UNIVERSITY OF CALIFORNIA
RIVERSIDE

Search for Heavy Majorana Neutrinos in $\mu^{\pm}\mu^{\pm}$+jets and $e^{\pm}e^{\pm}$+jets Events in pp Collisions at $\sqrt{s} = 7$ TeV

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Ferdinando Giordano

March 2013

CERN-THESIS-2013-024
08/02/2013

Dissertation Committee:

Dr. John Ellison, Chairperson
Dr. Robert B. Clare
Dr. Jose Wudka

Copyright by
Ferdinando Giordano
2013

The Dissertation of Ferdinando Giordano is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I would like to thank all the people that welcomed me in California and made it my second home for almost five years, without them I could not succeed in this experience. During the past years I met a lot of people who enriched my life and helped me in my work. I would like to thank all the people working at UCR for their help to get started with my life here, in particular the physics department where I have found more than a good working environment.

I still remember the tough first year and the help that came from all of my classmates, I would like to thank all of them, the ones that are still at UCR and the ones that already left. Then I moved from classmates to colleagues that soon became good friends, it would be difficult not to name Shin Chuan and Arun, for their support at work and their company during these years in America and Europe. Within the HEP group I have found a collaborative environment where it has always been easy to work thanks to great colleagues, expert post-doc and understanding professors, I would like to thank Bob and Steve for their friendly discussions and in particular my advisor John for his immense patience and priceless help during all my PhD.

Besides the physics department at UCR I met many other people that supported me during these years, some of them have already moved on and some are still here. I will never forget Christopher and Christian, the best house-mates I could wish to have, in different ways they helped me during my first year in Riverside and without them I could never be here today. I would like also to thank David and Wenchi that received me in their house not only as a tenant but as a friend.

After the first year I started my routing traveling schedule between CERN and UCR, moving was difficult but I had always good friends waiting for me, sometimes I

was so lucky to have them in both places. I would like to thank Wei Song for our long chats and gaming nights, Giacomo for our soccer matches, Alessandro for our dinners all together, thanks to them and all the other people that enriched my life during the periods I spent in Riverside.

On the other side of the Atlantic there were probably even more people waiting for me, long time friends and relatives, as well as new friends: my best friend and "brother" Pietro and my classmates from the high-school in Riola, all the physicists from my university in Bologna and all the others that I met during my PhD who soon became my close friends. And last but not least Michela who really took my hand and guided me better than Beatrice would have done, supporting and suffering me during this time.

Particular gratitude to all my relatives, also to the ones that are not here any more, I missed them during these years. My parents especially are the ones I have to thank the mosts, and after I dedicated my master thesis to my mother I want to dedicate this to my father for his silent and always present support during all my life and in all the choice I made.

to my father

ABSTRACT OF THE DISSERTATION

Search for Heavy Majorana Neutrinos in $\mu^{\pm}\mu^{\pm}$+jets and $e^{\pm}e^{\pm}$+jets Events in pp Collisions at $\sqrt{s} = 7$ TeV

by

Ferdinando Giordano

Doctor of Philosophy, Graduate Program in Physics
University of California, Riverside, March 2013
Dr. John Ellison, Chairperson

A search is performed for heavy Majorana neutrinos (N) using an event signature defined by two same-sign charged leptons of the same flavour and two jets. The data correspond to an integrated luminosity of 4.98 fb^{-1} of pp collisions at a centre-of-mass energy of 7 TeV collected with the CMS detector at the Large Hadron Collider. No excess of events is observed beyond the expected standard model background and therefore upper limits are set on the square of the mixing parameter, $|V_{\ell N}|^2$, for $\ell = e, \mu$, as a function of heavy Majorana-neutrino mass. These are the first direct upper limits on the heavy Majorana-neutrino mixing for $m_N > 90$ GeV.

