

# Universitá degli Studi di Padova

Dipartimento di Fisica e Astronomia "Galileo Galilei" Corso di Laurea Magistrale in Fisica

# Measurement of the Standard Model Higgs Boson in the Decay into Two Tau Leptons with the CMS Experiment at the LHC

**Relatore:** Prof. Franco Simonetto Candidato: Alberto Bragagnolo *Matricola:* 1109333

**Correlatore:** Prof. Dott. Elisabetta Gallo

Anno Accademico 2016 - 2017

#### Abstract

This thesis describes the commission of the physics objects and analysis for the  $H \to \tau \tau$ search, covering two of the six inclusive channels, the  $\mu \tau_{\rm h}$ - and the  $e \tau_{\rm h}$ -channel. For this purpose the standard candle process of Z production into  $\tau$  leptons has been studied. The result of this study was the measurement of the Z production cross section in proton-proton collisions.

The main original work presented in this thesis is the development of a Tag and Probe workflow used to measure the identification, isolation and trigger efficiency for electron and muon objects in the  $\tau\tau$  analyses context, needed to improve the agreement between data and simulation.

This analysis was performed with data collected between March and August 2016 with the CMS detector at the LHC, corresponding to a total integrated luminosity of 12.9 fb<sup>-1</sup>. The work presented in this thesis is part of the combined search of the SM and MSSM Higgs boson decay into two  $\tau$  leptons by the CMS analysis group at the Deutsches Elektronen-Synchrotron (DESY) research center in Hamburg.

ii

Dedicated to my grandmother, Maria Rina Furlan (†07.03.2015)

iv

# Contents

1

1	The	Stand	lard Model and the Higgs Boson	<b>5</b>
	1.1	Standa	ard Model and Electroweak Symmetry Breaking	5
		1.1.1	The Higgs Mechanism	7
			1.1.1.1 Bosons Masses	9
			1.1.1.2 Fermion Masses	10
	1.2	Cross S	Section Calculation at Hadron Colliders	12
	1.3	Higgs 1	Production and Decay at the LHC	14
		1.3.1	Higgs Decays	16
<b>2</b>	The	CMS	Experiment at the LHC	17
	2.1	The La	arge Hadron Collider	17
	2.2	The Co	ompact Muon Solenoid	22
		2.2.1	Overall design	22
	2.3	The Th	racking System	23
		2.3.1	The Pixel Tracker	24
		2.3.2	The Strip Tracker	24
	2.4	The E	lectromagnetic Calorimeter	25
	2.5	The H	adronic Calorimeter	28
	2.6	The St	uperconducting Solenoid	30
2.7 The Muon System		luon System	31	
		2.7.1	Drift Tube Chambers	32
		2.7.2	Cathode Strip Chambers	33
		2.7.3	Resistive Plate Chambers	34
	2.8	Data A	Acquisition and Trigger System	34
		2.8.1	Level 1 Trigger	34
		2.8.2	High Level Trigger	35
		2.0.2	Ingli Devel Inggel	50

3	Phy	vsics O	bject Reconstruction	<b>37</b>			
	3.1	Prima	ry Vertex Reconstruction	. 37			
	3.2	Partic	le Flow Reconstruction	. 38			
	3.3	Recon	struction and Identification of Muon	. 40			
		3.3.1	Muon Reconstruction	. 40			
		3.3.2	Muon Identification	. 41			
	3.4	Recon	struction and Identification of Electron	. 41			
		3.4.1	Electron Reconstruction	. 41			
		3.4.2	Electron Identification	. 43			
	3.5	5 Lepton Isolation					
	3.6	3.6 Reconstruction And Identification of Hadronic Tau					
		3.6.1	$\pi^0$ Reconstruction with Strip Object	. 45			
		3.6.2	Tau Reconstruction	. 46			
		3.6.3	Tau Isolation Discriminators	. 47			
		3.6.4	Tau Discriminator Against Electron	. 48			
		3.6.5	Tau Discriminator Against Muon	. 48			
	3.7	Missin	g Transverse Energy	. 49			
4	Eve	nt Sele	ection	51			
	4.1	Data S	Samples and Triggers	. 51			
	4.2	Simula	ated Samples and Background Modeling	. 52			
		4.2.1	Simulation in High Energy Physics	. 52			
		4.2.2	Monte Carlo Samples	. 53			
		4.2.3	Estimation of QCD Background: the ABCD method	. 55			
	4.3	Baseli	ne Event Selection	. 56			
		4.3.1	Muon and Electron Cuts	. 58			
		4.3.2	Tau Cuts	. 58			
		4.3.3	Lepton-Tau Pairs Cuts	. 58			
		4.3.4	Event Cuts	. 61			
			4.3.4.1 Sequence of the selection criteria	. 62			
	4.4	Correc	ztions	. 62			
		4.4.1	Pile-Up Correction	. 64			
		4.4.2	Identification and Isolation Efficiencies	. 65			
		4.4.3	Trigger Efficiency	. 65			
	4.5	Cut or	n the Transverse Mass	. 65			
5	Lep	ton ID	) and Trigger Efficiencies Measurement	69			
	5.1	1 Tag and Probe Method 60					
	0	5.1.1	Selection of Tag&Probe Pairs	. 70			
		5.1.2	Extraction of Efficiencies	.72			
			5.1.2.1 Fitting Model	. 72			
	5.2	Result	is	. 75			
		5.2.1	Identification and Isolation Scale Factor	. 75			
		5.2.2	Trigger Efficiency	. 75			

# CONTENTS

		5.2.3 Validation of the Efficiency Corrections	75
	5.3	Evaluation of the Systematic Error	82
6	Z-B	oson Production Cross Section Measurement	83
	6.1	Introduction	83
	6.2	Systematic Uncertainties	83
	6.3	Cross Section Measurement	85
Su	imma	ary and Outlook	87
Ac	cknow	wledgments	89
A	Dist	tributions After the Transverse Mass Cut	91
в	Elec	ctron Tag&Probe Invariant Mass Distributions	93
Bi	bliog	graphy	95

vii

#### Introduction

The discovery of a new boson at the Large Hadron Collider (LHC) announced by the ATLAS and CMS collaborations on 4th of July 2012 [1, 2] was a milestone in the half century long quest for understanding the origin of electroweak symmetry breaking and the mass of elementary particles. Subsequent measurements of both collaborations have indeed independently confirmed, that no significant deviations from the properties of the prognosticated Higgs boson, the last missing piece of the Standard Model, can be observed. Direct and indirect measurements of couplings of the new boson to Standard Model particles revealed compatibility, within uncertainties, with the predicted couplings of the Higgs boson.

The Standard Model (SM) is currently our best description of the particle physics phenomena. In the context of the SM, the electroweak symmetry breaking (EWSB), a mechanism which generates the masses of the vector bosons Z and  $W^{\pm}$ , is achieved by introducing a complex scalar doublet with non-zero vacuum state to spontaneously break the local gauge symmetry, leading to the prediction of the Higgs boson. The Higgs mechanism was developed by three independent groups in 1964: by Peter Higgs [3, 4]; by Robert Brout and François Englert [5]; and by Gerald Guralnik, Carl Hagen, and Tom Kibble [6]. It was incorporated into the framework of the  $SU(2)_L \times U(1)_Y$  Standard Model electroweak gauge theory by Steve Weinberg [7] and Abdus Salam [8] and the renormalizability of spontaneously-broken gauge theories was subsequently proven by Geradus 't Hooft and Martin Veltmann [9, 10]. Fermions are predicted to acquire masses by so-called Yukawa couplings to the broken scalar field, preserving chiral and gauge symmetry. This aspect of the theory is tested directly in the search of the Higgs boson decaying to fermion pairs [7].

As stated above, the theory predicts the existence of the Higgs boson, which is the only physical degree of freedom of the associated Higgs field introduced by the Higgs mechanism. The Higgs mass was measured by the ATLAS and CMS collaborations in the  $H \rightarrow \gamma \gamma$  and 4-lepton  $H \rightarrow ZZ$  channels. The combined measurement of the Higgs mass is  $m_H = 125.09 \pm 0.21(\text{stat}) \pm 0.11(\text{syst}) \text{ GeV/c}^2$  [11].

Both experiments found direct evidence for fermionic couplings of the new bosons,

and thereby establish strong evidence for the existence of Yukawa couplings [12, 13]. Via gluon fusion Higgs production processes, indirect evidence for couplings to up-type quarks has already been established. The  $H \to \tau \tau$  decay channel has the second largest branching fraction, after the decay channel to a pair of b quarks which however suffers from large QCD background. Therefore the  $H \to \tau \tau$  decay constitutes the most sensitive channel to study fermionic Higgs couplings.

However the SM is not perfect. There are many arguments that suggest that new physics is necessary to describe the universe as we observe it. Evidence for dark matter, the neutrino oscillation and gravity are not predicted by the SM. There are many theories attempting to extend the SM to new territories, one of the most popular ideas is Supersymmetry. The minimal supersymmetric scenario is the Minimal Supersymmetric Standard Model (MSSM) [14]. In the context of the MSSM, the Higgs decay into  $\tau$  pairs may be enhanced. Therefore the  $\tau\tau$  final state is sensitive to new physics.

The current thesis describes the commission of the physics objects and analysis for the  $H \rightarrow \tau \tau$  search, covering two of the six inclusive channels, the  $\mu \tau_{\rm h}$ - and  $e \tau_{\rm h}$ -channel. For this purpose the standard candle process of Z production into  $\tau$  leptons has been studied. The work presented in this thesis is part of the combined search of the SM and MSSM Higgs boson decay into two  $\tau$  leptons by the CMS analysis group at the Deutsches Elektronen-Synchrotron (DESY) research center in Hamburg. This analysis was performed with data collected between March and August 2016 with the CMS detector at the LHC, corresponding to a total integrated luminosity of 12.9 fb<sup>-1</sup>.

Within the CMS collaboration, one of the task of the  $H\tau\tau$  group at DESY is the measurement of the efficiency correction, needed to improve the agreement between data and simulation, for muon and electron objects in the  $\tau\tau$  analyses context. The main original work presented in this thesis is the development of a Tag and Probe workflow to measure the abovementioned efficiencies.

This thesis is organized in six chapters:

- Chapter 1 describes the Standard Model of particle physics, with emphasis to the EWSB and Higgs mechanism, and the Higgs production at the LHC, to state the hypothesis to be tested.
- Chapter 2 describes the LHC proton-proton collider and the CMS detector. General information is given on LHC and the CMS experiment, and CMS sub-detector components are described in slightly more detail, to show the general experimental circumstances, under which the analysis is performed.
- Chapter 3 is a description of the event and physics object reconstruction techniques relevant for this analysis. Particular attention is given to the reconstruction and identification of muons, electrons and hadronic tau leptons.
- Chapter 4 presents in detail the baseline selection for the  $Z/H \to \tau \tau \to \ell \tau_{had}$  $(\ell = e, \mu)$ . This chapter also describes the simulated samples used in the analysis,

the background estimation and the corrections applied to improve the agreement between data and simulation. An additional study on a transverse mass cut, in order to further reduce the W+Jets background, is also presented in this chapter.

- Chapter 5 presents the Tag and Probe workflow developed to measure the efficiency corrections. The measurement of the identification, isolation and trigger efficiencies for muon and electron, here described, is a major part of my thesis work and my personal contribution to the whole analysis group.
- Chapter 6 presents the final measurement of the Z boson production cross section in the  $\mu \tau_{had}$ - and  $e \tau_{had}$  decay channels. The systematic uncertainties that affect the abovementioned measurement are also discussed in this chapter.

# INTRODUCTION

4

# CHAPTER 1

# The Standard Model and the Higgs Boson

In particle physics the Universe is described as composed of particle. The elementary units interact with each other and give rise to more complex structure. By our current understanding, these elementary particles are physical manifestation of the quantum fields that also govern their interactions. The four fundamental forces known in nature are the electromagnetic force, the weak force, the strong force and gravity. The Standard Model (SM) [7, 8, 15–17] of particle physics provides the modern understanding of the behavior of the first three forces. The SM has been tested extensively and has so far withstood all tests exceptionally well. The latest triumph of the standard model was the discovery of a new boson with mass close to 125 GeV in July 2012, compatible with the predicted Higgs boson.

#### 1.1 Standard Model and Electroweak Symmetry Breaking

In the Standard Model, matter and its interactions are described by twelve elementary spin 1/2 particles, fermions, their antifermions and six gauge bosons. The particles predicted by the Standard Model are represented graphically in figure 1.1.

The SM is based on global Poincaré space-time symmetry and local gauge invariance under transformations of the direct product of groups  $SU(3)_C \times SU(2)_L \times U(1)_Y$ . Gauge bosons are structured in adjoint representations of the corresponding gauge groups:

- Strong force: the octet of gluons  $G^a_{\mu}$  for the  $SU(3)_C$  (*C* is color) gauge group establishes the interactions of Quantum Chromodynamics (QCD) by acting on  $SU(3)_C$  triplets of quarks.
- Electroweak force: the gauge bosons are represented by the  $SU(2)_L$  triplet  $W^i_\mu$  together with the singlet  $B_\mu$  of  $U(1)_Y$ .



Figure 1.1: The elementary particles of the Standard Model. From [18].

		Field		Spin
Oversla	$\begin{pmatrix} u \\ d \end{pmatrix}_I$	$\begin{pmatrix} c \\ s \end{pmatrix}_I$	$\begin{pmatrix} t \\ b \end{pmatrix}_{I}$	1/2
Quarks	$u_R$ $d_R$	$c_R$ $s_R$	$t_R$ $b_R$	$\frac{1/2}{1/2}$
Leptons	$\binom{\nu_e}{e}_L$	$\binom{\nu_{\mu}}{\mu}_{L}$	$ \begin{pmatrix} \nu_{\tau} \\ \tau \end{pmatrix}_{L} $	1/2
	$e_R$	$\mu_R$	$ au_R$	1/2
Higgs-doublet		$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}_L$		0
		$G^a_\mu$		1
Gauge bosons		$W^{i}_{\mu}$		1
		$B'_{\mu}$		1

Table 1.1: The field content of the Standard Model.

Matter particles, consisting of leptons and quarks, appear as right- and left-handed spinors living in irreducible representations of the Poincaré group. Left(right)-handed elementary fermion fields are doublets(singlet) under  $SU(2)_L$  and are charged under  $U(1)_Y$ . All quarks and leptons appear in three generations. The left-handed electron  $(e_L)$ , muon  $(\mu_L)$  and tau  $(\tau_L)$  together with the associated neutrinos  $\nu_L$ , form  $SU(2)_L$ doublets, whereas right-handed neutrinos are not observed in nature, leading to the right-handed  $SU(2)_L$  singlets  $e_R$ ,  $\mu_R$ ,  $\tau_R$ . Accordingly, left-handed up- and down-type quarks, u (up) and d (down), s (strange) and c (charm) as well as t (top) and b (bottom) form  $SU(2)_L$  doublets, and their right-handed counterparts form  $SU(2)_L$  singlets. The field content of the Standard Model is reported in table 1.1.

The lagrangian of the electroweak sector of the SM is

$$\mathcal{L}_{SM}^{\rm EW} = \mathcal{L}_{\rm gauge} + \mathcal{L}_{\rm fermion}, \tag{1.1}$$

7

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} F^{i}_{\mu\nu} F^{i\,\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \qquad (1.2)$$

$$\mathcal{L}_{\text{fermion}} = \overline{\psi^i}_{L,R} i \gamma^\mu D_\mu \psi^i_{L,R} \,. \tag{1.3}$$

where

- $F^i_{\mu\nu}$  and  $B_{\mu\nu}$  are the gauge field tensors of  $SU(2)_L$  and  $U(1)_Y$ , respectively<sup>1</sup>,
- $D_{\mu}$  is the covariant derivative for the  $SU(2)_L \times U(1)_Y$  SM gauge theory and can be written as

$$D_{\mu} = \partial_{\mu} - ig\mathbf{T}_{i}W^{i}_{\mu} - ig'\frac{Y}{2}B_{\mu},$$

with g and g' being the coupling constants of the  $SU(2)_L$  and  $U(1)_Y$  interaction respectively.  $\mathbf{T}^i$  and Y are the gauge group generators, a suitable representation can be written as  $(Y; \mathbf{T}_i) = (\mathbf{1}_2; \tau_i/2)$  with  $\mathbf{1}_2$  the bidimensional identity matrix and  $\tau_i = \sigma_i$  the Pauli matrices.

•  $\psi_{L,R}^i$  are the left- and right-handed fermion fields with *i* the fermion index<sup>2</sup>.

The concept of gauge symmetry in general forbids mass terms for all fields in the form of  $m_f \bar{\psi} \psi$ ,  $m_B^2 B_\mu B^\mu$  or  $m_W^2 W_\mu W^\mu$ , as these would directly spoil the gauge symmetry, which is in contradiction to the observed physically massive  $W^{\pm}$  and Z bosons mediating the weak force and the massive leptons and quarks.

#### 1.1.1 The Higgs Mechanism

In the Standard Model the masses of gauge bosons are generated by spontaneously breaking the corresponding gauge symmetries. The term spontaneous means that the symmetry is not broken explicitly by the interactions but rather by the asymmetry of the state of lowest energy, referred to as the vacuum quantum field theory. According to

<sup>&</sup>lt;sup>1</sup>Defined as  $F_{\mu\nu}^{i} = \partial_{\mu}W_{\nu}^{i} - \partial_{\nu}W_{\mu}^{i} + g\epsilon^{ijk}W_{\mu}^{j}W_{\nu}^{k}$  (i = 1, 2, 3) and  $B_{\mu\nu} = \partial_{\mu}B_{\nu} - \partial_{\nu}B_{\mu}$ . <sup>2</sup> $i = e, \mu, \tau, \nu_{e}, \nu_{\mu}, \nu_{\tau}, u, c, t, d, s, b$ .

the Goldstone theorem the spontaneous symmetry breaking creates massless bosons, so called Goldstone bosons. If the broken symmetry also corresponds to a gauge symmetry, then the associated Goldstone boson and the massless gauge boson combine to form a massive gauge boson. This is the famous Higgs (or BEH) mechanism, elaborated in 1964 independently by Peter Higgs [3, 4]; by Robert Brout and François Englert [5]; and by Gerald Guralnik, Carl Hagen, and Tom Kibble [6].

The simplest way to implement the Higgs mechanism is to add a weakly coupled spin-0 particle to the theory with a potential that is minimized at a non-zero field value (figure 1.2). The field of such a particle spontaneously breaks the electroweak  $SU(2)_L \times$  $U(1)_Y$  symmetry. The masses of the fermions are generated by Yukawa couplings to the introduced scalar field with degenerate vacuum state. The simplest choice for the new scalar field is a single complex scalar doublet called the Higgs field:

$$\phi = \begin{pmatrix} \phi^+\\ \phi^0 \end{pmatrix}. \tag{1.4}$$

The Higgs Lagrangian can be written as:

$$\mathcal{L}_{\text{Higgs}} = (D_{\mu}\phi)^{\dagger} (D^{\mu}\phi) - V(\phi^{\dagger}\phi) + \mathcal{L}_{\text{Yukawa}}, \qquad (1.5)$$

$$V\left(\phi^{\dagger}\phi\right) = +\mu^{2}\phi^{\dagger}\phi + \lambda(\phi^{\dagger}\phi)^{2}, \qquad (1.6)$$

Unitarity requires that the constants  $\lambda$  and  $\mu^2$  be real and stability demands that  $\lambda$  be positive. The covariant derivative must be

$$D_{\mu} = \left(\partial_{\mu} - \frac{i}{2}g\tau_i W^i_{\mu} - \frac{i}{2}g'B_{\mu}\right)\phi.$$

The electroweak SM Lagrangian is then:

$$\mathcal{L}_{SM}^{\rm EW} = \mathcal{L}_{\rm gauge} + \mathcal{L}_{\rm fermion} + \mathcal{L}_{\rm Higgs}, \qquad (1.7)$$

The parameter  $\mu^2$  in the Higgs potential  $V(\phi)$  is chosen negative,  $\mu^2 < 0$ , in order to realize the spontaneous breaking of the symmetry  $SU(2)_L \times U(1)_Y \to U(1)_Q$ , where  $U(1)_Q$  is the gauge symmetry group of electromagnetic interactions. This inevitably leads to a degenerate non-zero vacuum state of the Higgs field minimizing the potential:

$$\phi^{\dagger}\phi = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2},\tag{1.8}$$

with  $v \equiv \sqrt{-\frac{\mu^2}{\lambda}}$ . In order to have an electrically neutral vacuum the vacuum expectation value (VEV) of the Higgs field is due to  $\phi^0$ :

$$\langle \phi \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0\\ v \end{pmatrix}. \tag{1.9}$$

9



Figure 1.2: Illustration of the Higgs potential. The potential is symmetric with a local maximum in the centre. The energy state in the centre is symmetric, however it is an excited state and therefore unstable. Because of the form of the potential, the stable final state at the minimum, the vacuum state, is not symmetric anymore.

