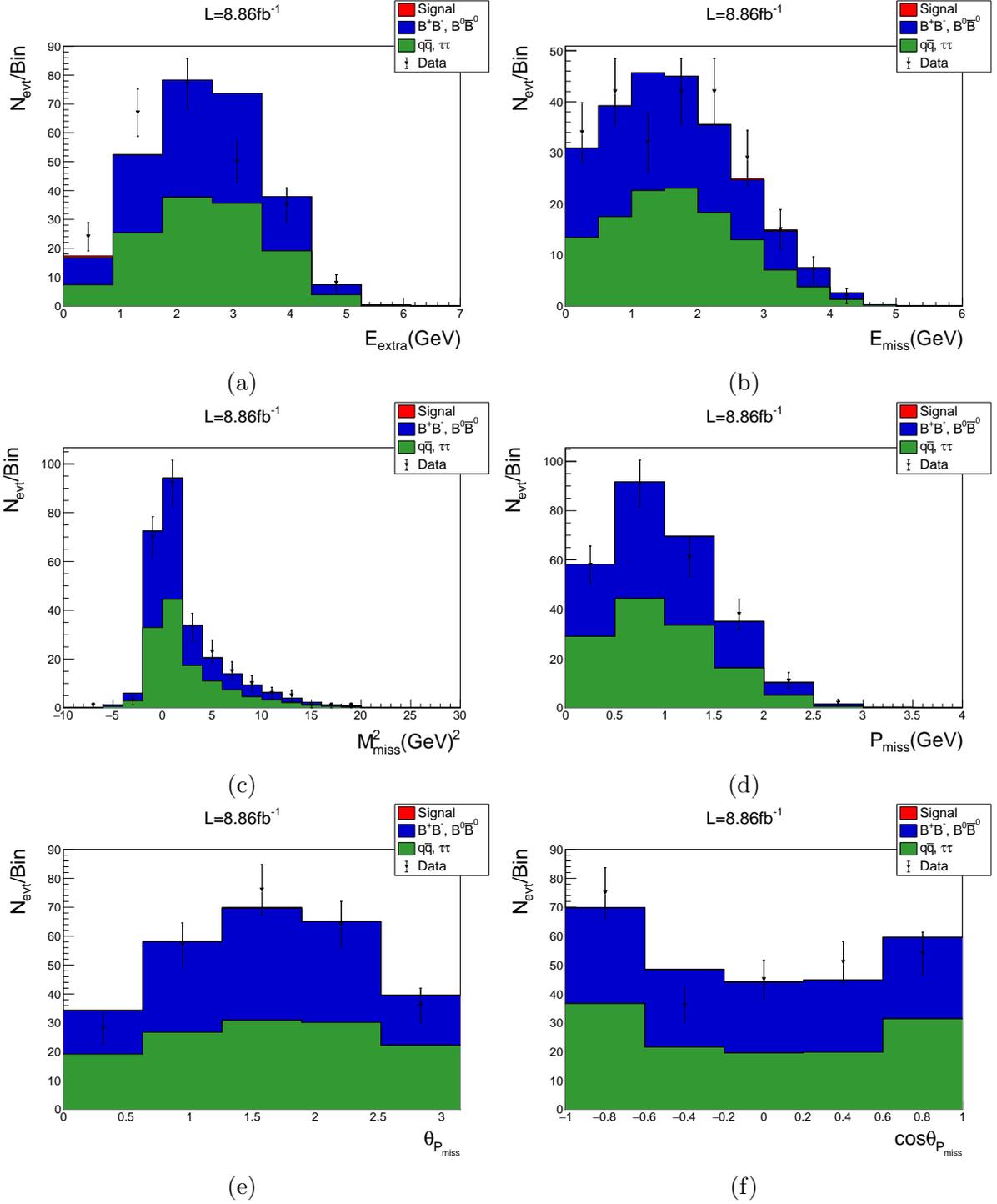


Figure 5.4: Plots with progressive cuts (each variable is cut only on the previous ones) of  $sigProb$  (a),  $|\cos\theta^{TAG,ROE}|$  (b) and  $M_{bc}$  (c) for the hadronic channels.

The observed variables are shown in Fig. 5.5.