Contents

List of Figures **xv**

List of Tables **xvii**

1 Introduction **1**

2 Standard Model and Beyond **4**
2.1 The Standard Model 4
2.1.1 Quantum electrodynamics, a $U(1)$ symmetry 6
2.1.2 Non-abelian gauge symmetries, path to QCD 7
2.1.3 Electroweak unification and the Higgs boson 9
2.1.4 Fermion masses 15
2.2 Current limits of the Standard Model 16
2.2.1 Neutrino oscillations 18
2.2.2 Neutrino masses 21
2.3 Possible extensions 21
2.3.1 Dirac mass term 21
2.3.2 Majorana mass term 22
2.3.3 Dirac and Majorana mass terms 23
2.3.4 The heavy neutrino and the see-saw mechanism 24

3 LHC and CMS **30**
3.1 The LHC collider 30
3.2 Main LHC components 32
3.3 The CMS detector 35
3.3.1 The magnet 38
3.3.2 The Tracker 40
3.3.3 The Electromagnetic Calorimeter 42
3.3.4 The Hadronic Calorimeter 45
3.3.5 The Muon System 47
3.3.5.1 The DTs 48
3.3.5.2 The CSCs 50
3.3.5.3 The RPCs 51
3.3.6 Trigger and Data Acquisition System (DAQ) 52
3.3.6.1 The Level 1 Trigger 53
3.3.6.2 The HTL Trigger 54

3.3.6.3 Data Acquisition (DAQ) . . . 55

4 CMS software and reconstruction 57
4.1 Software framework . . . 57
4.2 Simulation . . . 60
4.3 Muon Reconstruction . . . 65
4.3.1 Tracking and Tracker Muons . . . 65
4.3.2 Standalone muons . . . 69
4.3.3 Global muons . . . 70
4.4 Jet Reconstruction . . . 72
4.4.1 Calorimetric tower . . . 72
4.4.2 Anti-k_T cone jet clustering algorithm . . . 73
4.4.3 Jet energy correction and jet energy scale . . . 75
4.4.4 b-tagging . . . 78
4.5 Missing Transverse Energy . . . 80
4.6 Electrons and Photons . . . 82
4.6.1 Electron reconstruction . . . 83
4.6.2 Photon reconstruction . . . 85
4.7 Particle Flow . . . 85
4.7.1 Iterative tracking . . . 86
4.7.2 Calorimeter clustering . . . 87
4.7.3 Link algorithm . . . 88
4.7.4 Particle reconstruction improvements . . . 89

5 Event Selection 91
5.1 Data and MC samples . . . 91
5.1.1 2011 collision data . . . 91
5.1.2 Standard Model monte carlo samples . . . 92
5.1.3 Majorana neutrino signal monte carlo . . . 93
5.2 Signal Selection . . . 97
5.2.1 Trigger . . . 100
5.2.2 Primary vertex selection . . . 103
5.2.3 Muon selection . . . 103
5.2.4 Electron selection . . . 106
5.2.5 Jet selection . . . 108
5.2.6 Final selection requirements . . . 108
5.3 Selection Efficiency . . . 109
5.3.1 Offline selection efficiency . . . 109
5.3.2 Data-MC scale factor . . . 111

6 Background Estimation 116
6.1 Backgrounds due to SM sources of same-sign events . . . 116
6.2 Data-driven background prediction for fake muons . . . 119
6.2.1 The fake rate method . . . 119
6.2.2 Fake rate event selection . . . 122
6.2.3 Calibration of the background prediction . . . 125
6.2.4 Validation of the background prediction . . . 131
6.2.4.1 Monte Carlo closure tests . . . 131
6.2.4.2 Data closure test . . . 134

6.3 Charge misidentification background 141

7 Systematic Uncertainties **144**
7.1 General systematic uncertainties 144
7.2 Systematic uncertainty on fake muon background 146

8 Results and Conclusions **151**
8.1 Muon channel 151
8.1.1 Loose selection 151
8.1.2 Tight selection 155
8.2 Electron channel 158
8.2.1 Loose selection 158
8.2.2 Tight selection 158
8.3 Interpretation of results 160

Bibliography **163**

List of Figures

1.1 Faynamn diagram of the Majorana neutrino 3

2.1 Elementary partilces of standard model 6
2.2 Fermi diagram 9
2.3 EWK vertex 10
2.4 Higgs potential 13
2.5 Double-well potential 14
2.6 Majorana neutrino vertex diagrams 27
2.7 Majorana production via WW fusion 28
2.8 Majorana resonant production via W boson 28