Expanding around the minimum of its potential and applying a convenient gauge transformation (called *unitary gauge*), the scalar doublet can be written in the form

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0\\ v + H(x) \end{pmatrix} \tag{1.10}$$

where H(x) is a real scalar field, called the standard model Higgs boson field, that represent "radial" excitations of the Higgs VEV.

#### 1.1.1.1 Bosons Masses

Expressing the Higgs lagrangian 1.6 in the unitary gauge and omitting the kinetic and interaction terms<sup>3</sup>, we obtain the following mass lagrangian:

$$\mathcal{L}_m = -\lambda v^2 H^2 + \frac{v^2}{8} \{ g^2 [(W^1_\mu)^2 + (W^2_\mu)^2] + (gW^3_\mu - g'B_\mu)^2 \}.$$
(1.11)

<sup>&</sup>lt;sup>3</sup>Interaction terms are terms with a number of fields  $\geq 3$ .

The bosonic fields can be rotated, without loss of generality, into the physical states:

$$W^{\pm}_{\mu} = \frac{1}{\sqrt{2}} (W^1_{\mu} \mp i W^2_{\mu}) \quad \text{(charged currents)}, \tag{1.12}$$

$$Z_{\mu} = \cos \theta_W W_{\mu}^3 - \sin \theta_W B_{\mu} \quad \text{(neutral current)}, \tag{1.13}$$

$$A_{\mu} = \sin \theta_W W_{\mu}^3 + \cos \theta_W B_{\mu} \quad \text{(photon)}, \tag{1.14}$$

with  $\tan \theta_W = \frac{g'}{g}$ . The mass lagrangian is then given by

$$\mathcal{L}_m = -\lambda v^2 H^2 + m_w^2 W_\mu^+ W^{-\mu} + \frac{1}{2} m_Z^2 Z_\mu Z^\mu$$
(1.15)

The first term of eqn 1.15 is the mass term for the Higgs boson, from which the mass of the Higgs boson is given by

$$m_H = \sqrt{2\lambda v^2} = \sqrt{-2\mu^2}.$$
 (1.16)

Its value is not connected to other quantities already measured, hence the SM does not give a prediction for the value of the Higgs mass, which must be determined experimentally.

The second and third terms of eqn 1.15 are of fundamental importance, because they are mass terms for the W and Z gauge bosons. From them, we see that the masses of the W and Z gauge bosons in the SM are given by

$$m_W = \frac{gv}{2}$$
  $m_Z = \frac{v}{2}\sqrt{g^2 + g'^2} = \frac{gv}{2\cos\theta_W}.$  (1.17)

The photon field  $A_{\mu}$  remains massless after the rotation 1.14.

#### 1.1.1.2 Fermion Masses

Masses of fermions, and especially  $\tau$ -leptons, are generated via Yukawa couplings. The Yukawa Lagrangian can be written as:

$$\mathcal{L}_{\text{Yukawa}} = -\sum_{i,j=1}^{3} [y_{ij}^{\ell} \overline{\ell}_{L}^{i} \phi \ell_{R}^{j} + y_{ij}^{u} \overline{q}_{L}^{i} \tilde{\phi} u_{R}^{j} + y_{ij}^{d} \overline{q}_{L}^{i} \phi d_{R}^{j} + \text{h.c.}], \qquad (1.18)$$

where  $\tilde{\phi} = i\tau_2\phi^*$ ,  $f_\ell$ , and Yukawa terms for neutrino masses are omitted due to the absence of right-handed neutrinos. The complex  $3 \times 3$  matrices  $y_{ij}^{\ell}$ ,  $y_{ij}^{u}$  and  $y_{ij}^{d}$  are the Yukawa coupling matrices and allow for general flavour mixing with respect to the physical appearance of quarks and leptons as mass-eigenstates. Neglecting neutrino masses the flavour- and mass-eigenstates for electrons, muons and  $\tau$ -leptons coincide, such that in the unitary gauge the lepton sector is given by

$$\mathcal{L}_{\text{Yukawa}}^{\text{lepton}} = -\left(\frac{v+H}{\sqrt{2}}\right) \sum_{i,j=1}^{3} y_{ij}^{\ell} \overline{\ell}_{L}^{i} e_{R}^{j} + \text{h.c.}$$

$$= -\left(\frac{v+H}{\sqrt{2}}\right) \left[y^{e}(\overline{e}_{L}e_{R} + \overline{e}_{R}e_{L}) + y^{\mu}(\overline{\mu}_{L}\mu_{R} + \overline{\mu}_{R}\mu_{L}) + y^{\tau}(\overline{\tau}_{L}\tau_{R} + \overline{\tau}_{R}\tau_{L})\right],$$
(1.19)

and therefore the masses of the quarks are given by

$$m^{\ell} = \frac{y^{\ell}v}{\sqrt{2}} \tag{1.20}$$

and accordingly, a coupling of the Higgs boson to leptons appears, which is linearly proportional to the mass. The Yukawa couplings, and therefore lepton masses, are free parameters of the theory and can only be determined experimentally.

For quarks, the mass-eigenstates do not coincide with the flavour-eigenstate. Instead, the Yukawa matrices in equation 1.18 can be diagonalized via  $y_{ij}^{\prime d} = V_L^d y_{ij}^d V_R^{d\dagger}$  and  $y_{ij}^{\prime u} = V_L^u y_{ij}^u V_R^{u\dagger}$ , with  $V_L^u V_L^{d\dagger}$  being the Cabibbo-Kobayashi-Maskawa (CKM) matrix [19, 20]. After this change of basis from the weak-eigenstates into the mass-eigenstates the Yukawa terms in the unitary gauge are of the form

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}}^{\text{quark}} &= -\left(\frac{v+H}{\sqrt{2}}\right) \left[\sum_{i=d,s,b} y_i'^d \bar{q}_{iL}^d q_{iR}^d + \sum_{i=u,c,t} y_i'^u \bar{q}_{iL}^u q_{iR}^u\right] + \text{h.c.} \\ &= -\left(\frac{v+H}{\sqrt{2}}\right) \left[\sum_{i=d,s,b} y_i'^d (\bar{q}_{iL}^d q_{iR}^d + \bar{q}_{iR}^d q_{iL}^d) + \sum_{i=u,c,t} y_i'^d (\bar{q}_{iL}^u q_{iR}^u + \bar{q}_{iR}^u d q_{iL}^u)\right]. \end{aligned}$$
(1.21)

and therefore the masses of the leptons are given by

$$m_{\alpha} = \frac{y_{\alpha}^{\prime d} v}{\sqrt{2}} \quad (\alpha = d, s, b), \tag{1.22}$$

$$m_{\alpha} = \frac{y_{\alpha}^{\prime u} v}{\sqrt{2}} \quad (\alpha = u, c, t).$$
(1.23)

As in the case of lepons, since the quantities  $y_{\alpha}^{u,d}$  are unknown parameters of the SM, the masses of the quarks cannot be predicted and must be obtained from experimental measurements.

It is worth noticing that the heavier a particle is, the stronger it couples to the Higgs boson. Therefore the Higgs decays more often into more massive particles provided that such a decay is kinetically allowed.

In summary, the mass terms for all necessary elementary particles are generated

by introducing the complex scalar  $SU(2)_L$  doublet, the Higgs field, with the potential written in 1.6. In addition, a new physical scalar Higgs boson is predicted by the SM, with couplings to massive fermions and vector bosons particles:

$$g_{Hff} = \frac{m_f}{v} \quad f = (u, d, c, s, t, b, e, \mu, \tau), \qquad g_{HVV} = 2\frac{m_V^2}{v} \quad V = (W, Z).$$
(1.24)

Therefore the theory predicts a linear relation between masses and couplings, and the measurement of the couplings of the new boson discovered in Run-I constitutes a test for the Standard Model.

### **1.2** Cross Section Calculation at Hadron Colliders

The underlying theoretical framework of the calculation of cross sections for a hard process at a hadron-hadron collider is the parton model [21], in which a high energy hadron A, in our case a proton, is seen as a composition of quasi-free partons, i.e. quarks, and gluons, and has longitudinal hadron momentum  $p_A$ . The parton i has longitudinal momentum  $p_i$  and carries the momentum fraction  $x_i = p_i/p_A$ . The cross section calculation is based on the factorization theorem stating that the cross section is given by the convolution of the parton distribution functions (PDF)  $f_i(x, \mu^2)$  for the colliding hadrons (A, B) and the hard parton-parton cross section  $\sigma_{ij}$ :

$$\sigma(AB \to Z/\gamma^*) = \sum_{i,j} dx_i dx_j f_{i,A}(x_i, \,\mu^2) f_{j,B}(x_j, \,\mu^2) \cdot \sigma_{ij}(ij \to Z/\gamma^*, \hat{s}, \mu^2), \quad (1.25)$$

where the Drell-Yan process (schematised in figure 1.3) is taken as example, as the standard candle process of Z production into  $\tau$  leptons will be studied in this thesis. The hadrons A, B are pp at the LHC. The parton distribution function  $f_{i,A}(x_i, \mu_2)$  describes the probability density for finding a parton i inside the hadron A carrying a longitudinal momentum fraction  $x_i$ . The resulting squared partonic centre of mass energy is given by  $\hat{s} = x_1 x_2 (p_A + p_B)^2$ . The factorization scheme serves as a method to systematically eliminate collinear divergencies from the parton cross section  $\sigma_{ij}$  and absorb them into the parton distribution functions. The PDFs and the parton-parton cross section depend on the factorization and renormalisation scale  $\mu$ . The NNLO Drell-Yan production cross sections can be calculated with a precision of 1% for the renormalisation-scale and factorisation-scale uncertainty. For inclusive Z-boson production at the LHC the NLO-to-LO correction is sizeable (about 20%), while the NNLO-to-NLO correction amounts to 2% only. The DY cross section as a function of the invariant mass, as measured at  $\sqrt{s} = 7$  TeV and as predicted by NNLO calculations, is shown in figure 1.4.



(a) Drell–Yan process.



(b) Drell-Yan cross section as function of the invariant mass.

Figure 1.3: Left: Drell–Yan process, where a quark from one hadron and an antiquark from another hadron annihilate to create a pair of leptons through the exchange of a  $Z/\gamma^*$  boson. Right: DY invariant mass spectrum, normalized to the Z resonance region, as measured and as predicted by NNLO calculations, from [22].

## **1.3** Higgs Production and Decay at the LHC

The theoretical predictions for the properties of the Higgs boson are summarized in detail in [23–25]. Figure 1.4(a) shows a comparison of production cross-sections for Standard Model physics process of different type. As can be seen, the cross-section of Z-bosons exceeds the one for the Higgs boson by about five orders of magnitude, leading to overwhelming partially irreducible background for Higgs searches.

At the LHC, the main production processes are in order of importance: gluon fusion  $(gg \rightarrow H)$ , vector boson fusion  $(qqH \text{ or } q\bar{q}H)$  and Higgs boson production in association with a vector boson  $(W^{\pm}H \text{ or } ZH)$  or with a top-quark pair  $(t\bar{t}H)$ . The corresponding Feynmann diagrams are shown in figure 1.5. The corresponding cross-sections as predicted by the SM as a function of the center-of-mass energy are shown in figure 1.4(b).

**Gluon Fusion** The Higgs boson production via gluon-gluon fusion is mediated by triangular loops of heavy quarks. Due to the enhanced partonic gluon luminosity at high energies and the large couplings to especially top-quarks, the production mechanism has the largest contribution to the overall Higgs production cross-section, exceeding the other production mechanisms by at least one order of magnitude. The cross section is highly dependent on higher order QCD corrections.

**Vector Boson Fusion** The VBF production of the Higgs boson has the second largest contribution to the overall cross-section of the Higgs boson at the LHC. The process is induced by t- and u-channel<sup>4</sup> Z or W boson exchange interaction of two (anti-)quarks of the incoming protons, with the Higgs boson being radiated off the weak bosons, linking the two quarks. VBF production via W bosons exceeds the one with Z bosons by a factor of about 3, due to the larger W coupling to fermions. The VBF production mechanism leads to a distinct final state topology with two jets in the final state. Due two the absence of color exchange at leading order between the two incoming protons<sup>5</sup>, additional gluon radiation in the central rapidity region between the two final-state quarks is strongly surpressed. This leads to two high energetic jets at, on average, large pseudo-rapidities, with large pseudorapidity gaps and without additional high energetic jet activity in between, which is a distinguished feature in comparison to background processes and the other Higgs boson production mechanisms.

Z/W+H and tt+H The Higgs boson production via Bremstrahlung off weak vector bosons and heavy quark associated Higgs production processes, are additional important LHC Higgs production mechanism. The cross sections are low compared to the gluon fusion and VBF production cross section. The vector boson associated production offers the most sensitive possibility to measure Higgs boson decays into bottom quarks, where final state leptons from vector boson decays can be used to trigger events and suppress the otherwise overwhelming QCD background.

 $<sup>{}^{4}</sup>t$  and u are the Mandelstam variables, s-channel can be neglected.

<sup>&</sup>lt;sup>5</sup>Weak bosons are color singlets.



(a) Cross sections comparison.



(b) Higgs boson production cross sections as a function of the centre-of-mass-energies.

(c) Higgs Boson decay branching ratios.

Figure 1.4: Top: Comparison of several Standard Model cross sections as a function of collider energy, from [26]. Bottom: Standard Model Higgs boson production cross sections and decay branching ratios, from [27].



Figure 1.5: Main Higgs production processes driagrams in order of importance: gluon-fusion (top-left), vector boson fusion (top-right), W/Z associated production (bottom-left) and top-quark associated production (bottom-right). From [28].

#### 1.3.1 Higgs Decays

The partial decay widths are, once the Higgs boson mass is known, predicted by the SM, considering the couplings shown in equation 1.24. The couplings are proportional to the mass, and therefore the Higgs boson has the tendency to primarily decay into the heavier bosons and fermions, depending on the Higgs mass and phase space. In addition, decays into massless photons or gluons are possible mediated by loop diagrams, similiar to the production via gluon fusion. The Higgs boson decay Branching Ratios are reported in figure 1.4. At  $m_H = 125$  GeV the most prominent decays are  $H \rightarrow b\bar{b}$ ,  $H \rightarrow gg$  and  $H \rightarrow \tau \tau$ . In particular, due to the large irreducible QCD background for  $H \rightarrow b\bar{b}$ , the decay into tau-leptons is the most sensitive decay mode to study and measure fermionic Higgs decays and couplings, compared to the decay into muons with a decreased branching ratio by 2 orders of magnitude.

# CHAPTER 2

# The CMS Experiment at the LHC

The European Organization for Nuclear Research (CERN) is a European research organization based near Geneva, across the Swiss-French border. Established in 1954 it has 22 member states, involving the collaboration of more than 10 000 scientist representing over 600 universities and research facilities [29]. Several important achievements in particle physics have been made through experiments at CERN, such as the discovery of weak neutral currents in the Gargamelle bubble chamber [30], the discovery of W and Z bosons in the UA1 and UA2 experiments [31, 32] and the discovery of a boson consistent with the long sought Higgs boson [1, 2].

CERN's main function is to provide particle accelerators and other infrastructure needed for particle physics research. CERN hosts a proton linear pre-accelerator, a heavy ion linear pre-accelerator and corresponding accumulator ring, a small "booster" proton synchrotron, an antiproton accumulation ring, the Proton Synchrotron (PS), the Super Proton Synchrotron (SPS) and the Large Hadron Collider (LHC) (see figure 2.2).

### 2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a superconducting hadron collider and accelerator. It lies at a depth of over 100 m in the 26.7 km tunnel that previously hosted the LEP  $e^+e^-$  accelerator. The LHC is based on the synchrotron principle and it is deigned to accelerate two beams of proton, or heavy ions, in opposite directions to be collided at a center-of-masss energy of up to  $\sqrt{s} = 14$  TeV. The beams travel, in opposite direction, in two adjacent parallel beam pipes that intersect at four points. They are bent by 1232 superconducting NbTi dipole magnets and focused by an additional 392 quadrupole magnets, kept at 1.9 K by approximately 96 tonnes of superfluid Helium-4. The two beams are accelerated by 16 niobium filmed copper radio frequency cavities cooled down to 4.5 K increasing the proton energy by about 16 MeV per turn.



LHC DIPOLE : STANDARD CROSS-SECTION

Figure 2.1: Diagram of an LHC superconducting dipole magnet. From [33].

Acceleration The protons injected in the LHC are pre-accelerated by the succession of machines shown in the scheme in figure 2.2. The proton source is a bottle of hydrogen gas, where an electric field is used to strip hydrogen atoms of their electrons. The protons are first injected into the linear accelerator Linac2 and then transferred to the Proton Synchrotron Booster at an energy of 50 MeV. The PCB, which accelerates the beam to 1.4 GeV, is followed by the Proton Synchrotron, which pushes the beam to 25 GeV and arranges the particles into bunches with the correct time spacing. Each beam is organized in groups of protons called bunches with a nominal value of 2808 bunches, and a value up to 2220 in 2016. Bunches contain around  $1.2 \cdot 10^{11}$  particles and are separated between one another by a time interval of 25 ns. The last pre-accelerator before the LHC is the Super Proton Synchrotron where the beam is accelerated to 450 GeV. Finally the protons are accelerated by the LHC with 16 niobium filmed copper radio frequency cavities cooled down to 4.5 K increasing the proton energy by about 16 MeV per turn. The beams are accelerated to 6.5 TeV (0.999 999 991 the speed of light) in about 20 to 25 minutes.



# **CERN's Accelerator Complex**



**Luminosity** The rate of events generated in the LHC is given by

$$R_{\rm evt} = \mathcal{L}(t) \cdot \sigma_{\rm evt} \tag{2.1}$$

where  $\sigma_{\text{evt}}$  is the cross-section of the process and  $\mathcal{L}$  the instantaneous luminosity of the machine. The total number of events produced in a given time interval is given by the integral over time of the event rate:

$$N_{\text{evt}} = \int R_{\text{evt}}(t) dt = \int \sigma_{\text{evt}} \cdot \mathcal{L}(t) dt = \sigma_{\text{evt}} \cdot \int \mathcal{L}(t) dt = \sigma_{\text{evt}} \cdot L_{\text{int}}.$$
 (2.2)

The nominal instantaneous luminosity of the LHC is  $1.0 \cdot 10^{34}$  cm<sup>-2</sup>s<sup>-1</sup>. The machine is designed to maximize the integrated luminosity  $L_{\text{int}}$ .



Figure 2.3: Accumulative plots of integrated luminosity recorded by the CMS Experiment in 2016. From [35].



CMS Peak Luminosity Per Day, pp, 2016,  $\sqrt{s} =$  13 TeV

Figure 2.4: Peak delivered luminosity per day. From [35].

**Experiments** The LHC provide collisions for seven experiment. The beams are brought into collision in four *interaction points* (IP) around which the four big experiments were built:

- ALICE (A Large Ion Collider Experiment), is optimized to study heavy-ion (Pb-Pb nuclei) collisions and the properties of the quark-gluon plasma [36].
- ATLAS (A Toroidal LHC ApparatuS), is a general purpose detector designed to test the Standard Model, search the Higgs Boson and search for new physics [37].
- **CMS** (Compact Muon Solenoid), is a general purpose detector. It will be described in detail in 2.2.
- LHCb (Large Hadron Collider beauty), is a specialized b-physics experiment dedicated to the study of heavy-flavor physics, the production of new baryons, the study of CP violation and the search for rare decays predicted in theories beyond the standard model [38].