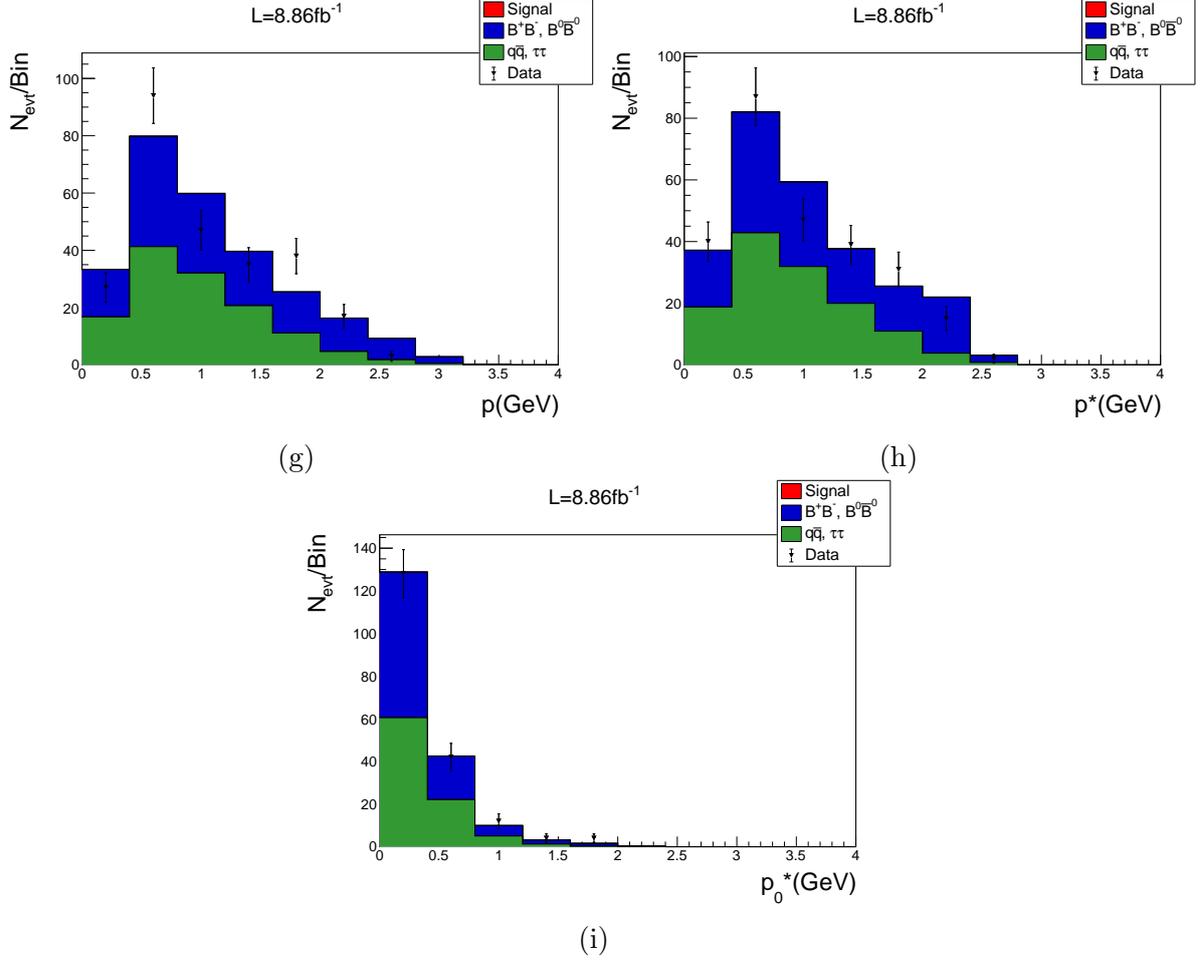


Figure 5.5: Plots with the loose cuts of  $E_{extra}$  (a),  $E_{miss}$  (b),  $M_{miss}^2$  (c),  $P_{miss}$  (d),  $\theta_{P_{miss}}$  (e),  $\cos \theta_{P_{miss}}$  (f),  $p$  (g),  $p^*$  (h) and  $p_0^*$  (i) for the hadronic channels.

The data and MC comparisons show a good level of agreement in the signal region.

### 5.3 Data-MC Discrepancy Considerations

A first contribution is due to the FEI on the tag-side: in fact, it has been observed that at low  $sigProb$  values, i.e. low probability of having a  $B_{tag}$ , the agreement between Data-MC worsens. Therefore, a correction factor  $\epsilon_{FEI}$  can be calculated by means of suitable data control samples.

A second contribution is due to the fact that the MC simulations do not reproduce the continuum background optimally and for this reason, especially  $|\cos \theta^{TAG,ROE}|$  and  $M_{bc}$  show disagreement where the continuum background contribution is greatest.

Another contribution is due to the *particleID*-MC mis-modeling, related to the MC mis-modeling of *particleID* algorithms. In order to correct the MC and to represent faithfully the experimental data, control samples are being identified to estimate a correction factor  $\epsilon_{PID}$  and its related systematic uncertainty.

# Conclusions

In the first part of this thesis project, we have described the sensitivity of the Belle II experiment to the search for  $B \rightarrow \tau\nu$  decays performed on Monte Carlo simulated samples corresponding to an integrated luminosity of  $1.0 \text{ ab}^{-1}$  and its BR estimation. Signal events were searched using the 1-prong decays of the  $\tau$  lepton and, after the selection optimization, an extended maximum likelihood fit for each of the  $\tau$  decay modes and a simultaneous fit are performed on  $E_{extra}$ . Pseudo-datasets are generated using MC PDFs and setting the BR to  $(1.09 \pm 0.24) \times 10^{-4}$ , reported in the last PDG review [19], and two BR measurements with a Toy-MC study are obtained.