3.1 LHC section 32
3.2 LHC injection sytem 33
3.3 CMS detector 35
3.4 CMS transverse and longitudinal view 39
3.5 CMS magnetic field map 40
3.6 Pixel detector 41
3.7 Strips detector 42
3.8 ECAL detector 43
3.9 ECAL detector 43
3.10 HCAL detector 45
3.11 Transverse view of the CMS muon system 47
3.12 Section of a DT chamber 49
3.13 DT cell 49
3.14 CSC 50
3.15 CSC layout 51
3.16 RPC section 52
3.17 L1 scheme 53
3.18 DAQ Architecture 55

4.1 Kalman filter 67
4.2 ECAL depth showering correction. 88
4.3 PF vs Calojets matching efficiency 89
4.4 PF vs Calojets resolutions 90

5.1 Faynamn diagram of the Majorana neutrino 97
5.2 Isolation cone 104

6.1 Number of vertices distribution in data . . . 117
6.2 Number of vertices distribution in Summer11 MC . . . 118
6.3 Number of vertices distribution in Fall11 MC . . . 118
6.4 Momentum of the fakable object . . . 124
6.5 Muon fake rate 2-D plot . . . 125
6.6 Simple functional form of the fake prediction vs $RelIso$ . . . 129
6.7 Derivative of the simple functional form of the fake prediction vs $RelIso$ 129
6.8 Efficiency and fake rate vs number of vertices . . . 131
6.9 $t\bar{t}$ closure test - $\mu_2$jj invariant mass, away-side jet $p_T > 40$ GeV . . . 133
6.10 $t\bar{t}$ closure test - $\mu_2$jj invariant mass, away-side jet $p_T > 80$ GeV . . . 133
6.11 $t\bar{t}$ closure test - μ momentum, away-side jet $p_T > 40$ GeV . . . 134
6.12 $t\bar{t}$ closure test - μ momentum, away-side jet $p_T > 80$ GeV . . . 134
6.13 Data closure test - $\mu_2$jj invariant mass for $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV (loose) . . . 136
6.14 Data closure test - μ momentum for $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV (loose) . . . 136
6.15 Data closure test - $\mu\mu$jj for $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV (loose) . . . 137
6.16 Data closure test - dijets mass for $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV (loose) . . . 137
6.17 Data closure test - $\mu_2$jj invariant mass for $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV (tight) . . . 139
6.18 Data closure test - μ momentum for $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV (tight) . . . 139
6.19 Data closure test - $\mu\mu$jj for $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV (tight) . . . 140
6.20 Data closure test - dijets mass for $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV (tight) . . . 140
6.21 electron charge missID probability . . . 143

7.1 Main background from single fake prediction . . . 149
7.2 Main background from single fake plus double fake prediction . . . 149
7.3 Main background from double fake prediction . . . 150

8.1 Muon channel signal selection (loose) - $\mu_2$jj invariant mass . . . 153
8.2 Muon channel signal selection (loose) - leading μ momentum . . . 153
8.3 Muon channel signal selection (loose) - $\mu\mu$jj invariant mass . . . 154
8.4 Muon channel signal selection (loose) - jj invariant mass . . . 154
8.5 Muon channel signal selection (tight) - $\mu_2$jj invariant mass . . . 156
8.6 Muon channel signal selection (tight) - leading μ momentum . . . 156
8.7 Muon channel signal selection (tight) - $\mu\mu$jj invariant mass . . . 157
8.8 Muon channel signal selection (tight) - jj invariant mass . . . 157
8.9 Electron channel signal selection (loose) - $e_2$jj invariant mass . . . 159
8.10 Electron channel signal selection (tight) - $e_2$jj invariant mass . . . 160
8.11 Esclusion plot muon channel . . . 161
8.12 Esclusion plot electron channel . . . 162

List of Tables

2.1 The three fundamental Interactions in the Standard model 5

3.1 The LHC design characteristics summary 34

5.1 Muon datasets list 92
5.2 Electron datasets list 92
5.3 SM monte carlo samples 94
5.4 Muon channel SM monte carlo samples 95
5.5 Electron channel SM monte carlo samples 96
5.6 Signal MC samples for the muon channel 98
5.7 Signal MC samples for the electron channel 99
5.8 Electron triggers IDs 102
5.9 Effective areas for muon isolation correction 105
5.10 VBTF selection criteria 107
5.11 Effective areas for electron isolation correction 107
5.12 Muon individual cut efficiencies for events passing the loose selection . . 109
5.13 Muon individual cut efficiencies for events passing the tight selection . . 110
5.14 Muon channel event selection efficiencies for the loose selection 110
5.15 Muon channel event selection efficiencies for the tight selection 111
5.16 Electron channel event selection efficiencies for the loose selection 112
5.17 Electron channel event selection efficiencies for the tight selection 113
5.18 Muon ID and Iso efficiencies for data and MC. 114
5.19 Electron ID and Iso efficiencies for data and MC. 115