The other three smaller experiments are located close to the large ones and are:

- **LHCf** (Large Hadron Collider forward), is a special-purpose experiment for astroparticle physics and is designed to study the particle showers generated in the "forward" region of very high energy collisions, helping to interpret the interaction of ultra-high-energy cosmic rays with the atmosphere [39].
- **MoEDAL** (Monopole and Exotics Detector At the LHC), has the prime goal to directly search for the magnetic monopole or dyon and other highly ionizing stable massive particles and pseudo-stable massive particles [40].
- **TOTEM** (Total Cross Section, Elastic Scattering and Diffraction Dissociation), observes the forward particles originating from the collisions at the center of CMS to measure the total scattering cross-section for protons and to study diffractive and elastic processes [41].

## 2.2 The Compact Muon Solenoid

The CMS experiment [42] is one of the two general purpose big experiments at the LHC. The detector, as suggested by the name itself, it's based around a solenoid that produces a strong magnetic field (currently 3.8 T). A stronger magnetic field bends paths of charged particles more and it allows, in combination with the precision position measurements in the tracker and muon detectors, precise momentum measurement, even for particles with very high energy. The aim of the CMS experiment is to investigate the physics at the TeV scale, the mechanism of the electroweak symmetry breaking and the search for physics beyond standard model (SUSY, dark matter, etc.). For this physics program excellent momentum and energy resolution, highest particle identification efficiencies for photons, electrons, muons, taus, jet, missing transverse energy up to the TeV scale and a large angular coverage ( $4\pi$  hermetic configuration) are mandatory.

#### 2.2.1 Overall design

Overall, the CMS detector is about 28.7 m long and has a diameter of 15.0 m with a weight of about 14000 tonnes. The subdetectors are arranged radially around the beam axis from the inside in an order such that those detectors that perform a nondestructive measurement of the particles are positioned in the inner part, followed by detectors that perform a destructive measurement (see fig. 2.5 and fig. 2.19 on page 36.). For superb charged-particle momentum resolution and detection of secondary vertices and thus identification of taus and b-jets, the pixel and silicon tracking system is placed in the center of the detector. The tracker is surrounded by the high granular electromagnetic calorimeter to absorb and measure the energy of electrons and photons with high-resolution to be followed by the hadronic calorimeter with large hermetic coverage for precise measurements of missing transverse energy and jet energies. The solenoid surrounds the calorimeters and produce the magnetic field of 3.8 T without causing further "dead" material in front of the calorimetric system. The outermost layer consists of the return-voke for the magnetic field interweaved with the muon chambers for the identification of muons with long radiation length and to provide improved energy resolution for TeV muons complementing the muon track measurements.

**Cartesian coordinate system** The  $\hat{x}$  coordinate axis is defined along the LHC radius pointing the centre of the ring. The  $\hat{y}$  coordinate axis is perpendicular to the  $\hat{x}$  axis and pointing to upwards. The  $\hat{z}$  coordinate axis is tangent to the beam with a direction that complete a right handed coordinate system.

**Polar coordinate system** The polar angle with respect to the positive  $\hat{z}$  axis direction is called  $\theta$  and goes counter-clockwise with respect to the positive  $\hat{z}$  axis. The azimuthal angle in the xy plane is called  $\phi$  and also goes counter-clockwise with respect to the positive  $\hat{x}$  axis.

#### 2.3. THE TRACKING SYSTEM



Figure 2.5: Illustration of the CMS detector with labeled subsystems. From [43].

"Particle physics" coordinate system Usually instead  $\phi$  of using  $\theta$  it is convenient to use the pseudorapidity defined as

$$\eta = -\ln(\tan(\frac{\theta}{2})).$$

It is preferred over  $\theta$  as the rate of particle production at hadron colliders is approximately constant as a function of  $\eta$  due to the Lorentz invariance of rapidity differences.



## 2.3 The Tracking System

The CMS tracking system [42, 44] is the most central part of the detector and is entirely based on silicon tracking technology. It is specifically designed to provide precise and efficient measurement of the tracks of charged particles as well as accurate reconstruction of secondary vertices. The CMS tracker covers a region of up to  $|\eta| < 2.5$ and it is cooled to operate at -10 °C. With 200 m<sup>2</sup> of active silicon area it is the largest semiconductor tracker ever built for a physics experiment. The CMS silicon tracker consists of two main parts: the pixel tracker and the silicon strip tracker.



Figure 2.6: The CMS silicon tracker in an rz view. Single-sided silicon strip module positions are indicated as solid light (purple) lines, double-sided strip modules as open (blue) lines, and pixel modules as solid dark (blue) lines. From [45]

#### 2.3.1 The Pixel Tracker

The pixel system is the part of the tracking system that is closest to the interaction region: its main purpose is to measures the origin of charged tracks. It achieves a high spatial resolution both in the xy plane, to identify secondary vertices as in b-quark decays, and along the  $\hat{z}$  axis, to separate particles originating from different p-p collisions in the same bunch crossing. It's composed by cylindrical layers of hybrid pixel detector modules (BPix) that surround the interaction point, and by two fan shaped disks of pixel modules (FPix) for each endcap (see figure 2.7).

The total pseudorapidity coverage is  $|\eta| < 2.5$ . The pixel detector contains 66 million pixel elements with the size of  $100 \times 150 \,\mu\text{m}^2$  across 1440 modules and delivers three high precision space points on each charged particle trajectory. The spatial resolution is 10 µm in the  $r - \phi$  plane and 20 µm in z direction in the barrel, and about 15 µm and 20 µm respectively in the endcaps.

#### 2.3.2 The Strip Tracker

The radial region between 20 cm and 116 cm is occupied by the silicon strip tracker that has the purpose to measure the track's momentum, complement the measurement of the pixel detector and follow the track up to the calorimeters and for the muons to point to their impact position in the muon system detectors. The strips are purposely misaligned by a tilt angle of 100 mrad and thanks to the crossing of overlapping strips it is possible to measure the z coordinate. The CMS strip tracker is composed of ten cylindrical layers in the barrel region (four inner ones and six outer ones), and twelve disks for each endcap (three inner and nine outer ones). The layout of the detector is



Figure 2.7: 3D view of the pixel tracker structure, showing the barrel layers (green) and endcap fans (pink). From [42].

shown in figure 2.6. The strip tracker provides precise spatial measurements (40 to 60 µm) in the  $r - \phi$  plane. The resolution in the  $\hat{z}$  direction is 500 µm. The momentum resolution for a 10 GeV charged particle track is 0.5%, and for a 100 GeV particle is about 1%. The total pseudorapidity coverage is  $|\eta| < 2.5$ .

#### 2.4 The Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) [42, 46] is a highly homogeneous and hermetic calorimeter made of 61200 lead tungstate (PbWO<sub>4</sub>) crystals mounted in the central barrel part, closed by 7324 crystals in each of the two endcaps (see figure 2.8 and 2.9). The ECAL purpose is, in combination with the tracking detector, to identify electrons and photons and facilitates high resolution energy measurements.

The high density ( $\rho = 8.28 \text{ g/cm}^3$ ), short radiation length ( $X_0 = 0.89 \text{ cm}$ ) and the small Molière radius ( $R_M = 2.2 \text{ cm}$ ) of the crystals allow for a compact design of the ECAL with about 25 interaction lengths and enables a fast calorimeter response with 80% of light being collected in 25 ns. The light output varies with temperature therefore, in order to maintain the energy resolution, a cooling system is installed to keep the crystal temperature stable at  $18 \pm 0.05$  °C. The energy resolution of the ECAL is

$$\frac{\sigma(E)}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{12\%}{E} \oplus 0.3\%.$$



Figure 2.8: Layout of the CMS electromagnetic calorimeter. From [42].



Figure 2.9: Transverse section through the ECAL, showing geometrical configuration in terms of  $\eta$ . From [47].


Figure 2.10: A lead tungstate crystal with the VPTs. From [48].

**ECAL Barrel** The ECAL barrel (EB) covers the pseudorapidity range  $|\eta| < 1.479$  and accommodates 61200 crystals to achieve a granularity of about  $(\Delta \phi, \Delta \eta) \approx (0.0174, 0.0174)$ . Scintillation photons emitted in electromagnetic showers are collected in magnetic-field-resistant silicon avalanche photodiodes (APDs).

**ECAL Endcap** The ECAL endcaps (EE) cover the range  $1.479 < |\eta| < 3.0$  and consist of  $2 \times 3662$  crystals per endcap, mounted in two disk halves. Photoelectrons are collected via vacuum phototriodes (VPTs), chosen for the high radiation environment and magnetic field orientation in the endcap regions. Figure 2.10 shows a crystal matched with a VPT.

For improved  $\pi^0$ /photon discrimination, a preshower detector (ES) is placed in the pseudorapidity region  $1.65 < |\eta| < 2.61$ . When neutral pions decay to a closely spaced pair of photons, the two photons are often difficult to separate by the use of the ECAL detector alone. The ES is designed to provide a measurement of the impact points of the photons with a higher resolution. It is a 20 cm thick sampling calorimeter, composed out of two alternating layers of lead and silicon strip sensors of orthogonal orientation. The ES is placed in front of the EE, adding 3 radiation lengths<sup>1</sup>. It also helps the identification of electrons against minimum ionising particles and improves the position determination of electrons and photons with high granularity.

<sup>&</sup>lt;sup>1</sup>About 95% of single incident photons start showering before the second sensor plane.

# 2.5 The Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) [42, 49] is a sampling calorimeter<sup>2</sup> that has the purpose to measure the energy of charged and neutral hadron, fundamental for jet reconstruction and missing transverse energy. These reasons demand a large hermetic coverage.

The HCAL is composed by four subdetectors arranged in the structure shown in figure 2.11. Surrounding the ECAL, and immersed in the 3.8 T magnetic field, the HCAL barrel (HB) and the HCAL encaps (HE) cover the pseudorapididy region of  $|\eta| < 1.3$  and  $1.3 < |\eta| < 3.0$  respectively. The forward calorimeter (HF) covers the region close to the beam axis  $3.0 < |\eta| < 5.2$ . To improve the central shower containment in the region  $|\eta| < 1.26$  an array of scintillators is located outside of the magnet in the outer barrel HCAL (HO). The effective thickness of the HCAL ranges between roughly 6 hadron interaction lengths to roughly 11 interaction lengths for the barrel ( $|\eta| < 1.3$ ), the magnet adds about 1 interaction length in front of the HO. The energy resolution of the HCAL is

$$\frac{\sigma(E)}{E} = \frac{70\%}{\sqrt{E}} \oplus 7\%.$$

**Encaps and barrel** The HB and HE are sampling calorimeters consisting of plastic scintillator interlayed with copper absorber plates. Since the HCAL is placed inside the solenoid, the absorber material are chosen to be non-magnetic, copper alloy and stainless steel. The scintillation light is collected by wavelength-shifting fibres and it is read with hybrid avalanche photodiodes (HPDs).

**Forward calorimeter** The purpose of the HF is to improve the hermicity of the HCAL system in order to reinforce the measurement of the missing trasverse energy. Due to the high radiation environment of about 10 MGy in this region, the HF uses more robust quartz fibres as active material and steel as absorber material. The scintillation light is collected with radiation resistant photomultiplier tubes (PMT).

**Outer calorimeter** The HO is placed outside the solenoid and within the return yoke, along with the barrel muon detector. It is composed only of scintillators, using the solenoid as absorber.

<sup>&</sup>lt;sup>2</sup>A sampling calorimeter is a calorimeter where the absorber material and the scintillating material are interleaved. The absorber material induces a hadronic shower and the scintillating material samples the shower along its length.

# 2.5. THE HADRONIC CALORIMETER



Figure 2.11: Longitudinal view of the CMS detector showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters. From [47].



Figure 2.12: One half of the HCAL Barrel (foreground) seen here in preparation for mounting inside the CMS magnet in 2006. From [50].

# 2.6 The Superconducting Solenoid

The superconducting solenoid magnet [42, 51] is the core part of the detector design and the part that gave the name for the experiment. Since the muon play a major role in the physics program of CMS, an efficient identification and a high precision measurement of their momentum is fundamental. This is achieved with a very wide and strong magnet.

The solenoid surrounds the tracker and the calorimetric system, it is 12.3 m long, has a diameter of 6 m and it is designed to operate at 4 T<sup>3</sup>. It is composed by four layer of coiled superconducting NbTi<sup>4</sup> cooled down to 4.6 K with liquid helium, conducting a nominal current of 19.41 kA. The CMS magnet is the most powerful single magnet ever built, storing a total energy of 2.6 GJ in its field.

The solenoid is complemented with a 10 kt iron return yoke in which the magnetic field saturates to a value of 1.8 T. The return yoke has the purpose to reduce the intensity of the magnet field on components far from the magnet, to guide the magnetic field outside the bore of the magnet and to enhance the resolution of muon momenta measurements. The return yoke also incorporates the muon system and absorbs particles different from muons. A precise mapping of the magnetic field in the CMS barrel yoke is reported in figure 2.13.



Figure 2.13: Value of |B| (left) and field lines (right) predicted on a longitudinal section of the CMS detector, for the underground model at a central magnetic flux density of 3.8 T. From [52].

 $<sup>^{3}</sup>$ In normal operation the magnet works at 3.8 T in order to lengthen the lifetime.

<sup>&</sup>lt;sup>4</sup>The same technology used in the LHC magnets (Rutherford cable).

## 2.7 The Muon System

The CMS muon system [42, 53] illustrated in figure 2.15 is designed to reconstruct the momentum and charge of muons. Muons do not annihilate with ordinary matter via electromagnetic force, since ordinary matter does not contain muons, and do not decay via the strong force, since they do not couple to gluons. Muons can only decay via the weak interaction into an electron and two neutrinos, this leads to a long lifetime of the order of  $\tau = 2.2 \ \mu s^5$ . Muons can lose their energy only through ionization up to momenta of around 100 GeV/*c* when a sizable amount of energy is lost due to bremsstrahlung; the emitted photons then start an electromagnetic shower [54]. In figure 2.14 the stopping power for positive muons in copper is reported as example. The energy loss for all muons produced at the LHC is small enough that they can cross all the detector and escape from it. Therefore in CMS, the identification of muons is accomplished by simply measuring hits in the muon chamber (see fig. v2.19).

The muon system, hosted in the magnet return yokes of CMS, is divided into the barrel,  $|\eta| < 1.2$ , and the endcaps  $|\eta| < 2.4$ , and is composed of three types of gaseous detectors: drift tube (DT) chambers in the barrel region, cathode strip chambers (CSC) in the endcap region and resistive plate chambers (RPC) in both the regions.

The transverse momentum resolution  $\sigma(p_{\rm T})/p_{\rm T}$  of the CMS muon system is ~ 1% at 100 GeV/c and ~ 10% at 1 TeV/c.



Figure 2.14: Stopping power  $(= -\langle dR/dx \rangle)$  for positive muons in copper as a function of  $\beta \gamma = p/Mc$ ). Solid curves indicate the total stopping power. The radiative contribution is the sum of the bremsstrahlung, direct pair production, and photonuclear interactions contributions. From [54].

 $<sup>^5{\</sup>rm The}$  muon lifetime is further length ened by relativistic time dilation in the laboratory frame by relativistic time dilation.



Figure 2.15: Layout of one quadrant of CMS. The four DT stations in the barrel (MB1-MB4, green), the four CSC stations in the endcap (ME1-ME4, blue), and the RPC stations (red) are shown. From [55].

## 2.7.1 Drift Tube Chambers

Drift tubes are installed in the barrel section of the detectors where the muon rate is smaller and the magnetic field weaker and more uniform. Drift tubes detect muons by measuring the track of ionization along their path. Electrons released in the ionization process drift towards the anode wire under the effect of a strong electric field and initiate an avalanche close to the wire. The drift of the charges in the avalanche generates a current, which is amplified and read out by the detector electronics. The drift tubes measure the position of the muon track from the time needed for electrons to drift to the wire.

The barrel muon system of the CMS detector consists of four stations integrated in the return yoke of the magnet (see fig. 2.15. In total, 60 drift chambers compose each one of the inner three stations, while 70 chambers are used in the outer station due to the presence of the yoke feet. Each chamber is composed by 3 "superlayers" (SL) of 4 staked layers of cells. Two SL measure the  $r - \phi$  coordinate and are located at the upper and lower face of the detector, one superlayer is positioned in between and measures



Figure 2.16: Sketch of a cell showing drift lines and isochrones. From [42].

the z coordinate. The fourth station does not contain the z-measuring layer and thus consists of only 2 SL. The drift chambers determine the position of the centroid of a track passing through them with a resolution of 1.5 mm, and can measure the angle of the track with respect to the anode wire with a precision of 20 mrad. The structure of a drift cell is shown in figure 2.16.

The DT system covers the pseudorapidity region  $|\eta| < 1.2$ .

## 2.7.2 Cathode Strip Chambers



7 trapezoidal panels forming 6 gas gaps

Figure 2.17: Layout of a CSC chamber. From [42].

In the barrel region of the detector, where the flux of the particles is higher and the magnetic field is stronger and inhomogeneous, Cathode Strip Chambers are chosen for their radiation hardness and fast response. The CSCs are multiwire proportional chambers, and use a gas under a strong electric field as active and amplification medium. There are four stations of CSCs in each endcap, with chambers positioned perpendicular to the beam line and interspersed between the flux return plates. All CSCs overlap in  $\phi$  to avoid gaps in the muon acceptance.

Each CSC is trapezoidal in shape and consists of six gas gaps, each gap having a plane of cathode strips and a plane of anode wires. The cathode strips in each chamber run radially outward and provide a precision measurement in the  $r - \phi$  bending plane. The anode wires are running almost perpendicularly to the strips, in order to provide measurements of  $\eta$  and the beam-crossing time of a muon. The structure of a CSC is reported in figure 2.17. Each CSC measures the space coordinates  $(r, \phi, z)$  in each of the six layers.

The CSC system covers the pseudorapidity region  $0.9 < |\eta| < 2.4$ .

#### 2.7.3 Resistive Plate Chambers

The Resistive Plate Chambers system (RPC) is a dedicated trigger system added in both the barrel and endcap regions. The RPCs provide a fast, independent and highly segmented trigger with a sharp  $p_{\rm T}$  threshold over a large rapidity range. The RPCs are composed of two parallel double-gap chambers filled with a mixture of gases<sup>6</sup>, operated in avalanche mode to ensure good operation at high rates.

# 2.8 Data Acquisition and Trigger System

The LHC provides proton-proton collisions at a nominal rate of 40 MHz, and every event has roughly the size, in terms of data, of 1 Mb. This leads to a raw data flow of 40 Tb/s, an amount which is impossible to store and process in a viable way. The CMS Data Acquisition and Trigger System (TriDAQ) [42, 56, 57] is designed and used to greatly reduce (reduction rate ~ 10<sup>5</sup>) the rate of events to be recorded and analyzed. In other words the task of the trigger system is to select and store on disk a small number of interesting collisions out of a huge number of trivial events.

The required rejection is too large to be achieved in a single processing step, so the full selection task is split into two steps: the Level 1 Trigger (L1), performing the initial selection, and the High Level Trigger (HLT), performing a more complex and finer selection.

#### 2.8.1 Level 1 Trigger

The Level 1 Trigger (L1), is designed to reduce the rate of events accepted for further processing to less than 100 kHz. The L1 is installed within a network of custom built hardware and firmware located in the sub-detectors front-end electronics nodes of specific trigger electronic devices. The L1 decision is based on the presence of local objects such as muons, photon, jets and electron, built from fast responding ECAL and HCAL

 $<sup>^{6}96.2\%</sup>$  R134a (C<sub>2</sub>H<sub>2</sub>F<sub>4</sub>), 3.5% iso-butane and 0.3% SF<sub>6</sub>.

sub-detector outputs and the muon system response. It also employs global sums of the transverse energy and missing transverse energy. Each of these items is tested against several  $p_{\rm T}$  or  $E_{\rm T}$  thresholds. The triggering of the L1 is in connection with a specific proton bunch crossing, and provokes the read-out of all the relevant signatures corresponding to the triggered bunch crossing from the detector. These global detector signatures serve as building block of trigger primitives.