The weighted average BR of the  $\tau$  decay channels considered is:

$$BR_{singFit} = (1.08 \pm 0.33) \times 10^{-4}$$

The BR from the simultaneous fits is:

$$BR_{simFit} = (1.08 \pm 0.33) \times 10^{-4}$$

Both the results are in good agreement with each other and with the value of BR used as input for this study (showing the bias is negligible). Future developments may concern the optimization of the signal selection to reduce the statistical uncertainty on the signal yield, refining the continuum rejection, and estimating the peaking background contribution in order to reject it.

In the final part of this thesis project, it was studied the agreement of the MC simulation with data collected in 2019 and corresponding to  $L = 8.86 \text{ fb}^{-1}$ . The data and MC comparisons have shown a good level of agreement in the signal region, even using quite loose cuts. Future developments may concern the estimate of correction factors due to the different behavior of data and MC simulations to further reduce the discrepancy.



# Bibliography

- [1] S.L. Glashow. “Partial Symmetries of Weak Interactions”. In: *Nucl. Phys.* 22 (1961), pp. 579–588.
- [2] Steven Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266.
- [3] Nicola Cabibbo. “Unitary Symmetry and Leptonic Decays”. In: *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531–533.
- [4] Makoto Kobayashi and Toshihide Maskawa. “CP-Violation in the Renormalizable Theory of Weak Interaction”. In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657.
- [5] Ling-Lie Chau and Wai-Yee Keung. “Comments on the Parametrization of the Kobayashi-Maskawa Matrix”. In: *Phys. Rev. Lett.* 53 (19 Nov. 1984), pp. 1802–1805.
- [6] C. Patrignani et al. “Review of Particle Physics”. In: *Chin. Phys. C* 40.10 (2016), p. 100001.
- [7] Lincoln Wolfenstein. “Parametrization of the Kobayashi-Maskawa Matrix”. In: *Phys. Rev. Lett.* 51 (21 Nov. 1983), pp. 1945–1947.
- [8] Andrzej J. Buras, Markus E. Lautenbacher, and Gaby Ostermaier. “Waiting for the top quark mass,  $K^+ \rightarrow \pi + \nu\nu^-$ ,  $B_s^0 - B_s^0$  mixing, and CP asymmetries in B decays”. In: *Physical Review D* 50.5 (Sept. 1994), pp. 3433–3446.
- [9] Schmidtler M. and Schubert K. R. “Experimental constraints on the phase in the Cabibbo-Kobayashi-Maskawa matrix”. In: *Zeitschrift für Physik C Particles and Fields* 53.2 (June 1992), pp. 347–353.
- [10] Cecilia Jarlskog and Raymond Stora. “Unitarity polygons and CP violation areas and phases in the standard electroweak model”. In: *Physics Letters B* 208.2 (1988), pp. 268–274.
- [11] Yuval Grossman, Yosef Nir, and Mihir P Worah. “A Model independent construction of the unitarity triangle”. In: *Physics Letters B* 407.3-4 (1997), pp. 307–313.
- [12] Tetsuo Abe et al. “Belle II Technical Design Report”. In: *arXiv preprint arXiv:1011.0352* (2010).
- [13] Emi Kou et al. “The Belle II physics book”. In: *Progress of Theoretical and Experimental Physics* 2019.12 (2019), p. 123C01.
- [14] M Bona et al. “SuperB: a high-luminosity asymmetric  $e^+ e^-$  super flavor factory. Conceptual Design Report (2007)”. In: *arXiv preprint arXiv:0709.0451* ().
- [15] Yuki Yoshi Ohnishi et al. “Accelerator Design at SuperKEKB”. In: *PTEP* 2013 (2013), 03A011.

- [16] A. Kuzmin. “Endcap calorimeter for SuperBelle based on pure CsI crystals”. In: *Nucl. Instrum. Meth. A* 623 (2010). Ed. by Hiroyuki Iwasaki, Takeshi K. Komatsubara, and Yasuhiro Sugimoto, pp. 252–254. DOI: 10.1016/j.nima.2010.02.212.
- [17] J. P. Lees et al. “Evidence of  $B^+ \rightarrow \tau^+ \nu$  decays with hadronic  $B$  tags”. In: *Phys. Rev. D* 88 (3 Aug. 2013), p. 031102.
- [18] A Abdesselam et al. “Measurement of the branching fraction of  $B^+ \rightarrow \tau^+ \nu_\tau$  decays with the semileptonic tagging method and the full Belle data sample”. In: *arXiv preprint arXiv:1409.5269* (2014).
- [19] M. Tanabashi et al. “Review of Particle Physics”. In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001.
- [20] Wei-Shu Hou. “Enhanced charged Higgs boson effects in  $B^- \rightarrow \tau \bar{\nu}, \mu \bar{\nu}$  and  $b \rightarrow \tau \bar{\nu} + X$ ”. In: *Phys. Rev. D* 48 (5 Sept. 1993), pp. 2342–2344.
- [21] N. Katayama. “Computing models of major HEP experiments: BELLE”. In: *Comput. Phys. Commun.* 110 (1998). Ed. by U. Gensch, L. Hagge, and W. Friebel, pp. 22–25.
- [22] Thomas Keck. “Full Event Interpretation”. In: *Machine Learning at the Belle II Experiment*. Springer, 2018, pp. 63–100.