6.1 Muon fake rates in QCD, jet $p_T > 20$ GeV 127
6.2 Muon fake rates in QCD, jet $p_T > 40$ GeV 127
6.3 Muon fake rates in QCD, jet $p_T > 60$ GeV 127
6.4 Muon fake rates from data, final selection 130
6.5 Muon fake rates from MC, final selection 130
6.6 Closure test summary for loose selection 132
6.7 Closure test summary for tight selection 135
6.8 Results for the $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV control region using loose selection . . . 135
6.9 Results for the $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV control region using tight selection . . . 138

7.1 Summary of systematic uncertainties. 146
7.2 Final prediction variation 148

8.1 Event yields for the muon channel loose signal selection 152
8.2 Event yields for the muon channel tight signal selection 155
8.3 Event yields for the electron channel loose signal selection 158
8.4 Event yields for the electron channel tight signal selection 159

Chapter 1

Introduction

The non-zero masses of neutrinos, confirmed from studies of their oscillations among three species, provide the first evidence for physics beyond the standard model (SM) [18]. The smallness of neutrino masses underscore the lack of a coherent formulation for the generation of mass of elementary particles. The leading theoretical candidate for accommodating neutrino masses is the so-called "seesaw" mechanism [81, 63, 94, 82], where the smallness of the observed neutrino masses (m_ν) is attributed to the largeness of a mass (m_N) of a new massive neutrino state N, with $m_\nu \approx y_\nu^2 v^2/m_N$, where y_ν is a Yukawa coupling of ν to the Higgs field, and v is the Higgs vacuum expectation value in the SM. In this model the SM neutrinos would also be Majorana particles. Owing to the new heavy neutrinos Majorana nature, it is its own antiparticle, which allows processes that violate lepton-number conservation by two units. Consequently, searches for heavy Majorana neutrinos are of fundamental interest.

The phenomenology of searches for heavy Majorana neutrinos at hadron colliders has been considered by many authors [72, 50, 46, 9, 84, 48, 66, 12]. Our search follows the studies in Refs. [66, 48] that use a model independent phenomenological

approach, with m_N and $V_{\ell N}$ as free parameters, where $V_{\ell N}$ is a mixing parameter describing the mixing between the heavy Majorana neutrino and the SM neutrino ν of flavour . Previous direct searches for heavy Majorana neutrinos based on this model have been reported by the L3 [8] and DELPHI [7] Collaborations at the Large Electron-Positron Collider. They have searched for $Z \to \nu_\ell N$ decays and set limits on $|V_{\ell N}|^2$ as a function of m_N for heavy Majorana-neutrino masses up to approximately 90 GeV. The ATLAS Collaboration at the Large Hadron Collider (LHC) has also reported limits on heavy Majorana neutrino production [5, 6] in the context of an effective Lagrangian approach [49] and the Left-Right Symmetric Model [60, 69]. Indirect limits on $|V_{\ell N}|^2$ have been obtained from the non-observation of neutrinoless double beta decay [16], resulting in 90% confidence level (CL) limits of $|V_{\ell N}|^2/m_N < 7 \times 10^{-5}$ TeV^{-1} . Precision electroweak measurements have been used to constrain the mixing parameters resulting in indirect 90% CL limits of $|V_{eN}|^2 < 0.0066$, $|V_{\mu N}|^2 < 0.0060$, and $|V_{\tau N}|^2 < 0.016$ [75].