The L1 accepts the following two types of inputs as decision for issuing a trigger:

- Energy deposited in the ECAL and HCAL and variables that can be rapidly calculated from this primary information. Examples of these variables are the number of jets, the total hadronic transverse momentum, the number of hadronic  $\tau$  lepton decays, the number of electron/photon candidates and calorimetric missing transverse energy.
- Information from the muon system. Muon tracks are reconstructed using the hits detected by the muon system. An event containing muons is automatically triggered.

The data fragments from the same L1 are then collected from the various read out buffers and assembled into a single event by the "event builder", to be sent to the HLT filter farm.

#### 2.8.2 High Level Trigger

The High Level Trigger (HLT), is a software framework designed to reduce the Level 1 Accept (L1A) rate of 100 kHz to a final output rate of approximately 1 kHz. It runs on a dedicated server farm installed in the vicinity of the CMS detector. The data fragments of each sub-detector pass the event building network, to be rebuilt as a single event in several stages. The fact that the HLT is implemented in software allows great flexibility in the choice of algorithms, for example the HLT can reconstruct a track from the hits in the silicon tracker and then trigger events on the base of the track topology. The last stage of HLT processing does reconstruction and event filtering with the primary goal of making datasets of different physics signatures.

If the HLT decision is positive, the event is forwarded to the CERN central Tier-0 site for full offline event reconstruction and data archival of several reconstruction steps distributed over the world-wide LHC Grid network.

The CMS TriDAQ system therefore consists of the CMS detector front-end readout electronics, L1 and HLT trigger system, online event building and data and detector system quality monitoring. The structure of the data acquisition system is shown in figure 2.18.



Figure 2.18: Architecture of the CMS DAQ system. The trigger (left), data acquisition (center), and detector control systems (right) work in parallel to select and collect event. From [42].



Figure 2.19: Illustration of paths and interactions of different type of particles in the CMS subdetectors. From [58].

# Chapter 3

# Physics Object Reconstruction

In this chapter a brief description of the event and object related reconstruction techniques with the CMS detector, relevant for this thesis, is discussed. Particular attention is given to the reconstruction and identification of muons, electrons and taus.

# 3.1 Primary Vertex Reconstruction

In the typical experimental conditions at the LHC many protons collide in the same bunch crossing. The position in space of each of these collision is denominated *vertex*. Only a very small fraction of the overall number of collisions is interesting, and these are usually surrounded by a number of non interesting *pile-up collisions*. Figure 3.1 shows an event with a very high number of pile-up interactions. For each bunch crossing the interesting collision is named *primary vertex* and it is selected by requiring the largest sum of  $p_T^2$  of all particles pointing to this vertex.



Figure 3.1: Event with 78 reconstructed vertices in one beam crossing, obtained from a special high pile-up run. From [59].

Vertices are reconstructed from a collection of good quality tracks selected from the event. Good quality tracks are tracks with at least five hits in the silicon detector, and at least two hits in the pixels, impact parameter significance in the xy plane smaller than 5 and normalized  $\chi^2$  smaller than 20. The vertex reconstruction can be split in two stages:

- 1. an optimal set of vertices is chosen as the origin for the provided tracks, this stage is perfomed by the Deterministic Annealing (DA) algorithm [60];
- 2. the position of each vertex is determined by a fit on the relative tracks, this stage uses a weighted least squares method called "adaptive fitting" [61].

The vertex must be compatible with the beamline within 1 cm.

## **3.2** Particle Flow Reconstruction

The Particle Flow reconstruction algorithm [62–64] allows combining the information provided by all the subdetectors to improve the identification of particles beyond the capabilities offered by a single subdetector, to provide a consistent global description of the collision and to get an improved measurement of their kinematic observables. The Particle Flow reconstruction is divided in four main steps:

- 1. **Tracking**, an iterative tracking algorithm [65], based on a Kálmán filter, is run on the set of hits measured in the silicon tracker.
- 2. Clustering the energy deposit in the ECAL and HCAL.
- 3. Linking, through a *link algorithm* that processes the clusters. This step relates reconstructed tracks to clusters of energy in the ECAL and HCAL. A graphical example is shown and described in figure 3.2.
- 4. Building of particles candidates.
  - Muon candidates are built from previously reconstructed muons and their expected energy deposits in the calorimeters are removed.
  - Electron candidates are built following up their trajectory in the ECAL together with potential radiated bremsstrahlung photons. A more detailed description is discussed in section 3.4. The electron clusters are then removed from the set of clusters not yet associated to a particle.
  - Unassociated ECAL clusters are then identified as photons, and remaining tracks ending in the calorimeters are identified as charged hadrons.
  - Finally, energy deposits in the HCAL that were not yet accounted for are identified as neutral hadrons.



Figure 3.2: Event display for an hadronic jet. A  $\pi^+$  is identified from its reconstructed track (T2), the absence of deposits in the ECAL and the observation of a shower in the HCAL (cluster H2). The dual cluster in the ECAL (E2, E3), identifies the photons from a  $\pi^0$ . The short (i.e. straight, energetic) track T1 that starts a large shower in the ECAL (cluster E1) and continues it into the HCAL (cluster H1), is identified as a  $\pi^-$ . An isolated ECAL cluster E4 is deduced to be a  $K_L^0$  that decayed to  $3\pi^0$  from the shape of its shower. From [62].

# 3.3 Reconstruction and Identification of Muon

In this section the muon reconstruction and identification techniques used in the CMS experiment are briefly described.

## 3.3.1 Muon Reconstruction

In CMS a muon can be identified and reconstructed in two approaches:

- *inside-out* where the reconstruction starts from a track in the tracker system and then is matched with a track in the external detectors when it is propagated out;
- *outside-in* where the reconstruction starts from a track identified in the muon system and then matched with a track in the silicon tracker after a propagation of its trajectory inside.

This distinction leads to four different types of muon reconstruction strategies:

- Standalone muons are muons reconstructed from hits in the muon system, independently from the tracking system. This type of muons usually have high purity<sup>1</sup>, but low precision in their momentum measurement, due to multiple scattering in the yoke material and limitation in momenta measurement of the muon system. Since the tracking system is very efficient muons of this type originating from collisions are very rare (~ 1%). Usually Standalone Muons are cosmic ray muons traveling through the muon system<sup>2</sup>. These muons are used for alignment and calibration of the muon system.
- Global muons are Standalone Muons with a matching track found in the inner tracking system. The track is then refitted combining the two information improving the measurement of the muon momentum. This type of muons have high purity, while the efficiency heavily depends on the muon momentum, reaching a plateau for  $p_T > 8$  GeV.
- **Tracker muons** are muon reconstructed from a track left in the tracking system that matches a muon segment<sup>3</sup> in the muon system. Muons of this type have a good purity and a very high precision determination of their momentum.
- Calorimeter muons are muons reconstructed from a track left in the tracking system and an amount of energy, compatible with a muon, deposited in the calorimetric system. These muons have very low purity and easily mimicked by other particles. Calorimeter Muons are usually not used in physics analyses, except in the case where low  $p_{\rm T}$  muons are needed.

<sup>&</sup>lt;sup>1</sup>Fraction of *genuine* muons in the sample of reconstructed muons.

<sup>&</sup>lt;sup>2</sup>The muon system acceptance volume is  $\sim 10^3$  times larger than the one of the silicon tracker.

<sup>&</sup>lt;sup>3</sup>A muon segment is a line connecting matched hits in the DT and CSC subdetectors.

## 3.3.2 Muon Identification

There are different type of muon identification selection criteria (with increasing level of purity):

- 1. Loose Muon selection The muon is required to be identified as a muon by the PFlow event reconstruction, and also to be reconstructed either as a Global Muon or as a Tracker Muon. This identification criteria is designed to be highly efficient for prompt muons and for muon from heavy and light quark decays.
- 2. Medium Muon selection Loose Muon with additional track-quality and muonquality requirements:
  - fraction of valid hits in the tracking system higher than  $0.8(^4)$ .
  - be reconstructed as a Global Muon.
  - normalized  $\chi^2$  of the Global track smaller than 3.
  - segment compatibility<sup>5</sup> higher than 0.303.
  - $\chi^2$  from the kink-finder on the inner track<sup>6</sup> smaller than 20.
  - $\chi^2$  of the standalone-tracker position match smaller than 12.

This identification criteria is designed to be highly efficient for prompt muons and for muons from heavy quark decays. This is the identification criteria used in the analysis presented in this thesis. A more detailed description of the criteria can be found in [66].

3. **Tight Muon selection** A muon reconstructed both as a Global Muon and a Tracker Muon with additional muon-quality requirements. This identification criteria is designed to have high purity.

# 3.4 Reconstruction and Identification of Electron

In this section the electron reconstruction and identification techniques used in the CMS experiment are briefly described.

### 3.4.1 Electron Reconstruction

The typical electron signature in CMS consists of hits in the silicon tracker and one or more clusters in the ECAL. Electrons are very light particles and by interacting with the tracker material they can radiate a considerable fraction of their energy via bremsstrahlung. More specifically, about one third of the electrons at 10 GeV radiate

<sup>&</sup>lt;sup>4</sup>For the analysis presented in this thesis this value has be lowered to 0.49 due to problems in tracker efficiency due to high ionizing particles.

<sup>&</sup>lt;sup>5</sup>The segment compatibility evaluates the compatibility of a (tracker) muon candidates with the muon hypothesis from hits in the muon chambers.

 $<sup>{}^{6}\</sup>chi^{2}$  based algorithm to reject muons from  $\pi/K$  decays.



Figure 3.3: Illustration of the bremsstrahlung electron energy loss and relative clusters. From [69].

more than 70% of their energy before reaching the ECAL, and in 10% of the cases the energy loss exceeds 95% [67]. To obtain a correct estimation of the initial electron energy, the energy deposited by the electron in the ECAL is summed with the energies deposited by the bremsstrahlung photons. Since the electron track is bent in the magnetic field the emitted photons lie along the tangents to the electron trajectory and result in ECAL cluster energy deposits spread in the  $\phi$  direction relative to the energy cluster of the electron itself, with about the same value of eta as the electron track. This is shown in figure 3.3.

The electron is reconstructed in the following steps:

- the energy is reconstructed building superclusters, i.e. cluster of clusters, starting from the highest energetic seed cluster using the *hybrid and island clustering algorithms* [67] extending to a maximum of 0.3 rad in  $\phi$  direction.
- the energy weighted mean position of the supercluster in the  $\eta \phi$  plane is propagated in the pixel detector for both charge hypotheses, searching for a first compatible hit within a  $\Delta \phi \Delta z$  window.
- the track seeded by the supercluster is then reconstructed using the pattern recognition technique called Gaussian Sum Filter (GSF) [68], in combination with the Bethe-Heitler model of bremsstrahlung.
  - This cluster based seeding technique is complemented with a tracker based seeding one for low  $p_{\rm T}$  electrons, where the track to cluster matching and track quality criteria are looser to reduce the number of possibilities for computation.
- to determine the sign of the electron charge the majority decision of the charge from the GSF track, the matched Kalman Filter track and the relative position of the innermost track hit to the supercluster is used.

A more detailed description can be found in [67].

### 3.4.2 Electron Identification

Certain physics processes that do not directly involve electrons can mimic electrons signatures, leading to non-negligible background contributions. Example of these processes are:

- photon conversion: photons can go through asymmetric conversion in the tracker producing one electron carrying the majority of the photon energy.
- semileptonic decay of heavy flavor hadrons: heavy hadrons can decay semileptonically into an electron carrying the majority of the momentum of the initial particle.
- inelastic charge exchange of charged pions and kaon:  $K^{\pm}$  or  $\pi^{\pm}$  can interact with the nuclei of the ECAL producing a  $\pi^0$  which then decays into photons.

Since there is no single variable that can be used to discriminate fake electron from the genuine ones with good purity and efficiency a multivariate (MVA) Boosted Decision Tree (BDT) method is used in order to improve the reconstructed electrons collection [70]. The BDT training used in the electron selection in this thesis is the *NonTrig Electron ID*. It is based on information such as shower or cluster shape, track information and track to cluster matching variables for electron candidates, which can be significantly different for real electrons and faking processes.

## 3.5 Lepton Isolation

Isolation is a measurement of the amount of activity in the proximity of a given lepton. Isolation is defined inside a cone  $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$  around the trajectory of the lepton. It is used to discriminate leptons produced in hard process from the ones produced in the decay of hadrons, since the last ones are usually associated with jets in the proximity. Two different type of isolation can be defined:

- Absolute Isolation  $I_{abs}$  is defined as the sum of the transverse momenta of the other particles in the  $\Delta R$  cone.
- **Relative Isolation** *I*<sub>rel</sub> is defined as follows:

$$I_{\rm rel} = \frac{\sum p_{\rm Ti}}{p_{\rm T\ell}} = \frac{\sum p_{\rm Ti} + \sum p_{\rm Ti} + \sum p_{\rm Ti} + \sum p_{\rm Ti}}{p_{\rm T\ell}} \qquad (3.1)$$

The calculation of the isolation is very easy using the Particle Flow algorithm, as the right-hand sum in 3.1 can be calculated summing over the transverse momenta of the different particles in the PF collections.

 $\Delta\beta$  correction for PU interaction Since we want to use isolation as propriety to investigate the production mechanism behind a particle in a specific collision, it is desirable to correct the equation 3.1 for the contribution of particles originating in pile-up vertices. While pile-up removal for charged particles is straightforward since they can be easily labeled with their primary vertices, for photons and neutral hadrons the correction has to rely on a estimator called " $\Delta\beta$ ". The current estimate for the amount of transverse momentum carried by PU neutral hadrons and photons combined is half the transverse momentum carried by PU charged hadrons. This estimation comes from the fact that the majority of the hadrons produced consists in pions and, due to isospin symmetry, the production ratio between charged pions and neutral pions is 2:1. The expression for relative isolation with the  $\Delta\beta$  correction becomes:

$$I_{\text{rel}\Delta\beta} = \frac{\sum p_{\text{T}i}}{p_{\text{T}\ell}} = \frac{\sum_{\text{charged}} p_{\text{T}i}}{p_{\text{T}\ell}} + \frac{\max\left(0, \sum_{\text{neutral had}} p_{\text{T}i} + \sum_{\gamma} p_{\text{T}i} - 0.5 \cdot \sum_{\text{PU charged had}} p_{\text{T}i}\right)}{p_{\text{T}\ell}},$$
(3.2)

where the  $\Delta R$  labels are omitted.

# 3.6 Reconstruction And Identification of Hadronic Tau

The  $\tau$  lepton has a short lifetime (0.29 ps), so it decays before reaching the detector elements. In ~ 65% of the cases,  $\tau$  leptons decay semi-hadronically, typically into one or three charged  $\pi$  mesons, often accompanied by neutral pions (decaying via  $\pi^0 \to \gamma \gamma$ ), and a  $\nu_{\tau}$ . The major  $\tau$  hadronic decays are listed in table 3.1.

Table 3.1: Major hadronic tau decays and corresponding branching ratio. The intermediate meson resonance is reported in the second column.

Decay mode	Resonance	Branching Ratio
$\pi^{-}\pi^{0}\nu_{ au}$	$ ho^-$	26.0%
$\pi^-  u_{ au}$		11.6%
$\pi^-\pi^+\pi^- u_ au$	$\alpha_1^-$	9.8%
$\pi^-\pi^0\pi^0 u_ au$	$\alpha_1^-$	9.5%
$\pi^-\pi^+\pi^-\pi^0\nu_{ au}$		4.8%
Other hadronic		1.7%
Total hadronic		64.8%

The tau signatures are typically isolated and form collimated jets with low charged track multiplicity, as illustrated in figure 3.4. Since  $\nu_{\tau}$  takes significant fraction of momentum the visible mass is lower than the actual lepton mass,  $m_{\tau} = 1776.86 \pm 0.12$  MeV.



Figure 3.4: Comparison between a  $\tau$  jet and a QCD jet. From [71].

Figure 3.5: Illustration of Run-1 strips with fixed size.

The main algorithm used in most CMS  $\tau$  analyses is the Hadron Plus Strip (HPS) algorithm [72, 73]. The algorithm uses PF-particles and is designed to optimize the performance of the  $\tau_{\text{had}}$  identification and reconstruction by considering the different hadronic decay modes of the  $\tau$  lepton individually. The algorithm starts the reconstruction of a  $\tau_{\text{had}}$  candidate from PF-Jet, which is used as an initial seed for the  $\pi^0$  components.

## **3.6.1** $\pi^0$ Reconstruction with Strip Object

Neutral mesons are reconstructed in ECAL as objects called "strips". The  $\pi^0$  decays mainly into a pair of photons that usually go through photon conversion in the tracker material. The electron tracks are bent in the magnetic fields, resulting in the broadening of the signal in the azimuthal direction. In the HPS algorithm this effect is taken into account by reconstructing photon in strips in the ECAL as follows:

- 1. the most energetic electromagnetic particle<sup>7</sup> within the PF-Jet is used as a seed and the initial position is centered around it.
- 2. the next highest  $p_{\rm T} e/\gamma$  that is within a window of size<sup>8</sup>

$$\Delta \eta = f(p_{\rm T}^{\gamma}) + f(p_{\rm T}^{\rm strip}),$$
$$\Delta \phi = g(p_{\rm T}^{\gamma}) + g(p_{\rm T}^{\rm strip})$$

around the strip center is merged to the strip. The functions f and g are determined using single  $\tau$  MC samples and have the following expressions:

$$f(p_{\rm T}) = 0.20 \cdot p_{\rm T}^{-0.66},$$
  
$$g(p_{\rm T}) = 0.35 \cdot p_{\rm T}^{-0.71}.$$

<sup>&</sup>lt;sup>7</sup>An "electromagnetic particle" is an electron or a photon.

<sup>&</sup>lt;sup>8</sup>In Run-1 the window had a fixed size of  $\Delta \eta = 0.05$  and  $\Delta \phi = 0.2$  around the strip center, as shown in figure 3.5.

The upper limit of the strip size is set to 0.3 in  $\Delta \phi$  and 0.15 for  $\Delta \eta$ , the lower limit is set to 0.05 for both  $\Delta \phi$  and  $\Delta \eta$ .

- 3. the strip position is recomputed as a  $p_{\rm T}$  weighted average of all  $e/\gamma$ .
- 4. the procedure is reiterated until no further  $e/\gamma$  is within the  $\eta \times \phi$  window.

The strips with  $p_{\rm T} > 2.5$  GeV are kept as  $\pi^0$  for further processing.

## 3.6.2 Tau Reconstruction

The  $\pi^0$  candidates are then combined with charged hadrons, reconstructed with the Particle Flow algorithm, to reconstruct the decay mode and calculate the  $\tau$  lepton four-momentum and isolation quantities. The hadronic tau reconstruction proceeds according to the following criteria:

- all the reconstructed hadrons and strips have to be contained into a cone of  $\Delta R = (3.0 \text{ GeV}/c)/p_{\text{T}}^{\tau}$ .
- the  $\vec{p}_{\rm T}$  of the reconstructed  $\pi$  is required to match the direction of the input PF-Jet.
- the decay must be compatible with a corresponding resonance ( $\rho$  or  $\alpha_1$ ) hypothesis.
- reconstructed candidates are required to be isolated.

The decay mode reconstruction criteria are presented in table 3.2. A real  $\pi^0$  decaying in two photons can result into zero, one or two strips depending on its energy and the angular separation between the two photons. A low energy  $\pi^0$  will not be usually reconstructed into a strip, whereas a  $\pi^0$  decaying into two well separated photons can be reconstructed into two different strips.

Table 3.2: HPS decay mode reconstruction. The " $\nu_{\tau}$ " stands for both the tau neutrino and tau antineutrino, depending on the  $\tau$  charge sign.

HPS topology	tau decay mode
1 hadron	$\tau^{\pm} \to \pi^{\pm} \nu_{\tau}$
	$\tau^{\pm} \to \pi^{\pm} \pi^0 \nu_{\tau}$ , with low energy $\pi^0$
1  hadron + 1  strip	$\tau^+ \to \pi^+ \pi^0 \nu_{\tau}$ , with both photons inside one strip
1  hadron + 2  strips	$\tau^{\pm} \to \pi^{\pm} \pi^0 \nu_{\tau}$ , with photons well separated
3 hadron	$\tau^{\pm} \to \pi^{\mp} \pi^{\pm} \pi^{\mp} \nu_{\tau}$

The HPS algorithm achieves an identification efficiency of 50 - 60%, with a misidentification probability ranging between per cent and per mille levels when combined with the MVA-isolation discriminator described in the next section.

#### 3.6.3 Tau Isolation Discriminators

Requiring  $\tau_{had}$  candidates to pass isolation requirements provides a strong reduction of the jet  $\rightarrow \tau_{had}$  misidentification probability as shown in figure 3.4. Two types of  $\tau_{had}$ isolation discriminators are used in CMS, namely the isolation-sum and the MVA-based discriminators. Only the MVA-based discriminator is described since it is the one used in the analysis presented in this thesis.

**MVA discriminator against jets** The MVA  $\tau_{had}$  identification discriminator combines the isolation and shape variables with variables sensitive to  $\tau$ -lifetime information to provide the best possible discrimination of  $\tau_{had}$  decays from QCD jets. A BDT is used to achieve a strong reduction in the jet  $\rightarrow \tau_{had}$  misidentification probability. Different working points, defined by cuts in the BDT, are implemented corresponding to different  $\tau_{had}$  identification efficiencies. The misidentification probability as a function of  $\tau_{had}$  identification efficiency is shown in figure 3.6 for both isolation-sum and MVA-based discriminators. The MVA discriminator demonstrates a reduction of misidentification probability by a factor of 2 at a similar efficiency with respect to the cut-based approach. The MVA identification discriminator is discussed more in detail in [72, 73].



Figure 3.6: Misidentification probability of the isolation discriminators as a function of  $\tau_{had}$  identification efficiency, evaluated using  $H \to \tau \tau$  and QCD MC samples. The points correspond to WPs of the discriminators. For the MVA-based discriminators the WPs are: Very Loose, Loose, Medium, Tight, Very Tight, and Very Very Tight working point, respectivey. The misidentification probability is calculated with respect to jets, which pass minimal  $\tau$  reconstruction requirements. From [73].

#### 3.6.4 Tau Discriminator Against Electron

Isolated electrons have a high probability to be misidentified as  $\tau_{had}$ 's that decay to either  $\pi^{\pm}$  or  $\pi^{\pm}\pi^{0}$ . In particular bremsstrahlung photons from electron that cross the tracker material can mimic  $\pi^{0}$ s in the decay mode reconstruction. In order to reduce the  $e \rightarrow \tau_{had}$  misidentification probability, while maintaining a high efficiency on genuine  $\tau_{had}$  decays, a MVA electron discriminator has been implemented [72, 73]. The BDT is trained with simulated samples containing genuine  $\tau_{had}$  and electrons, and has different working point according to the efficiency for real  $\tau_{had}$  to pass the discriminator. The expected efficiency and misidentification rate are reported in table 3.3.

Table 3.3: Expected efficiencies and misidentification rates of the  $\tau_{had}$  MVA discriminator against electrons, averaged over  $p_T$  and  $\eta$ .

WP	Efficiency	$e \rightarrow \tau_{\rm h}$ misidentification rate
Very loose	94.3%	$2.38\times 10^{-2}$
Loose	90.6%	$4.43 \times 10^{-3}$
Medium	84.8%	$1.38 \times 10^{-3}$
Tight	78.3%	$6.21 \times 10^{-4}$
Very tight	72.1%	$3.54\times10^{-4}$

#### 3.6.5 Tau Discriminator Against Muon

A MVA discriminator has also been trained to separate  $\tau_{had}$  decays from muons. It combines the calorimeter energies of any charged particle or photon constituting the  $\tau_{had}$ candidate with information from the muon system. The BDT is trained using samples of simulated Drell-Yan, W,  $t\bar{t}$  and Higgs events. Different WP are defined by changing the cutoff on the MVA output. The expected efficiency and misidentification rate are reported in table 3.4. The BTD discriminator against muon is described more in detail in [72].

Table 3.4: Expected efficiencies and misidentification rates of the  $\tau_{had}$  MVA discriminator against muons, averaged over  $p_T$  and  $\eta$ .

WP	Efficiency	$\mu \rightarrow \tau_{\rm h}$ misidentification rate
Loose	99.5%	$5.20 \times 10^{-4}$
Medium	99.0%	$3.67 \times 10^{-4}$
Tight	98.0%	$3.18 \times 10^{-4}$

## 3.7 Missing Transverse Energy

At the LHC the momentum of the colliding particles is aligned with the  $\ddot{Z}$  axis of the experiment, this leads to an initial transverse momenta of about zero. Since the linear momentum is conserved the sum of the transverse momenta of the collision products has to be zero,

$$\sum_{\text{initial}} \vec{p}_{T_i} = \sum_{\text{final}} \vec{p}_{T_i} \approx 0.$$

If the transverse momentum of the observed products is significantly different from the typical scale of the colliding partons, this condition can be interpreted as the result of one or more particles escaping undetected. The Missing Transverse Energy (MET), indicated as  $\not\!\!\!E_T$ , is the module of the transverse momentum that would be needed to balance to zero the sum of the collision products,

$$\sum_{\text{observed}} \vec{p}_{T_i} + \vec{E}_T = 0.$$

In the case of the measurement of the Higgs boson decaying into a pair of  $\tau$  leptons, MET occurs due to the presence of tau neutrinos.

# CHAPTER 4

# **Event Selection**

The goal of this thesis is the commissioning and study of the physics objects and analysis methods used in the search for the SM Higgs boson decays into a pair of tau leptons, in particular for the channels  $H \rightarrow \tau \tau \rightarrow \mu \tau_{had} 2\nu$  and  $H \rightarrow \tau \tau \rightarrow e \tau_{had} 2\nu$  (the "2 $\nu$ " notation will be dropped for the rest of the thesis). For this purpose the standard candle process of Z production into tau leptons<sup>1</sup> is studied, as it shares a lot of similarities with the Higgs decay process and consists in a major source of background to the Higgs measurement. The signature searched for consists in a lepton candidate (muon or electron) and an hadronic tau candidate. In figure 4.1 a schematic view of the  $\mu \tau_{had}$  channel is shown.

This chapter reports on the samples used for the analysis, on the selection criteria and the corrections needed to improve the agreement between the simulation and the data.



Figure 4.1:  $\mu \tau_{had}$  decay channel.

## 4.1 Data Samples and Triggers

The data sample used in this analysis was collected by the CMS detector between April and August 2016 during the runs B, C and D, for a total integrated luminosity of 12.9 fb<sup>-1</sup>. The data are saved in the following datasets:

• SingleMuon\_Run2016BCD-PromptReco-v2 for the  $\mu \tau_{had}$  channel. This dataset is formed by all the events that passed at least one of the single muon High Level Triggers, in other words if there is at least one passing muon trigger object. The specific trigger used in the analysis is IsoMu22, this trigger requires a well isolated muon candidate with a transverse momentum of at least 22 GeV.

<sup>&</sup>lt;sup>1</sup>What is actually measured is the Drell Yan (DY) process  $q\bar{q} \rightarrow Z/\gamma^* \rightarrow \tau\tau$ .

• SingleElectron\_Run2016BCD-PromptReco-v2 for the  $e\tau_{had}$  channel. This dataset is formed by all the events that passed at least one of the single electron High Level Triggers. The specific trigger used in the analysis is Ele25\_eta2p1 (Tight WP), this trigger requires an electron candidate with a transverse momentum of at least 25 GeV and pseudorapidity lower than 2.1, the *Tight* working point is used for better purity.

**Trigger Matching** The triggers are associated to trigger objects, which are particle candidates built from the reconstruction algorithm in the High Level Trigger [56]. This algorithm trades high precision in the reconstruction for speed in the execution of the code. Trigger objects are, indeed, the objects on which the HLT algorithms runs. Since later we will evaluate and apply corrections for trigger efficiency it is important that the lepton on which the efficiency measurement is performed is really the one responsible for the trigger. To guarantee that the leptons are responsible for the triggering of an event, the trigger objects associated to the trigger used are matched to the reconstructed muon or electron in the analysis. This operation, called trigger matching, requires that the trigger object lies in a cone centered on the direction of flight of the lepton. The size of the cone in the  $\eta - \phi$  plane in this analysis is equal to  $\Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2} = 0.5$ .

**Data Quality** We require that the analyzed events are recorded in "good" conditions. The events are defined as good if they were collected in experimental conditions where the LHC machine and all the detectors were in proper conditions to perform physics measurements. This certification is performed by the CMS Data Quality Monitoring and Certification group [74].

# 4.2 Simulated Samples and Background Modeling

#### 4.2.1 Simulation in High Energy Physics

The simulation of a high energy particle collision in general is divided into the following subsequent steps (illustrated in figure 4.2):

- 1. Matrix element calculation of the hard scattering process. This step consists in the calculation of the hard process at a fixed order of  $\alpha_s$ , using the momenta of the incoming partons, which are distributed according to the input Parton Density Functions (PDFs). The generators used are POWHEG[75] AMCNLO MG5[76] and LO MG5[76] depending on the process (see table 4.1 for the list of generated process). The parton density function used is CTEQ6 [77].
- 2. Simulation of the parton shower. This step consists in simulation of the higher order QCD effects, together with Initial State Radiation (ISR) and Final State Radiation(FSR). The simulation of the parton shower is managed by PYTHIA 8 [78].



Figure 4.2: Schematic view of a event-generation process for a p-p collision. From [82].

- 3. Simulation of the hadronization. This step consist in the recursively splitting of the radiated partons into pairs of  $q\bar{q}$  or gg pairs, reducing the energy scale of the event. The hadronization is modeled with the Lund string model [79] implemented in the PYTHIA 8 parton shower generator.
- 4. Simulation of the detector response. This step consists in the simulation of the interactions of long lived particles with the detector and the response of the latter to these particles. It is implemented in the Geant4 software framework [80, 81], which is interfaced to the CMS software framework. The output has the same format as real collision events, so that simulated data can be analyzed with the same software tools as collision events.

## 4.2.2 Monte Carlo Samples

The simulated samples used in this analysis are reported in table 4.1. The various process can be classified in the following groups:

- Simulated backgrounds:
  - Diboson events;
  - $t\bar{t}$  and single top events;

- W + Jets events, where  $W \rightarrow \ell \nu$  and the jet fakes a tau;
- Drell-Yan events resulting in ee or  $\mu\mu$  final states, where a lepton is misidentified as a tau;
- Simulated signal: Drell-Yan events resulting in  $\tau\tau$  final state.

The QCD background, where one jet fakes a tau and another jet fakes a lepton, is derived exploiting some control sample extracted from collision data (hereafter, data-driven) as described in section 4.2.3. All the simulated samples are subject to the same selection criteria used for the selection of events in data apart for the trigger matching since triggers were not simulated yet at the time of this thesis.

Table 4.1: List of simulated samples used in the analysis, the generator used, the cross section and the order of computation.

Process <sup>1</sup>	Generator	$\sigma[{\rm pb}]$	Order <sup>2</sup> <sup>3</sup>
$VV \rightarrow 2\ell  2\nu$	AMCNLO MG5	11.95	NLO
$WZ \rightarrow 1\ell  1\nu  2q$	AMCNLO MG5	10.71	NLO
$WZ \to 2\ell  2q$	AMCNLO MG5	5.595	NLO
$ZZ \to 2\ell  2q$	Amenlo $MG5$	3.22	NLO
$ZZ \to 4\ell$	POWHEG	1.256	NLO
$WW  ightarrow 1\ell  1 \nu  2q$	Amenlo $MG5$	1.212	NLO
$t\bar{t}$	POWHEG	831.76	NLO
Single $t$ (t-channel)	POWHEG	136.02	NLO
Single $\bar{t}$ (t-channel)	POWHEG	80.95	NLO
Single $t$ (tW-channel)	POWHEG	35.6	NLO
Single $\bar{t}$ (tW-channel)	POWHEG	35.6	NLO
$W + \text{Jets} \to \ell \nu \text{ (inclusive)}$	LO MG5	61526.7	LO
$W + \text{Jets} \rightarrow \ell \nu \ (1 \text{ jet f.s.})$	LO MG5	9644.5	LO
$W + \text{Jets} \rightarrow \ell \nu \ (2 \text{ jets f.s.})$	LO MG5	3144.5	LO
$W + \text{Jets} \rightarrow \ell \nu \ (3 \text{ jets f.s.})$	LO MG5	954.8	LO
$W + \text{Jets} \rightarrow \ell \nu $ (4 jets f.s.)	LO MG5	485.6	LO
Drell-Yan + Jets $\rightarrow \ell \ell (10 \text{ GeV} < m_{\ell \ell} < 50 \text{ GeV})$	LO MG5	18610	LO
Drell-Yan + Jets $\rightarrow \ell \ell \ (m_{\ell \ell} > 50 \text{ GeV}, \text{ inclusive})$	LO MG5	5765.4	LO
Drell-Yan + Jets $\rightarrow \ell\ell \ (m_{\ell\ell}50 \text{ GeV}, 1 \text{ jet f.s.})$	LO MG5	1012.5	LO
Drell-Yan + Jets $\rightarrow \ell \ell \ (m_{\ell \ell} > 50 \text{ GeV}, 2 \text{ jets f.s.})$	LO MG5	332.8	LO
Drell-Yan + Jets $\rightarrow \ell \ell \ (m_{\ell \ell} > 50 \text{ GeV}, 3 \text{ jets f.s.})$	LO MG5	101.8	LO
Drell-Yan + Jets $\rightarrow \ell \ell \ (m_{\ell \ell} > 50 \text{ GeV}, 4 \text{ jets f.s.})$	LO MG5	54.8	LO

<sup>1</sup> f.s. stands for "final state".

 $^2$  LO stands for Leading Order, NLO for Next to Leading Order.

<sup>3</sup> POWHEG is a NLO generator up to 1 jet in the f.s, and a LO one for 2 or more jets in the f.s..

The simulated events are weighted according to the cross section of the process  $\sigma$ , the integrated luminosity  $\mathcal{L}$  that corresponds to the data and the number of processed MC events  $N_{\text{processed}}$ . The weights are given by the following formula:

$$w_{\rm MC} = \frac{\sigma \cdot \mathcal{L}}{N_{\rm processed}}.\tag{4.1}$$

#### 4.2.3 Estimation of QCD Background: the ABCD method

It is very difficult to properly simulate the QCD background. In this analysis the QCD background is modeled in a data-driven way, using a control sample of events different from the one used to perform the measurement. Events for this sample are selected from the data used for the analysis, using the same selection imposed on events in the inclusive category (this selection will be described in section 4.3), with one fundamental exception: the lepton charge and the tau charge need to be of the same sign. The sample is referred as *same sign* (SS) sample, while the one selected by imposing opposite charge of the lepton-tau pair is called *opposite sign* (OS) sample. These two samples are by construction statistically independent.

There is no guarantee that the QCD SS sample has the same yield as the QCD OS one, therefore a correction for the normalization is implemented. To derive this correction we use an improved ABCD method. This method uses a pair of weakly correlated variables to estimate the yield of events in one of four regions, defined by imposing rectangular cuts on the two variables. The four regions, illustrated in figure 4.3, are defined as a function of the sign of the charges in the pair and the isolation of the lepton.

- A This is the signal region, where we want to estimate the QCD background.
- B This region corresponds to the SS sample described in the previous paragraph.
- C This region contains OS events with an inverted isolation cut  $I_{\rm rel} > 0.5$ , these events are called anti-isolated.
- D This region contains anti-isolated SS events.

The estimation for the QCD background in the signal region (A) is:

$$f_A(m) = f_B(m) \cdot \frac{N_C}{N_D},\tag{4.2}$$

where

- $f_A(m)$  and  $f_B(m)$  are the  $\tau \tau^2$  invariant mass distribution of events belonging to classes A and B respectively;
- $N_C$  and  $N_D$  are the number of events in the regions C and D.

<sup>&</sup>lt;sup>2</sup>In our case the  $\tau\tau$  system is  $\mu\tau_{had}$  or  $e\tau_{had}$ .



Figure 4.3: The four region defined in the ABCD method. From [83].

To get a more robust estimation of the ratio  $\frac{N_C}{N_D}$ , the latter is taken as the median of the three different ratios that can be obtained by inverting the isolation of the lepton only, the hadronic tau only, or both:

$$\frac{N_C}{N_D} = \text{median}\left(\left(\frac{N_C}{N_D}\right)_{\text{inv }\ell}, \left(\frac{N_C}{N_D}\right)_{\text{inv }\tau_{\text{had}}}, \left(\frac{N_C}{N_D}\right)_{\text{inv both}}\right).$$
(4.3)

The  $\frac{N_C}{N_D}$  scale factor value in this analysis is found to be  $\frac{N_C}{N_D} = 1.06$ .

# 4.3 Baseline Event Selection

The baseline selection for the  $\mu \tau_{had}$  ( $e \tau_{had}$ ) is formed by events that contain a leptontau pairs passing the baseline selection. The cuts applied on the lepton are summarized in table 4.2, those applied on the tau are summarized in table 4.3, and those applied on the event are summarized in table 4.4. These selection criteria are discussed in the following sections. Events passing these cuts are referred to as "inclusive category". In addition to the selection cuts, some correction factors are applied to simulated events. These corrections are needed to improve the description of the Monte Carlo samples when there are disagreements with data, as discussed in section 4.4.

Variable	Muon ( $\mu \tau_{had}$ ch.)	Electron $(e\tau_{\text{had}} \text{ ch.})$	
Impact parameter <sup>1</sup>	$ d_{xy}  < 0.45 \text{ mm and }  d_z  < 2 \text{ mm}$		
Transverse momentum Pseudorapidity Rel. Isolation Trigger to match	$\begin{array}{c} p_{\rm T} > 23 \ {\rm GeV} \\  \eta  < 2.4 \\ I_{\rm rel}^{\Delta R=0.4} < 0.15 \\ {\rm IsoMu22} \end{array}$	$p_{\rm T} > 26 {\rm ~GeV} \  \eta  < 2.1 \ I_{\rm rel}^{\Delta R=0.3} < 0.1 \ {\rm Ele25\_eta2p1} {\rm (Tight~WP)}$	
Identification Medium $ID^2$		NonTrig MVA Id (WP80) No missing hits in pixel detector PFlow electron conversion veto	
<sup>1</sup> IP, projected on the respective plane/axis, with respect to the primary vertex.			

Table 4.2: Lepton selection.

ŀ

 $^2$  As described in section 3.3.2.

Table 4.3: Hadronic tau selection.

Variable	Cut	
Transverse momentum Pseudorapidity Impact parameter <sup>3</sup>	$p_{ m T} > 20  { m GeV} \  \eta  < 2.3 \  d_z  < 2  { m mm}$	
Identification	By Decay Mode Finding	
$^{3}$ IP of the leading charged hadron candidate w.r.t. the PV.		

Table 4.4: Requirements for the event to be accepted in the inclusive category.

Variable	$\mu\tau_{\rm had}$ channel	$e\tau_{\rm had}$ channels	
Leptons in the event	Only one muon	Only one electron	
Charge	$q_{\ell} \cdot q_{\tau} = -1  (^5)$		
Tau isolation discriminator	MVA Isolation (Tight WP)		
Tau discriminator vs muon WP	Tight	Loose	
Tau discriminator vs electron WP	Loose	Tight	
$^{5}$ i.e. the lepton charge and the tau charge must have opposite signs.			