We report on a search for the production of a heavy Majorana neutrino in proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 7$ TeV at the LHC using a set of data of integrated luminosity 4.98 ± 0.11 fb^{-1} collected with the Compact Muon Solenoid (CMS) detector. The principal Feynman diagram for this process is shown in Fig. 1.1. The heavy Majorana neutrino can decay to a lepton with positive or negative charge, leading to events containing two leptons with the same or opposite sign. Same-sign events have much lower backgrounds from SM processes and therefore provide an accessible signature of heavy Majorana neutrino production. We search for events with two isolated leptons of same sign and flavour ($\mu^\pm\mu^\pm$ or $e^\pm e^\pm$) and at least two accompanying jets. Contributions from SM processes to such dilepton final states are very small and the background is dominated by processes such as multijet production, in which leptons from b-quark decays or from jets are misidentified as isolated prompt

leptons.

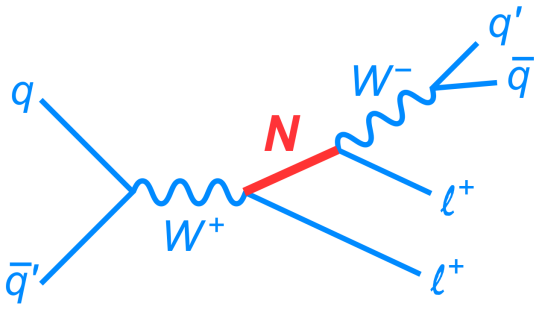


Figure 1.1: The lowest-order Feynman diagram for production of a heavy Majorana neutrino N. The charge-conjugate diagram also contributes and results in a $\ell^-\ell^- q\bar{q}'$ final state.

Chapter 2

Standard Model and Beyond

2.1 The Standard Model

The Standard Model (SM) [64, 80, 22] is, so far, the most successful theory to describe the fundamental particles and their interactions. It is based on the Quantum Field Theory (QFT) [78] formalism and it consists of two main categories: fermions and bosons. The fermions are the constituents of matter while the bosons are the particles responsible for their interactions. The newly discovered boson [26], with a mass of approximately 125 GeV is consistent with the so-called Higgs boson, belied to be responsible for the nonzero particle masses and is needed to fulfill the general description of the electroweak symmetry breaking.

The main difference between the two main groups is that fermions are described by Fermi-Dirac statistic having spins 1/2, and the bosons by Bose-Einstein statistics having spin 1 (the photon and $W^{\pm}$ and Z bosons), spin 2 (the graviton), or spin 0 (the Higgs boson).

In the current view the SM includes three of the known forces of nature summarized in Table 2.1. In order they are the *strong force*, the *electromagnetic*, and the

weak force. Gravity, compared to the the other, has a coupling of the order of 10^{-38} and can be neglected when dealing with sub-atomic particles.

Force	Gauage boson	Coupling	Theory
Strong	gluon (g)	~ 1	QCD
Electromagnetic	photon (γ)	$\sim 10^{-2}$	QED
Weak	W^+, W^- and Z	$\sim 10^{-13}$	GSW Theory

Table 2.1: The three fundamental Interactions in the Standard Model listed in order of the respective strength at low energy or large distances. These values change with energy due to the well known running coupling constants effect.

Matter is described by the three fermion generations as shown in Fig. 2.1. The fermions are divided into two main sectors: quarks and leptons. The quarks are subject to all forces. Each generation is identical to the previous one except for the mass of the particles .Across the three quark generations we can identify the *up* and *down* families. Within each family the quarks have the same charge, $+2/3$ for the up family and $-1/3$ for the down family. The leptons, on the other hand, are not subject to the strong force so they only interact via electromagnetic and weak forces. The leptons, as for the quarks, can be divided into two groups: charged leptons (e, μ, τ) and neutral leptons $(\nu_e, \nu_\mu, \nu_\tau)$.

This picture is not complete yet, since we still need to talk about the newly discovered Higgs boson. Before introducing it, and fully understanding its importance, it is useful to briefly illustrate the quantum field theory formalism used in the SM to describe the interaction of fermions and bosons. For a full description of the standard model see the references mentioned above, the formalism and notation used in this chapter follow Ref. [78].