#### 4.3.1 Muon and Electron Cuts

The requirements applied to select leptons are reported in table 4.2. Selected muons (electron) should lie in the geometrical acceptance of the muon (electron) trigger  $|\eta| < 2.4$   $(|\eta| < 2.1)^3$ . The lepton transverse momentum must be at least 1 GeV larger than the trigger threshold ( $p_{\rm T} > 23$  GeV for muons,  $p_{\rm T} > 26$  GeV for electrons), in order to avoid systematic errors related to the trigger turning point efficiency modeling. The leptons are required to be matched to a related trigger object.

Real muons or electron, not directly produced in the physical event of interest, can emerge from the decay of hadrons. The main processes are  $\pi \to \mu \nu_{\mu}$  and  $K \to \mu \nu_{\mu}$  for muons and Dalitz and double Dalitz decays<sup>4</sup> for electrons. For electron another example is the conversion of photons into electron-positron pairs in the detector material. These leptons tend to originate far from the primary vertex, thus to improve the rejection of these kinds of leptons the cut on impact parameter is tightened with respect to the one in the identification criterias:  $|d_{xy}| < 0.45$  mm and  $|d_z| < 2$  mm. After these impact parameter requirements the main QCD sources of real muons and electrons are the semileptonic decays of beauty and charm quarks.

**Electron only cuts** Every electron track is required to have an associated hit in each of the traversed layers of the pixel detector in order to reduce the fraction of *fake tracks* and photon conversion. Fake tracks are reconstructed trajectories that do not correspond to any real particle. These are originated from randomly distributed hits produced by noise in the tracker electronics or the occurrence of nuclear interactions in the detector material.

#### 4.3.2 Tau Cuts

The requirements applied to select tau candidates are reported in table 4.3. The identification is done through the decay mode finding identification described in section 3.6.2. Additional discriminators, such as isolation or discriminator against lepton, are applied on an event basis and described in section 4.3.4.

## 4.3.3 Lepton-Tau Pairs Cuts

Pairs candidate are formed from the selected leptons and taus that passed the quality cuts described in 4.3.1 and 4.3.2. In order to constitute a pair candidate a lepton and a tau have to be closer than  $\Delta R_{\text{pair}} = 0.5$  in the  $\eta - \phi$  plane. In case of multiple possible candidates, one single pair for the event is chosen with the following criterias in descending order of importance: smaller value of the lepton isolation variable, lepton isolation, higher lepton  $p_T$ , smaller value of the tau isolation variable, higher tau  $p_T$ . The  $p_T$  and  $\eta$  distributions for both the channels studied in this thesis are shown in figure 4.4 and 4.5.

<sup>&</sup>lt;sup>3</sup>The silicon tracker geometrical acceptance is larger, with a pseudorapidity coverance of  $|\eta| < 2.5$ . <sup>4</sup>Dalitz decay:  $\pi^0 \to \gamma e^+ e^-$ . Double Daliz decay:  $\pi^0 \to e^+ e^- e^+ e^-$ .





(b) Hadronic tau  $p_T$  distribution in the  $\mu \tau_{had}$  channel.



Figure 4.4:  $p_T$  and  $\eta$  distributions for  $\mu \tau_{had}$  channel in the inclusive category. The points represent the data with statistical uncertainties; the background and signal contributions are shown as stacked histograms, normalized to the data luminosity and after all corrections. The QCD contribution is estimated from the same-sign data. The lower plots show the ratio between data and MC, while the band shows the systematic uncertainty on the MC contributions as described later in the text. All the corrections that will be discussed in section 4.4 are applied. These indications hold for all the distribution reported in this chapter.

nel.



Figure 4.5:  $p_T$  and  $\eta$  distributions for  $e\tau_{had}$  channel in the inclusive category.

#### 4.3. BASELINE EVENT SELECTION

#### 4.3.4 Event Cuts

The cuts applied on a event basis are summarized in table 4.4. Selected events for the  $\mu \tau_{had}$  ( $e \tau_{had}$ ) channel are required to have only one muon (electron) and no extra leptons, neither muon of electron. Tau discriminators against muon and electron<sup>5</sup> are required in order to reduce the background from  $Z \to \ell \ell$  events where one lepton mimics an hadronic tau signal. In order to reduce the background from QCD events the selected tau has to pass the MVA isolation discriminator<sup>6</sup>, the Tight Working Point is used.

The invariant visible mass of the lepton-tau pair is shown in figure 4.6 for both channels. Due to the neutrinos, the visible mass peaks to lower values than the actual  $Z^0$  mass (about 91 GeV). The Missing Transverse Energy, computed using the MVA MET method, is shown in figure 4.7 for both channels.

The data are well reproduced by the MC predictions inside the uncertainties in all distributions. At this point the main contributions are the  $Z \to \tau \tau$  signal and the W+Jets background contribution. The peak structure around 90 GeV (blue histogram) in the visible mass is due to  $Z \to ee \ (\mu\mu)$  events where one electron (muon) is misidentified as hadronic tau. These events are "leftovers" of the tau discriminator against electron (muon) (see next section for a discussion on the impact of each selection criteria).



Figure 4.6:  $m_{\rm vis}$  distributions in the inclusive category.

 $<sup>^{5}</sup>$ As described in section 3.6.4 and 3.6.5.

<sup>&</sup>lt;sup>6</sup>As described in section 3.6.3.



Figure 4.7: MET distributions in the inclusive category.

#### 4.3.4.1 Sequence of the selection criteria

In order to qualitatively understand how each cut reduces the background a cutflow for the  $\mu \tau_{had}$  is reported in table 4.5 and figure 4.8. The QCD and W+jets backgrounds are mainly reduced by the cut on the tau isolation, respectively of a factor 60 and 30. After the isolation cut the  $Z \rightarrow \ell \ell$  ( $\ell = e, \mu$ ) background events composes over the 95% of the selected events. This background is greatly reduced, by a factor of 400, by the tau fake discriminators. Diboson and W+jets background are reduced by both cut levels in a similar magnitude. W+jets is further reduced by a cut in the transverse mass. This cut will be discussed separately in section 4.5 since the inclusive category is defined without cut on the transverse mass. The results for the  $e\tau_{had}$  channel are very similar.

# 4.4 Corrections

A number of correction factors are applied to the Monte Carlo samples in order to improve the agreement with the data. The corrections described in this section address the differences in the pile-up distributions, the differences in the muon and electron identification efficiencies and the fact that we do not require the trigger matching in the simulated samples<sup>7</sup>.

<sup>&</sup>lt;sup>7</sup>The triggers were not simulated at the time of this analysis.
### 4.4. CORRECTIONS

Table 4.5: Cutflow for the different background sources and for the signal for  $\mu \tau_{had}$  channel. The number of events are normalized to the data luminosity. The cut levels are defined as: (1) opposite sign pair selection; (2) tau isolation discriminator; (3) veto for extra leptons and tau discriminators against fake rates.

Cut level	QCD	di-boson	W+jets	t $\bar{t}$ +single top	$Z\ell\ell$	$Z\tau\tau$ (signal)
1	2010370	64632	3146140	873733	17070300	152436
1 + 2	33844	16530	110129	101679	8634210	81977
1 + 2 + 3	25005	2409	92794	21935	21669	71744



(a) Number of events from each different process at each cut level.



(b) Relative composition of the selected events. (c) Relative composition of the selected events. Logarithmic scale.

Figure 4.8: Number of events from each different process in the selected  $\mu \tau_{had}$  events at each cut level (defined in tab. 4.5).

### 4.4.1 Pile-Up Correction

The distribution in the number of pile-up vertices in the Monte Carlo samples is different with respect to the one observed in data. Production Monte Carlo samples are generated with distributions for the number of pileup interactions which is meant to roughly cover, though not exactly match the conditions expected for each data-taking period. Thus an event-by-event correction is required to improve the agreement.

The MC pileup distribution is reweighed to the data pileup distribution, derived from the instantaneous luminosity measured in each lumi section, over a certain period of data-taking. The MC event weight is calculated as

$$w_{\rm PU} = \frac{\rm PU_{data}(nPU_{event})}{\rm PU_{MC}(nPU_{event})}$$
(4.4)

where  $nPU_{event}$  is the number of Pile-Up interactions of the event to be reweighted,  $PU_{MC}(nPU_{event})$  is the value of the normalized<sup>8</sup> PU distribution of the simulation at  $nPU_{event}$ , and  $PU_{data}(nPU_{event})$  is the value of the normalized PU distribution of the data at  $nPU_{event}$ . The spectrum of the number of Primary Vertex distribution after the correction is shown in figure 4.9. The figure shows that on average ~ 15 primary vertices are expected in these data, with a tail up to 40. The agreement data-MC is reasonable inside the uncertainties, which assume an uncertaintly of  $\pm 5\%$  on the minimum bias cross section for the pile-up interactions.



Figure 4.9: Number of Primary Vertex distribution for  $\mu \tau_{had}$  in the inclusive category.

<sup>8</sup>Normalized to one.

### 4.4.2 Identification and Isolation Efficiencies

The efficiency for electrons, muons and hadronic tau to pass the identification and isolation cuts (Id&Iso) may not be modeled properly by the simulation. For this reason a weight is applied to each simulated event as a function of the particle  $\eta$  and  $p_T$ , in order to bring the simulation to agreement with the data. The weight is expressed as

$$w_{\mathrm{IdIso}\,\ell}(p_T,\eta) = \frac{\epsilon_{\mathrm{IdIso}\,\ell}^{\mathrm{data}}(p_T,\eta)}{\epsilon_{\mathrm{IdIso}\,\ell}^{\mathrm{MC}}(p_T,\eta)}, \ (\ell = \mu, \, e, \, \tau_{\mathrm{had}})$$
(4.5)

where the efficiencies  $\epsilon$ , binned in  $p_T$  and  $\eta$ , are estimated with the Tag And Probe method. This weight is often called *Scale Factor* (SF).

The measurement of the identification and isolation efficiencies for muon and electron is a major part of my thesis work and my personal contribution to the whole analysis group, and it will be described in detail in chapter 5. The measured  $\tau_{had}$  SF is found to be constant in  $p_T$  and  $\eta$  and has a value of 0.83 with 6% uncertainty.

### 4.4.3 Trigger Efficiency

The trigger response is not simulated in the Monte Carlo samples used in this analysis. On the other hand in the real data we require that the selected leptons match the respective HLT, as stated in section 4.1. Since triggers do not have 100% efficiency a correction is needed to account this "extra" selection present only in the data. The trigger response usually is strongly  $p_T$  and  $\eta$  dependent, a typical efficiency response curve for a trigger is shown in figure 4.10, the *turning point* of the threshold at  $p_T \sim 22$ GeV is clearly visible.

The efficiency curves for the triggers used in this analysis are measured with the same Tag And Probe method used for the Id&Iso efficiencies, with the exception that, since only data efficiencies are available, these are applied directly and not in the form of a scale factor. The weight applied to each MC event is

$$w_{\rm trig}(p_T, \eta) = \epsilon_{\rm trig}^{\rm data}(p_T, \eta). \tag{4.6}$$

The measurement of the triggers efficiencies will be described in detail in chapter 5.

## 4.5 Cut on the Transverse Mass

The transverse mass for a lepton is a physical observable defined as

Looking at the spectra in figure 4.11 one can observe that the  $Z \to \tau \tau$  signal and the W+jets background have very different distributions: the W+jets background peaks around the W mass (~ 70 - 80 GeV), while the signal is concentrated mostly at lower  $m_T$  values and clustered around zero.



Figure 4.10: IsoMu22 trigger efficiency for selected  $\eta$  bins as a function of the muon  $p_T$ .

We can then use a cut on the value of the transverse mass as discriminator to reduce the W+jets background. The value of the cut is chosen as the one that maximizes the significance of the signal defined as

significance = 
$$\frac{S}{\sqrt{S+B}}$$
 (4.8)

where S and B stand respectively for the number of signal and background events in the 36-84 GeV visible mass region of the Z peak. The significance as a function of the  $m_T$  cut is shown in figure 4.12 and an optimal cut value for both channel is  $m_T < 50$ GeV. The spectrum of the visible mass before and after the cut in  $m_T$  is presented in figures 4.13 and 4.14. The reduction of the W+jets background is clearly visible in both channels. The other distributions with the  $m_T$  cut applied are reported in appendix A.



Figure 4.11:  $m_T$  distributions for both channels in the inclusive category.



Figure 4.12: Significance of the signal in the 36-84 GeV visible mass region as a function of the cut in  $m_T$ . The maximum is at a  $m_T^{\text{cut}} = 50$  GeV for both channel.



Figure 4.13: Visible mass distribution spectrum for  $\mu \tau_{had}$  before and after the  $m_T < 50$  GeV cut.



Figure 4.14: Visible mass distribution spectrum for  $e\tau_{had}$  before and after the  $m_T < 50$  GeV cut.

# Chapter 5

## Lepton ID and Trigger Efficiencies for $Z \rightarrow \tau \tau$ Measurement

In this chapter the measurement of the lepton<sup>1</sup> Id&Iso Scale Factor (SF) and of the Triggers efficiencies, needed to improve the agreement between simulation and data as described in section 4.4.2, is presented. In order to measure these efficiencies a *Tag* and Probe workflow, within the CMS software environment, has been developed. The efficiency scale factors are derived for the selection criteria reported in table 5.1. The efficiencies have been calculated using  $Z \to \ell \ell$ , ( $\ell = e, \mu$ ) events in the same data-taking period as the analysis of this thesis and Drell-Yan + Jets  $\to \ell \ell$  MC samples.

### 5.1 Tag and Probe Method

The Tag and Probe (T&P) method allows the nearly unbiased measurement of any user defined object efficiency from data (or simulation) by exploiting leptons resonances like  $Z^0$  or  $J/\Psi$ . The efficiency is represented by the fraction of particles, in our case electrons or muons, passing a given selection, in our case the identification criteria and the isolation cut or the matching with a trigger object.

In order to avoid bias on the set of particle candidates the measurement is performed on pairs of identical particles, in our case muons or electrons, emerging from the decay of well known massive particles, in our case the  $Z^0$  boson. A tight selection is imposed on one (and only one) of the leptons in the pair, the "tag" lepton. This tight selection ensures an acceptable  $Z^0$  signal to background ratio in the measurement dataset. The other lepton in the pair, called "probe", is subject to loose selection criteria and it is the one on which the efficiency measurement is performed. The loose selection is used to keep the bias to the measurement small. The detailed description of the T&P steps is reported in the next sections.

<sup>&</sup>lt;sup>1</sup>Hereafter "lepton" stands for "electron" or "muon".

Object type	Electron	Muon
Identification	NonTrig MVA Id (WP80) No missing hits in pixel detector PFlow electron conversion veto	Medium ID
Isolation	$I_{\rm rel}^{\Delta R=0.3} < 0.10$	$I_{\rm rel}^{\Delta R=0.4} < 0.15$
Trigger	Ele25_eta2p1 (Tight WP)	IsoMu22

Table 5.1: List of the selection criteria which efficiency has been measured.

#### 5.1.1 Selection of Tag&Probe Pairs

The first step of the T&P method is building the candidate collections of tags and probes. The selection criteria for the tags and the probes are reported in table 5.2 (for  $Z \rightarrow \mu\mu$ ) and 5.3 (for  $Z \rightarrow ee$ ). The selection of the tags is very similar to the one of the leptons in the baseline selection, already discussed in section 4.3. The probes candidates have to pass very loose kinematic and impact parameter cuts; neither trigger matching, isolation cut or passing the identification criteria is required.

The next step is building, on event basis, the T&P pairs from the collection of tags and probes. The pair requirements, reported in table 5.4, are chosen in order to select pairs from  $Z^0$  decay: the two leptons must be well separated and have opposite sign charges, and the pair is required to have a reconstructed visible mass higher than 50 GeV. The pairs are formed in combinatorial fashion, i.e. the collection is formed by all the possible combinations of a tag and a probe that pass the pair requirements<sup>2</sup>, in a given event.

**Trigger efficiency T&P pairs** Since we want to factorize the selection efficiency into Id&Iso efficiency and trigger efficiency, the collection of probe for trigger efficiency measurements has the additional requirement that the probes must pass the Id&Iso requirements. The tag and pair requirements remain the same as described above.

**Passing and failing probes** A probe is labeled as "passing probe" for a given object if it passes the corresponding selection criteria (reported in table 5.1), otherwise it is labeled as "failing probe". For Id&Iso efficiency measurement a passing probe probe must pass the identification criteria and have an isolation lower than the isolation cut. For trigger efficiency measurement a passing probe must be correctly matched to the trigger to be evaluated<sup>3</sup>.

<sup>&</sup>lt;sup>2</sup>—Since all the tags pass the probe requirements if more than one tags are present in the event tag-tag pairs are selected. This is intentional.

<sup>&</sup>lt;sup>3</sup>The trigger matching was described in section 4.1. In brief it means that that the trigger object must lie in a  $\Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2} = 0.5$  cone centered on the direction of flight of the lepton.

Variable	Tag	Probe
Transverse momentum	$p_{\rm T} > 23~{\rm GeV}$	$p_{\rm T} > 10 {\rm ~GeV}$
Pseudorapidity	$ \eta  < 2.1$	$ \eta  < 2.4$
$d_{xy}$ impact parameter	$ d_{xy}  < 0.45 \text{ mm}$	$ d_{xy}  < 2 \text{ mm}$
$d_z$ impact parameter	$ d_z  < 2 \text{ mm}$	$ d_z  < 5 \text{ mm}$
Trigger to match	IsoMu22	/
Rel. Isolation	$I_{\rm rel}^{\Delta R=0.4} < 0.15$	/
Identification	Medium ID	/

Table 5.2: Tags and probes selection criteria for muon  $(Z \rightarrow \mu \mu)$ . The *slash* stands for no requirement on the given variable.

Table 5.3: Tags and probes selection criteria for electron  $(Z \rightarrow ee)$ . The *slash* stands for no requirement on the given variable.

Variable	Tag	Probe
Transverse momentum Pseudorapidity $d_{xy}$ impact parameter $d_z$ impact parameter Trigger to match Rel. Isolation	$p_{\rm T} > 28 \; { m GeV} \  \eta  < 2.1 \  d_{xy}  < 0.45 \; { m mm} \  d_z  < 2 \; { m mm} \ { m Ele25\_eta2p1} \; ({ m Tight WP}) \ I_{ m rel}^{\Delta R=0.3} < 0.10$	$\begin{array}{c} p_{\rm T} > 10 \ {\rm GeV} \\  \eta  < 2.5 \\  d_{xy}  < 2 \ {\rm mm} \\  d_z  < 5 \ {\rm mm} \\ \hline \end{array}$
Identification	NonTrig MVA Id (WP80) No missing hits in pixel detector PFlow electron conversion veto	/

Table 5.4: Selection criteria for the tag and probe pairs for both channel.

Variable	$\ell\ell$ -pair
Charges signs Distance in the $n - \phi$ plane	opposite $\Delta R > 0.5$
Visible mass	$m_{\rm vis} > 50 {\rm ~GeV}$

### 5.1.2 Extraction of Efficiencies

The efficiency is measured as a function of  $p_T$  and  $\eta$ 

$$\epsilon(p_T, \eta) = \frac{N_{\text{passing}}(p_T, \eta)}{N_{\text{passing}}(p_T, \eta) + N_{\text{failing}}(p_T, \eta)},$$
(5.1)

where the number of passing and failing probes are derived from the integral of the dilepton mass distribution in the interval 80-102 GeV around the Z mass. The pseudorapidity regions are divided in the ranges:

- muon  $|\eta| \in [0 0.9], [0.9 1.2], [1.2 2.1], [2.1 2.4];$
- electron  $|\eta| \in [0 1.48], [1.48 2.1], [2.1 2.5].$

The efficiencies are derived in data and MC and the MC efficiency is then rescaled to the data efficiency using the scale factors:

$$\mathrm{SF}_{\mathrm{IdIso}\,\ell}(p_T,\eta) = \frac{\epsilon_{\mathrm{IdIso}\,\ell}^{\mathrm{data}}(p_T,\eta)}{\epsilon_{\mathrm{IdIso}\,\ell}^{\mathrm{MC}}(p_T,\eta)}, \ (\ell=\mu,e).$$
(5.2)

#### 5.1.2.1 Fitting Model

The invariant mass distributions for  $Z \to \mu\mu$  for passing and failing probes, in selected  $p_T$  ranges, are shown in figure 5.1 and 5.2. The equivalent distributions for  $Z \to ee$ are reported in appendix B. For the fitting the signal is modeled as 2 asymmetrical gaussian functions with the same mean, while the background is modeled with an exponential function. An FSR component  $(Z \to \ell \ell + \gamma)$ , modeled as a gaussian function with left shifted central value, is optionally added at the failing probes signal for the Id&Iso efficiency measurement. This component is added in the following kinematic regions, as it better describes the signal shape:  $25 < p_T(\text{GeV}) < 40$  for muons and  $25 < p_T(\text{GeV}) < 40$ GeV for electrons, independently on the pseudorapidity region. Examples of invariant mass distributions in these  $p_T$  ranges with FSR fitting are shown in figure 5.3 and 5.4.

Overall the  $Z^0$  peak is well described by the fitting model in all  $p_T$  and pseudorapidity regions.





Figure 5.1:  $Z \rightarrow \mu\mu$  invariant mass distribution for Muon Id&Iso efficiency measurement for data, with  $1.2 < |\eta| < 2.1$ and  $10 < p_T(\text{GeV}) < 15$ . The red line represent the signal, the dashed blue line the background.

Figure 5.2:  $Z \rightarrow \mu\mu$  invariant mass distribution for Muon Id&Iso efficiency measurement for data, with  $1.2 < |\eta| < 2.1$ and  $40 < p_T(\text{GeV}) < 60$ . The red line represent the signal, the dashed blue line the background.





Figure 5.3:  $Z \rightarrow \mu\mu$  invariant mass distribution for Muon Id&Iso efficiency measurement for data, with  $1.2 < |\eta| < 2.1$ and  $30 < p_T(\text{GeV}) < 40$ . The red line represent the signal with the FSR component, the dashed blue line the background.

Figure 5.4:  $Z \rightarrow ee$  invariant mass distribution for Electron Id&Iso efficiency measurement for data, with  $1.48 < |\eta| < 2.5$  and  $20 < p_T(\text{GeV}) < 25$ . The red line represent the signal with the FSR component, the dashed blue line the background.

## 5.2 Results

### 5.2.1 Identification and Isolation Scale Factor

**Muon** The measured muon identification and isolation efficiency, with relative scale factor, is shown in figure 5.5 for  $p_T^{\mu} < 60$  GeV. Scale factors range from 0.92 to 0.98, indicating a good modeling of the muon Id&Iso objects in the simulation. The difference between data and simulation tends to diminish with increasing momentum.

**Electron** The measured electron identification and isolation efficiency, with relative scale factor, is shown in figure 5.6 for  $p_T^e < 60$  GeV. Scale factors range from 0.80 to 0.98, with efficiency discrepancy up to 20%. Once again difference between data and simulation tends to diminish with increasing momentum.

### 5.2.2 Trigger Efficiency

**IsoMu22** The measured trigger efficiency is shown in figure 5.7 for  $p_T^{\mu} < 100$  GeV. Besides the barrel region ( $|\eta| > 2.1$ ), the efficiency reaches the plateau just few GeV above the  $p_T = 22$  GeV threshold, justifying the  $p_T$  cut in the baseline selection. The plateau efficiency values range from 0.8 to 0.9 and decrease with increasing pseudorapidity.

**Ele25\_eta2p1 Tight WP** The measured trigger efficiency is shown in figure 5.8 for  $p_T^e < 100$  GeV. The trigger shows a slow turning point, with 50% efficiency at the  $p_T = 25$  GeV threshold and the plateau at  $p_T > 50$  GeV. The efficiency is higher (about 10%) in the barrel compared to the endcaps, and drops for  $|\eta| > 2.1$ , as expected<sup>4</sup>.

### 5.2.3 Validation of the Efficiency Corrections

In figure 5.9 ( $\mu \tau_{had}$  ch.) and 5.10 ( $e \tau_{had}$  ch.) the lepton  $p_T$  and  $m_{vis}$  distributions are show before and after applying the Id&Iso SF and trigger efficiency corrections<sup>5</sup>. A great improvement in the agreement between data and simulation is evident, with a normalized  $\chi^2$  smaller than 1 after the corrections.

<sup>&</sup>lt;sup>4</sup>We remember that this trigger has a  $|\eta| < 2.1$  cut.

<sup>&</sup>lt;sup>5</sup>As defined in equation 4.5 and 4.6 in section 4.4.



Figure 5.5: Muon identification and isolation efficiency as function of  $p_T$  in different pseudorapidity regions. Drell-Yan + Jets  $\rightarrow \ell \ell$  ( $\ell = e, \mu, \tau; m_{\ell \ell} > 50 \text{ GeV}$ ) constitutes the MC events. The scale factor is calculated as the ratio between data and MC efficiencies in a given  $p_T$  bin.



Figure 5.6: Electron identification and isolation efficiency as function of  $p_T$  in different pseudorapidity regions. Drell-Yan + Jets  $\rightarrow \ell \ell$  ( $\ell = e, \mu, \tau; m_{\ell \ell} > 50$  GeV) constitutes the MC events. The scale factor is calculated as the ratio between data and MC efficiencies in a given  $p_T$  bin.



Figure 5.7: IsoMu22 HLT efficiency as function of  $p_T$  in different pseudorapidity regions.



Figure 5.8: Ele25\_eta2p1 (*Tight* WP) HLT efficiency as function of  $p_T$  in different pseudorapidity regions.

GeV 40000

m 35000

Stepts 25000

20000

15000

10000

5000

1.0

0.

0

20

data/MC 1.5



(a) Muon  $p_T$  distribution before applying efficiency corrections.



(c)  $m_{\rm vis}$  distribution before applying efficiency corrections.

(b) Muon  $p_T$  distribution after applying efficiency corrections.

60

80

40

CMS Private Work µt ch.

12.9 fb<sup>-1</sup> (13 TeV)

+ Data

Utt+single top

dibosons DOCD

χ<sup>2</sup>/d.o.f. = 0.24

100 120 pT<sub>µ</sub> [GeV]



(d)  $m_{\rm vis}$  distribution after applying efficiency corrections.

Figure 5.9:  $\mu \tau_{had}$  channel selected distributions before and after applying muon Id&Iso trigger efficiency corrections. PileUp and  $\tau_{had}$  Id&Iso corrections are applied in all distributions.



(a) Electron  $p_T$  distribution before applying efficiency corrections.



(c)  $m_{\rm vis}$  distribution before applying efficiency corrections.



60

80

100

120 pT [GeV]

40

CMS Private Work et ch.

16000

14000

Events / 12000

8000

6000

4000

2000

data/MC 1.5

0

1.0

0.5

0

20

3 GeV



(d)  $m_{\rm vis}$  distribution after applying efficiency corrections.

Figure 5.10:  $e\tau_{had}$  channel selected distributions before and after applying electron Id&Iso trigger efficiency corrections. PileUp and  $\tau_{had}$  Id&Iso corrections are applied in all distributions.

12.9 fb<sup>-1</sup> (13 TeV)

Data

Utt+single

diboson:

χ<sup>2</sup>/d.o.l. = 0.31

## 5.3 Evaluation of the Systematic Error

The systematic error on the measurement of the scale factor has been estimated varying the integration range of the  $Z^0$  peak, the results are shown in figure 5.11. For the electron scale factor the maximum relative variation, with respect the the nominal one<sup>6</sup>, is 5% only in one bin. The mean variation is about ~ 1 - 2% at low  $p_T$  and at subpercent level difference at high  $p_T$ . For the muon scale factor the maximum relative variation, with respect the the nominal one, is 2.5% only in one bin and less than 1% elsewhere.



Figure 5.11: Id&Iso scale factor measurement varying the integration range of the  $Z^0$  peak in different pseudorapidity regions. Different colors represent different integration ranges (legend on top). The SFs are functions of  $p_T$  (scale not shown).

<sup>&</sup>lt;sup>6</sup>The nominal scale factor is the one measured with a [80-102] GeV integration range.

## CHAPTER 6

Z-Boson Cross Section Measurement in the  $Z \rightarrow \tau \tau \rightarrow \mu \tau_h$  and  $Z \rightarrow \tau \tau \rightarrow e \tau_h$  Decay Channels

### 6.1 Introduction

In this chapter the measurement of  $Z \to \tau \tau$  cross section<sup>1</sup> in the  $\tau \tau \to \mu \tau_{had}$  and  $\tau \tau \to e \tau_{had}$  channel is described.  $Z \to \tau \tau$  plays an important role in the measurement of the Standard Model Higgs boson in the decay into a pair of  $\tau$  leptons and in the LHC program as:

- It provides an experimental benchmark for the commissioning of the physics objects described in chapter 3, the analysis described in chapter 4 and the efficiency corrections measured in chapter 5.
- It constitutes a major source of irreducible background to the search for neutral Higgs bosons.
- It provides a number of tests to the Standard Model such as the decay properties of the τ lepton, predictions for the Z boson production cross section in proton-proton collisions, and the lepton universality.

## 6.2 Systematic Uncertainties

Various sources of systematic uncertainties can affect the number of events counted in each bin of the mass histogram. The systematic uncertainties taken into account in the Z production cross section measurement are summarized in table 6.1 and described hereafter.

<sup>&</sup>lt;sup>1</sup>What is measured is  $\sigma(pp \to Z) \cdot Br(Z \to \tau\tau)$ .

Source	Value	Correlation
Luminosity	6.2%	Yes
Tau ID efficiency	6%	Yes
Lepton Id&Iso SF	2%	No
Lepton trigger efficiency	2%	No
Tau energy scale	3%	Yes
Muon energy scale	$\sim 0\%$	No
Electron energy scale	$\sim 0\%$	No
$E_T$ energy scale	2%	Yes
QCD norm. ( $\mu \tau_{had}$ ch.)	4%	No
QCD norm. $(e\tau_{had} ch.)$	12%	No
Diboson norm.	5%	Yes
W+Jets norm.	10%	No
$t\bar{t} + \text{single top norm.}$	6%	Yes
$Z\ell\ell$ norm.	30%	No
$Z\tau\tau$ norm.	4%	Yes

Table 6.1: Values for the systematic uncertainty taken into account. The last column reports if the given uncertainty is correlated between the two decay channels investigated.

**Luminosity** The integrated luminosity  $\mathcal{L}$  is a multiplicative term used to normalize all Monte Carlo samples, as stated in equation 4.1, and therefore has a profound effect on the matching between data and simulation. CMS can measure the instantaneous luminosity at the collision point using various kinds of detectors and techniques. The currently used one combines the measurement of the instantaneous visible event rate<sup>2</sup> with the physical measurement of the beam size in the transverse plane<sup>3</sup>. The relative uncertainty on the total integrated luminosity has been measured to be 6.2% [84]. This affects in a coherent way all the background sources estimated from simulation.

Efficiency corrections As described in section 4.4.2, MC simulation is corrected for measured differences in ID and isolation efficiency in data and MC simulation as a function of  $p_T$  and  $\eta$ . Additionally the efficiencies of the trigger measured in data are applied to the MC simulation. The systematic uncertainties on the lepton Id&Iso scale factor and trigger efficiency are found to be around 2% (see section 5.3). These values are applied as a systematic to the signal and all simulated backgrounds. Similarly the uncertainty on the tau ID efficiency measurement is 6%.

**Energy scale** The uncertainty in the muon and electron energy scale has a negligible effect for both the channels studied in this thesis, while the uncertainty in the tau energy

<sup>&</sup>lt;sup>2</sup>Obtained through the measurement of the occupancy in the pixel detector at each bunch crossing. <sup>3</sup>This measurement is performed through a *Van der Meer scan*.

### 6.3. CROSS SECTION MEASUREMENT

scale is assigned a value of 3%. An uncertainty on the resolution and response of  $E_T$ scale uncertainty is derived in studies of the hadronic recoil in  $Z \to \mu\mu$  events in data and MC simulation. This amounts to a 2% normalization uncertainty and affects the event yields for all the samples due to the  $m_T$  selection.

**Simulated sample normalization** A number of systematic uncertainties, which affect the normalization of a given process, are considered [85]:

- Cross section These uncertainties are due to uncertainties in the theoretical calculations, such as uncertainties in the parton distribution functions (PDF), variations of the renormalization and factorization scales, and uncertainties in the modelling of the underlying event and parton showers, and in part to the measurement. The magnitude of the uncertainty depends on the production process. The uncertainty on the normalization of the  $Z/\gamma^* \to \tau\tau$  and  $Z/\gamma^* \to \ell\ell$  amounts to 4% due to the theoretical uncertainty on the NNLO cross-section. Similarly the uncertainty on the diboson and single top contribution is 5%. For the  $t\bar{t}$  background the crosssection uncertainty is 6%. A systematic uncertainty of 10% affects the W+Jets normalization.
- Fake rate An uncertainty of 30% on the fake rate<sup>4</sup> measurement is applied to the  $Z/\gamma^* \to \ell\ell$  background.
- QCD OS/SS ratio A 4% (12%) systematic uncertainty on the QCD OS/SS ratio (see section 4.2.3) affects the QCD normalization affect in the  $\mu \tau_{had}$  ( $e \tau_{had}$ ) channel.

## 6.3 Cross Section Measurement

After the selection procedure described in chapter 4, and with the additional transverse mass cut described in section 4.5, the final event sample is determined. The cross section is measured evaluating the ratio between the measured and simulated signal in the  $m_{\rm vis} \in [44, 88]$  GeV visible mass region. The visible mass distribution for both channel is shown in figure 6.1. Defining  $\sigma(Z \to \tau \tau) = \sigma(pp \to Z) \cdot \text{Br}(Z \to \tau \tau)$ , the topological cross section is measured as

$$\sigma^{\text{data}}(Z \to \tau\tau, \, m_{\ell\ell} > 50 \,\text{GeV}) = \frac{\sigma^{\text{th}}(pp \to Z \to \ell\ell)}{3} \cdot \frac{N_{\text{data}} - N_{\text{bkg}}}{N_{\text{MC}}}, \qquad (6.1)$$

where

•  $\sigma^{\text{th}}(Z \to \ell \ell, m_{\ell \ell} > 50 \text{ GeV}) = 5.7654 \pm 0.032 \text{ nb}$  is the theoretical NNLO prediction used for the generation of the simulated DY events  $Z/\gamma^* \to \ell \ell (\ell = e, \mu, \tau)$  with  $m_{\ell \ell} > 50 \text{ GeV}^5$ ,

<sup>&</sup>lt;sup>4</sup>Rate at which electrons and muons are reconstructed as taus.

<sup>&</sup>lt;sup>5</sup>The reconstructed visible mass is lower than 50 GeV due to the neutrinos escaping detection.



Figure 6.1:  $m_{\rm vis}$  distribution in the inclusive category with the additional  $m_T < 50 \text{ GeV}$  cut (as described in section 4.5) for both  $\mu \tau_{\rm had}$  and  $e \tau_{\rm had}$  channels. The colored region between the vertical two red lines shows the integration region  $m_{\rm vis} \in [44, 88]$  GeV.

- the <sup>1</sup>/<sub>3</sub> factor comes from  $\operatorname{Br}(Z \to \ell \ell) = \operatorname{Br}(Z \to ee) + \operatorname{Br}(Z \to \mu \mu) + \operatorname{Br}(Z \to \tau \tau) \simeq 3 \cdot \operatorname{Br}(Z \to \tau \tau)$ , with  $\operatorname{Br}(Z \to ee) \simeq \operatorname{Br}(Z \to \mu \mu) \simeq \operatorname{Br}(Z \to \tau \tau)^6$ ,
- $N_{\text{data}}$  is the number of events counted in the  $m_{\text{vis}} \in [44, 88]$  GeV visible mass region,
- $N_{\text{bkg}} = N_{Z\ell\ell} + N_{t\bar{t}+t} + N_{W+j} + N_{diboson} + N_{QCD}$  is the number of estimated background events counted in the  $m_{\text{vis}} \in [44, 88]$  GeV visible mass range,
- $N_{\rm MC} = N_{Z\tau\tau}$  is the number of simulated  $Z \to \tau\tau$  signal events counted in the  $m_{\rm vis} \in [44, 88]$  GeV visible mass region.

Using equation 6.1 we find the following Z production cross section:

$$\begin{split} &\sigma_{\mu\tau_{\rm had}}^{\rm data}(Z \to \tau\tau, \, m_{\ell\ell} > 50\,{\rm GeV}) = 1.88 \pm 0.02\,({\rm stat}) \,\pm 0.26\,({\rm syst}) \,\pm 0.17\,({\rm lumi})\,[{\rm nb}], \\ &\sigma_{e\tau_{\rm had}}^{\rm data}(Z \to \tau\tau, \, m_{\ell\ell} > 50\,{\rm GeV}) = 1.82 \pm 0.03\,({\rm stat}) \,\pm 0.33\,({\rm syst}) \,\pm 0.19\,({\rm lumi})\,[{\rm nb}], \end{split}$$

respectively for the  $\mu \tau_{had}$  channel and  $e \tau_{had}$ . The measured values are compatible<sup>7</sup> with each other and with the NNLO prediction,  $1.922 \pm 0.011$  nb [87, 88].

<sup>&</sup>lt;sup>6</sup>From lepton universality and the negligible differences in the phase space. Measured values: Br( $Z \rightarrow ee$ ) = (3.363 ± 0.004)%, Br( $Z \rightarrow \mu\mu$ ) = (3.366 ± 0.007)%, Br( $Z \rightarrow \tau\tau$ ) = (3.370 ± 0.008)% [86]

<sup>&</sup>lt;sup>7</sup>Correlations between systematic uncertainties have been taken in account.

## Summary and Outlook

This thesis presents the study of the  $Z \rightarrow \tau \tau$  decay, where one  $\tau$ -lepton decays hadronically and the other into a lighter charged lepton, namely an electron or a muon. The analysis was performed using the data collected between March and August 2016 with the CMS experiment, corresponding to an integrated luminosity of 12.9 fb<sup>-1</sup> at 13 TeV. We measured the Z boson production cross section in proton-proton collisions to be:

$$\begin{aligned} \sigma_{\mu\tau_{\rm had}}^{\rm data}(Z \to \tau\tau, \ m_{\ell\ell} > 50 \,{\rm GeV}) &= 1.88 \pm 0.02 \,({\rm stat}) \pm 0.26 \,({\rm syst}) \pm 0.17 \,({\rm lumi}) \,[{\rm nb}], \\ \sigma_{\ell\tau_{\rm had}}^{\rm data}(Z \to \tau\tau, \ m_{\ell\ell} > 50 \,{\rm GeV}) &= 1.82 \pm 0.03 \,({\rm stat}) \pm 0.33 \,({\rm syst}) \pm 0.19 \,({\rm lumi}) \,[{\rm nb}]. \end{aligned}$$

The measurements are consistent with each other, but also in agreement with theoretical NNLO prediction.

A number of corrections were applied to the MC samples in order to account for the efficiency of Identification, Isolation and High Level Trigger objects. One of the task of the  $H \rightarrow \tau \tau$  CMS group at DESY is the evaluation of these corrections for muon and electron objects in the  $\tau \tau$  analyses context. In the work presented in this thesis a Tag and Probe workflow has been developed in order to measure the abovementioned corrections. The workflow is currently used in the research group at DESY and the results are shared with the rest of the CMS Collaboration involved in the search of the SM Higgs Boson, or a MSSM Higgs Boson, decay into a pair of  $\tau$ -lepton.

The second run of the LHC has already provided an integrated luminosity of 41.07 fb<sup>-1</sup> [35] and is expected to provide over 100 fb<sup>-1</sup> by the end of the run in 2018. This outstanding performance of the LHC and the CMS detector will allow the precise measurement of the SM Higgs Boson coupling to  $\tau$ -lepton, investigating any deviation from the Standard Model predictions.

## SUMMARY AND OUTLOOK

### Acknowledgments

This thesis would not have been possible without the help and the support of many people, I owe them my sincere gratitude.