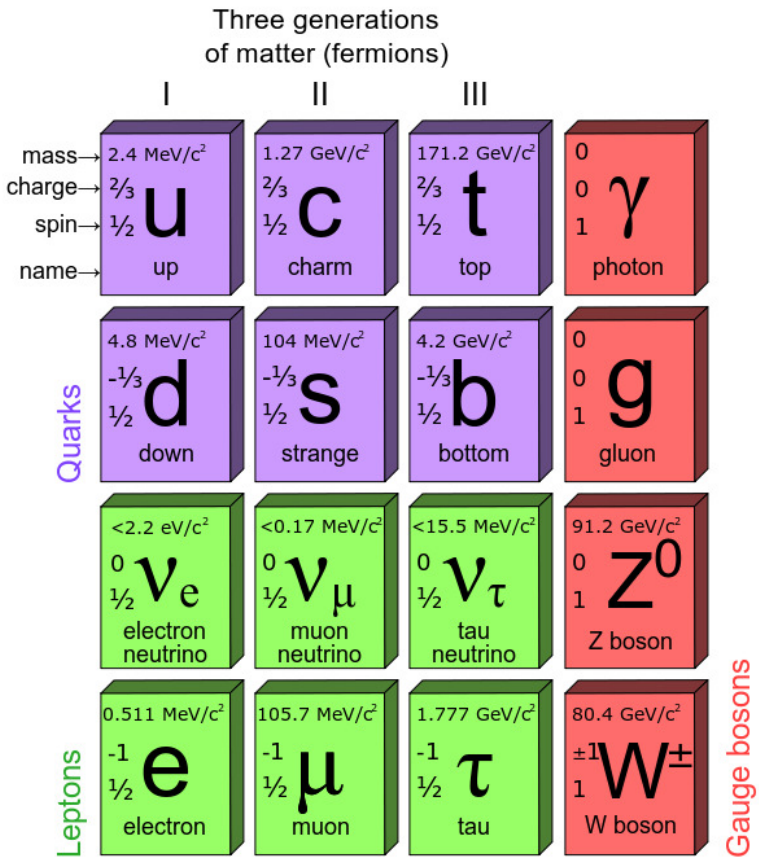


Figure 2.1: Table of elementary particles of standard model. In red are the vector bosons, in violet the quarks and in green the leptons. The mass, charge, and spin is also shown for all the particles.

2.1.1 Quantum electrodynamics, a $U(1)$ symmetry

Quantum electrodynamics (QED) is one of the best examples of the achievements of the standard model. QED describes the interactions between charged fermions via the photon, the vector-boson of the electromagnetic field. It is the synthesis of the work done by Maxwell, continued by Einstein, and ended by Dirac.

Starting from the Dirac Lagrangian $\mathcal{L}_{Dirac} = \bar{\Psi}\left(i\partial\!\!\!/ - m\right)\Psi$ and the Maxwell Lagrangian $\mathcal{L}_{Maxwell} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2$ for the free fields and imposing global and local $U(1)$ symmetry, it is possible to derive easily the Lagrangian for QED. We impose U(1) gauge symmetry using the following equations for the particle and electro-

magnetic fields

$$\Psi \to e^{iq\theta}\Psi, \tag{2.1}$$

$$A_\mu \to A_\mu - \partial_\mu \theta, \tag{2.2}$$

where θ is a parameter that can depend on the local coordinates x. These will give natural rise to the interactions between electromagnetic field and particles as we can see in the full QED Lagrangian:

$$\mathcal{L}_{QED} = \bar{\Psi}\left(i\partial\!\!\!/ - m\right)\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2 - eA_\mu\bar{\Psi}\gamma^\mu\Psi \tag{2.3}$$

the first term can be recognized as the Dirac Lagrangian where Ψ is a Dirac field representing the electron, the second and third terms represent the Lagrangian of the free electromagnetic field, where $F_{\mu\nu}$ is the Maxwell tensor, and the last term describes the interaction between the current $\bar{\Psi}\gamma^\mu\Psi$ and A_μ.

QED describes the interaction of spin-1/2 particles with photons and it has many nice properties that make it a very successful theory, for details see Refs. [70, 93, 76]. In particular here we underline that in this theory there is only one generator with one degree of freedom in the symmetry. The net result is the presence of one vector-boson, the photon, that is strictly massless due to the imposed symmetries.

2.1.2 Non-abelian gauge symmetries, path to QCD

We have not yet described all the interactions in the SM, and to do so we have to make use of other gauge groups besides the $U(1)$ used for the QED. This was first done by Yang and Mills [95] that derived a complete Lagrangian for a generic $SU(N)$ symmetry with $N^2 - 1$ generators, this will be then used to derive the Lagrangian for

the quantum chromodynamics (QCD) that describes the strong interactions in the SM.