I want to thank my DESY supervisor, Prof. Dr. Elisabetta Gallo, for her patient guidance through my six months traineeship at DESY. Her support has proved essential not only in the actual work, but also in enjoying Hamburg and in going through all the traineeship paperwork. I am also very grateful to my internal supervisor at the University of Padova, Prof. Franco Simonetto, for his invaluable support in reviewing this thesis during these last months.

My heartfelt thanks go to Dr. Alexei Raspereza, Dr. Francesco Costanza and Valeria Botta, for the all the time they spent supporting, helping and teaching me. I would like also to thank the rest of the *CMS Higgs in Tau Tau* group at DESY, it was great to work with you: Dr. Alexis Kalogeropoulos, Dr. Rainer Mankel, Vallary Bhopatkar, Illya Bobovnikov and Yiwen Wen.

I am deeply indebted to thank my family, in particular my parents, Carla and Renato, for their continuous moral and financial support during my studies in Padua and traineeship at DESY, and for all the opportunities they gave me throughout life. I could not have done it without them. Last but not least, I would like to spend a moment in remembering my grandma Rina, who passed away shortly after my Bachelor graduation. I would not be where I am now without her love and support during the past two decades.

## SUMMARY AND OUTLOOK

# APPENDIX A

## Variables Distributions After the Transverse Mass Cut



Figure A.1:  $\mu \tau_{had}$  channel baseline selection distributions after the  $m_T < 50$  GeV cut.



Figure A.2:  $e\tau_{had}$  channel baseline selection distributions after the  $m_T < 50$  GeV cut.

# Appendix B

## Tag&Probe $Z \rightarrow ee$ Invariant Mass Distributions





Figure B.1:  $Z \rightarrow ee$  invariant mass distribution for Electron Id&Iso efficiency measurement for data, with 1.48 <  $|\eta|$  < 2.1 and 10 <  $p_T(\text{GeV})$  < 15. The red line represent the signal, the dashed blue line the background.

Figure B.2:  $Z \rightarrow ee$  invariant mass distribution for Electron Id&Iso efficiency measurement for data, with  $|\eta| < 1.48$  and  $40 < p_T(\text{GeV}) < 60$ . The red line represent the signal, the dashed blue line the background.

94APPENDIX B. ELECTRON TAG&PROBE INVARIANT MASS DISTRIBUTIONS

## Bibliography

- [1] The ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020.
- The CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 30–61.
   DOI: 10.1016/j.physletb.2012.08.021.
- [3] P. W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Physics Letters* 12.2 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9.
- [4] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: Phys. Rev. Lett. 13 (Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.
- [5] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (Aug. 1964), pp. 321–323. DOI: 10.1103/Phy sRevLett.13.321.
- G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (20 Nov. 1964), pp. 585-587. DOI: 10.1103/PhysRevLett.13.585. URL: http://link.aps.org/doi/10.1103/ PhysRevLett.13.585.
- S. Weinberg. "A model of leptons". In: *Phys. Rev. Lett.* 22 (1967), pp. 1264–1266.
   DOI: doi:10.1103/PhysRevLett.19.1264.
- [8] M. A. Salam and J. C. Ward. "Weak and electromagnetic interactions". In: Il Nuovo Cimento 11 (Issue 4 Feb. 1959), pp. 568–577. DOI: 10.1007/BF02726525.
- G. 't Hooft. "Renormalizable Lagrangians for massive Yang-Mills fields". In: Nuclear Physics B 35.1 (1971), pp. 167–188. ISSN: 0550-3213. DOI: http://dx.doi. org/10.1016/0550-3213(71)90139-8. URL: http://www.sciencedirect.com/ science/article/pii/0550321371901398.

- G. 't Hooft and M. J. G. Veltman. "Regularization and Renormalization of Gauge Fields". In: *Nucl. Phys.* B44 (1972), pp. 189–213. DOI: 10.1016/0550-3213(72) 90279-9.
- [11] G. Aad et al. "Combined Measurement of the Higgs Boson Mass in pp Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments". In: *Phys. Rev. Lett.* 114 (2015), p. 191803. DOI: 10.1103/PhysRevLett.114.191803. arXiv: 1503.07589 [hep-ex].
- [12] S. Chatrchyan and other. "Evidence for the 125 GeV Higgs boson decaying to a pair of  $\tau$  leptons". In: Journal of High Energy Physics 2014.5 (2014), pp. 1–72. ISSN: 1029-8479. DOI: 10.1007/JHEP05(2014)104. URL: http://dx.doi.org/10.1007/JHEP05(2014)104.
- G. Aad et al. "Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector". In: JHEP 04 (2015), p. 117. DOI: 10.1007/JHEP04(2015) 117. arXiv: 1501.04943 [hep-ex].
- [14] S. P. Martin. "A Supersymmetry Primer". In: Perspectives on Supersymmetry II, pp. 1-153. DOI: 10.1142/9789814307505\_0001. eprint: http://www.worldsc ientific.com/doi/pdf/10.1142/9789814307505\_0001. URL: http://www. worldscientific.com/doi/abs/10.1142/9789814307505\_0001.
- [15] S. L. Glashow. "Partial symmetries of weak interactions". In: Nuclear Physics 22.4 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.
- [16] M. Gell-Mann. "A schematic model of baryons and mesons". In: *Physics Letters* 8 (1964), pp. 214–215. DOI: doi:10.1016/S0031-9163(64)92001-3.
- [17] G. Zweig. "An SU(3) model for strong interaction symmetry and its breaking. Version 2". In: Developments in the Quark Theory of Hadrons. Ed. by D. Lichtenberg and S. P. Rosen. Vol. 1. 1964, pp. 22–101. URL: http://inspirehep.net/record/4674/files/cern-th-412.pdf.
- [18] A. Purchell. Go on a particle quest at the first CERN webfest. 2012. URL: http: //cds.cern.ch/record/1473657/?ln=it.
- [19] N. Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531-533. DOI: 10.1103/PhysRevLett.10.531. URL: http://link.aps.org/doi/10.1103/PhysRevLett.10.531.
- [20] M. Kobayashi and T. Maskawa. "CP Violation in the Renormalizable Theory of Weak Interaction". In: Prog. Theor. Phys. 49 (1973), pp. 652–657.
- [21] R. P. Feynman. "The Behavior of Hadron Collisions at Extreme Energies". In: *Third International Conference on High Energy Collision*. (Stony Brook University, New York, Sept. 5–6, 1969). Gordon and Breach Science Publishers, Sept. 1969, pp. 237–249.

- [22] The CMS Collaboration. "Measurement of the Drell-Yan Cross Section in pp Collisions at  $\sqrt{s} = 7$  TeV". In: J. High Energy Phys. 10.arXiv:1108.0566. CMS-EWK-10-007. CERN-PH-EP-2011-093 (Aug. 2011), 007. 37 p. URL: http://cds.cern.ch/record/1372196.
- [23] LHC Higgs Cross Section Working Group et al. "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables". In: *CERN-2011-002* (CERN, Geneva, 2011).
   DOI: 10.5170/CERN-2011-002. arXiv: 1101.0593 [hep-ph].
- [24] LHC Higgs Cross Section Working Group et al. "Handbook of LHC Higgs Cross Sections: 2. Differential Distributions". In: CERN-2012-002 (CERN, Geneva, 2012).
   DOI: 10.5170/CERN-2012-002. arXiv: 1201.3084 [hep-ph].
- [25] LHC Higgs Cross Section Working Group et al. "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties". In: CERN-2013-004 (CERN, Geneva, 2013). DOI: 10.5170/CERN-2013-004. arXiv: 1307.1347 [hep-ph].
- [26] J. Stirling. Parton Luminosity and Cross Section Plots. 2013. URL: http://www. hep.ph.ic.ac.uk/~wstirlin/plots/plots.html.
- [27] The CMS Collaboration. LHC Higgs Cross Section Working Group. 2016. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG.
- [28] J. M. H. Salfeld-Nebgen. "Search for the Higgs Boson Decaying into  $\tau$ -Leptons in the Di-Electron Channel". PhD thesis. U. Hamburg, Dept. Phys., 2014. URL: http://www-library.desy.de/cgi-bin/showprep.pl?thesis14-019.
- [29] CERN Rapport annuel 2014. Tech. rep. Geneva: CERN, 2015. URL: https://cds. cern.ch/record/2026820.
- [30] F. J. Hasert et al. "Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment". In: *Phys. Lett.* B46 (1973), pp. 138–140. DOI: 10.1016/0370-2693(73)90499-1.
- [31] M. Banner et al. "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the {CERN} pp collider". In: *Physics Letters B* 122.5–6 (1983), pp. 476–485. DOI: http://dx.doi.org/10.1016/0370-2693(83)91605-2.
- [32] G. Arnison et al. "Experimental observation of isolated large transverse energy electrons with associated missing energy at s=540 GeV". In: *Physics Letters B* 122.1 (1983), pp. 103–116. DOI: http://dx.doi.org/10.1016/0370-2693(83) 91177-2.
- [33] A. Team. "Diagram of an LHC dipole magnet. Schéma d'un aimant dipôle du LHC". June 1999. URL: http://cds.cern.ch/record/40524.
- [34] J. Haffner. The CERN accelerator complex. Complexe des accèlèrateurs du CERN. Oct. 2013. URL: http://cds.cern.ch/record/1621894.
- [35] D. Marlow. Public CMS Luminosity Information. CMS Experiment at LHC, CERN. 2016. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublic Results.

- [36] The ALICE Collaboration. "The ALICE experiment at the CERN LHC". In: Journal of Instrumentation 3.08 (2008), S08002. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08002.
- [37] The ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: Journal of Instrumentation 3.08 (2008), S08003. URL: http://stac ks.iop.org/1748-0221/3/i=08/a=S08003.
- [38] The LHCb Collaboration. "The LHCb Detector at the LHC". In: Journal of Instrumentation 3.08 (2008), S08005. URL: http://stacks.iop.org/1748-0221/ 3/i=08/a=S08005.
- [39] The LHCf Collaboration. "The LHCf detector at the CERN Large Hadron Collider". In: Journal of Instrumentation 3.08 (2008), S08006. URL: http://stacks. iop.org/1748-0221/3/i=08/a=S08006.
- [40] The MoEDAL Collaboration. Technical Design Report of the MoEDAL Experiment. Tech. rep. CERN-LHCC-2009-006. MoEDAL-TDR-001. June 2009. URL: http://cds.cern.ch/record/1181486.
- [41] The TOTEM Collaboration. "The TOTEM Experiment at the CERN Large Hadron Collider". In: Journal of Instrumentation 3.08 (2008), S08007. URL: http:// stacks.iop.org/1748-0221/3/i=08/a=S08007.
- [42] The CMS Collaboration. "The CMS experiment at the CERN LHC". In: Journal of Instrumentation 3.08 (2008), S08004. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08004.
- [43] T. Sakuma. 3D SketchUp images of the CMS detector. Sept. 2012. URL: https: //cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=11514.
- [44] The CMS Collaboration. The CMS tracker system project: Technical Design Report. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds. cern.ch/record/368412.
- S. Chatrchyan et al. "Alignment of the CMS Silicon Tracker during Commissioning with Cosmic Rays". In: JINST 5 (2010), T03009. DOI: 10.1088/1748-0221/5/ 03/T03009. arXiv: 0910.2505 [physics.ins-det].
- [46] The CMS Collaboration. The CMS electromagnetic calorimeter project: Technical Design Report. Technical Design Report CMS. Geneva: CERN, 1997. URL: https: //cds.cern.ch/record/349375.
- [47] The CMS Collaboration. CMS Physics: Technical Design Report Volume 1: Detector Performance and Software. Technical Design Report CMS. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/922757.
- [48] J. Williams. CMS ECAL Endcaps photographs. STFC Particle Physics Department. June 2008. URL: http://hepwww.rl.ac.uk/CMSvpt/bestphotos/UK/ slides/UK%20RAL%20-%20VPT%20and%20Crystal.htm.
- [49] The CMS Collaboration. The CMS hadron calorimeter project: Technical Design Report. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds. cern.ch/record/357153.
- [50] M. Brice, L. Veillet, and D. Lazic. CMS prepares for Pixel and HCAL upgrades. CMS Experiment at LHC, CERN. Oct. 2012. URL: http://cms.web.cern.ch/ news/cms-prepares-pixel-and-hcal-upgrades.
- [51] The CMS Collaboration. CMS, the magnet project: Technical design report. 1997.
- [52] The CMS Collaboration. "Precise mapping of the magnetic field in the CMS barrel yoke using cosmic rays". In: *Journal of Instrumentation* 5.03 (2010), T03021. URL: http://stacks.iop.org/1748-0221/5/i=03/a=T03021.
- [53] The CMS Collaboration. The CMS muon project: Technical Design Report. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/ record/343814.
- [54] D. E. Groom, N. V. Mokhov, and S. I. Striganov. "Muon stopping power and range tables 10-MeV to 100-TeV". In: Atom. Data Nucl. Data Tabl. 78 (2001), pp. 183–356. DOI: 10.1006/adnd.2001.0861.
- [55] M. S. Kim et al. "CMS reconstruction improvement for the muon tracking by the RPC chambers". In: *PoS* RPC2012 (2012). [JINST8,T03001(2013)], p. 045. DOI: 10.1088/1748-0221/8/03/T03001. arXiv: 1209.2646 [physics.ins-det].
- [56] The CMS Collaboration. CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems. Technical Design Report CMS. URL: https://cds.cern. ch/record/706847.
- [57] S. Cittolin, A. Rácz, and P. Sphicas. CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and dataacquisition project. Technical Design Report CMS. Geneva: CERN, 2002. URL: https://cds.cern.ch/record/578006.
- [58] D. Barney. CMS slice raw illustrator files. Version 1. CMS Experiment at LHC, CERN. 2011. URL: https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ ShowDocument?docid=5581.
- [59] A. Holzner. "78 reconstructed vertices in event from high-pileup run 198609". CMS Collection. Sept. 2012. URL: http://cds.cern.ch/record/1479324.
- [60] E. Chabanat and N. Estre. "Deterministic Annealing for Vertex Finding at CMS". In: (2005). URL: https://cds.cern.ch/record/865587.
- [61] R. Frühwirth et al. "New vertex reconstruction algorithms for CMS". In: *Computing in High Energy and Nuclear Physics* 24-28 (Mar. 2003).
- [62] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Tech. rep. CMS-PAS-PFT-09-001. 2009. Geneva: CERN, Apr. 2009. URL: https://cds.cern.ch/record/1194487.

- [63] Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector. Tech. rep. CMS-PAS-PFT-10-001. 2010. URL: https://cds.cern.ch/record/1247373.
- [64] Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV. Tech. rep. CMS-PAS-PFT-10-002. Geneva: CERN, 2010. URL: https://cds.cern.ch/record/1279341.
- [65] W. Adam et al. Track Reconstruction in the CMS tracker. Tech. rep. CMS-NOTE-2006-041. Geneva: CERN, Dec. 2006. URL: https://cds.cern.ch/record/ 934067.
- [66] C. Botta and G. Petrucciani. Proposal of a new "Medium" Muon Id. CERN, Dec. 8, 2014. URL: https://indico.cern.ch/event/357213/contributions/1769745/ attachments/710701/975626/muonid-pog081214.pdf.
- [67] S. Baffioni et al. "Electron reconstruction in CMS". In: The European Physical Journal C 49 (Issue 4, Mar. 2007), pp. 1099–1116. DOI: 10.1140/epjc/s10052-006-0175-5.
- [68] W. Adam et al. Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC. Tech. rep. CMS-NOTE-2005-001. Geneva: CERN, Jan. 2005. URL: https://cds.cern.ch/record/815410.
- [69] J. Incandela. Status of the CMS SM Higgs Search. University of California, Santa Barbara and CERN, July 3, 2012. URL: https://cms-docdb.cern.ch/cgibin/PublicDocDB/RetrieveFile?docid=6125&filename=CMS\_4July2012\_Inca ndela.pdf.
- [70] The CMS Collaboration. "Measurement of the properties of a Higgs boson in the four-lepton final state". In: *Phys. Rev.* D89.9 (2014), p. 092007. DOI: 10.1103/ PhysRevD.89.092007. arXiv: 1312.5353 [hep-ex].
- [71] R. Boniecki. Tau identification and reconstruction at CMS. University of Warsaw, Jan. 9, 2013. URL: http://epiphany.ifj.edu.pl/epiphany.2013/pres/day3\_ 3\_Boniecki\_Tau\_identification\_and\_reconstruction\_at\_CMS.pdf.
- [72] The CMS collaboration. "Reconstruction and identification of  $\tau$  lepton decays to hadrons and  $\nu_{\tau}$  at CMS". In: Journal of Instrumentation 11.01 (2016), P01019. URL: http://stacks.iop.org/1748-0221/11/i=01/a=P01019.
- [73] The CMS Collaboration. "Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2". July 6, 2016. URL: http://inspirehep.net/record/1474218.
- [74] F. De Guio and the Cms Collaboration. "The CMS data quality monitoring software: experience and future prospects". In: Journal of Physics: Conference Series 513.3 (2014), p. 032024. URL: http://stacks.iop.org/1742-6596/513/i=3/a= 032024.

- S. Frixione, P. Nason, and C. Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: JHEP 11 (2007), p. 070.
  DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].
- [76] J. Alwall et al. "MadGraph 5 : Going Beyond". In: JHEP 06 (2011), p. 128. DOI: 10.1007/JHEP06(2011)128. arXiv: 1106.0522 [hep-ph].
- J. Pumplin et al. "New generation of parton distributions with uncertainties from global QCD analysis". In: JHEP 07 (2002), p. 012. DOI: 10.1088/1126-6708/2002/07/012. arXiv: hep-ph/0201195 [hep-ph].
- [78] T. Sjöstrand et al. PYTHIA. 2016. URL: http://home.thep.lu.se/~torbjorn/ Pythia.html.
- [79] B. Andersson. The Lund Model: Cambridge: Cambridge University Press, Oct. 1998. ISBN: 9780511524363. DOI: 10.1017/CB09780511524363. URL: https:// www.cambridge.org/core/books/the-lund-model/F9442FBA5696D5DE0D0E8D 46F70C34C7.
- [80] S. Agostinelli et al. "Geant4—a simulation toolkit". In: Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: http: //dx.doi.org/10.1016/S0168-9002(03)01368-8.
- [81] J. Allison et al. "Geant4 developments and applications". In: *IEEE Transactions on Nuclear Science* 53.1 (Feb. 2006), pp. 270–278. ISSN: 0018-9499. DOI: 10.1109/TNS.2006.869826.
- [82] N. Bartosik and A. Geiser. "Associated Top-Quark-Pair and b-Jet Production in the Dilepton Channel at  $\sqrt{s} = 8$  TeV as Test of QCD and Background to tt+Higgs Production". Presented 03 Jul 2015. PhD thesis. Hamburg: Hamburg U., Aug. 2015. URL: https://cds.cern.ch/record/2047049.
- [83] L. Calligaris and A. B. Meyer. "Search for a Heavy Neutral Higgs Boson decaying into Tau Pairs with the CMS Experiment at the LHC". PhD thesis. Hamburg U., 2015. URL: https://cds.cern.ch/record/2103663.
- [84] CMS Luminosity Measurement for the 2015 Data Taking Period. Tech. rep. CMS-PAS-LUM-15-001. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/ 2138682.
- [85] Search for a neutral MSSM Higgs boson decaying into  $\tau\tau$  with 12.9 fb<sup>-1</sup> of data at 13 TeV. Tech. rep. CMS-PAS-HIG-16-037. Geneva: CERN, 2016.
- [86] K. A. Olive et al. "Review of Particle Physics". In: Chin. Phys. C38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001.
- [87] H. Li, T. Seva, and H. Tong. Standard Model Cross Sections for CMS at 13 TeV. 2016. URL: https://twiki.cern.ch/twiki/bin/viewauth/CMS/StandardModel CrossSectionsat13TeV.

[88] Y. Li and F. Petriello. "Combining QCD and electroweak corrections to dilepton production in FEWZ". In: *Phys. Rev.* D86 (2012), p. 094034. DOI: 10.1103/ PhysRevD.86.094034. arXiv: 1208.5967 [hep-ph].

102