For the QCD sector the selection of the $SU(3)$ is unique in view of:

- The group must admit complex representations to account for both quarks and anti-quarks and distinguish them. In fact there are mesons which can be conveniently described as $q\bar{q}$ bound states, but not any qq bound state.
- There must be a color singlet completely antisymmetric representation made up of qqq in order to solve the statistics puzzle of the low-lying baryons of spin 1/2 and 3/2.
- The number of colors must be in agreement with the data on the total e^+e^- hadronic cross section and the $\pi^0 \to \gamma\gamma$ decay rate.

This led to the QCD Lagrangian:

$$\mathcal{L}_{QCD} = \bar{\Psi}^{\alpha,A} \left(i\partial\!\!\!/ - m_A \right) \Psi^{\alpha,A} - \frac{1}{4} F^a_{\mu\nu} F^{a\,\mu\nu} + g A^a_\mu \bar{\Psi}^{\alpha,A} \gamma^\mu T^a_{\alpha\beta} \Psi^{\beta,A} \tag{2.4}$$

as before we can recognize the Dirac term, where now the Ψ spinor has a color index $\alpha = 1, 2, 3$ (or β) and a flavor index $A = u, d, c, s, t, b$ indicating the six quarks, the index a instead distinguishes among the eight possible QCD generators T^a.

The QCD seems a simple extension of the QED with only more generators, but actually it has some fundamental difference as the self interaction of the gluons vector-boson. The $SU(3)$ is non-abelian and therefore the eight gluons, corresponding to the eight QCD generators, can self interact carrying color charge; while the photon does not carry electric charge and cannot self interact. Similarly to the QED the QCD is a perfect symmetry and the gluons are strictly massless.

2.1.3 Electroweak unification and the Higgs boson

In the current view of the SM, the mentioned electromagnetic theory is unified with the weak force to form the so-called Electroweak force (EWK). In this section the unified picture of the EWK interactions will be briefly discussed.

The weak force was first investigated by Fermi, who wrote an effective theory valid at low energy, in particular when the exchanged momentum is less than the mass of the W. In this case the two vertices of Fig. 2.3 can be collapsed in the single vertex as shown in Fig. 2.2. Another important characteristic of the weak force is that violates parity, therefore we need a chiral group, that is, a symmetry that distinguishes between left-handed and right-handed components of fields. The chiral parts of the field are defined starting from the Dirac representation and selecting only the left or right component:

$$\Psi_L = \frac{1-\gamma^5}{2}\Psi, \qquad \Psi_R = \frac{1+\gamma^5}{2}\Psi \tag{2.5}$$

where we used the chiral or Weyl representation for γ^5 (1 represents a 2×2 unitary matrix):

$$\gamma^5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.6}$$

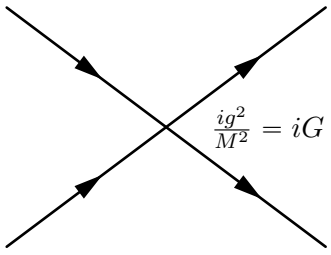


Figure 2.2: Fermi EWK effective theory vertex valid for $p^2 \ll M^2$.

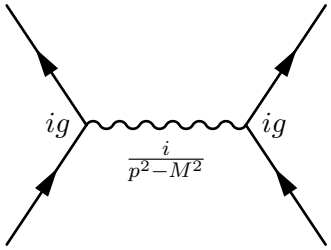


Figure 2.3: Real weak force interaction vertex.

The Fermi effective Lagrangian is therefore given by:

$$\mathcal{L}_{Fermi} = -\frac{G_F}{\sqrt{2}} \left[\bar{\Psi}_\mu \gamma^\mu \left(1-\gamma^5\right) \Psi\right] \left[\bar{\Psi} \gamma_\mu \left(1-\gamma^5\right) \Psi\right] \tag{2.7}$$

where G_F is the Fermi constant, that is related to the actual strength of the weak force and the mass of the W boson by:

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2} = 8.24 \times 10^{-6}\ \text{GeV}^{-2} \tag{2.8}$$

here is clear the nature of the weakness of this force, indeed the weak coupling $\alpha_w = g^2/4\pi$ is not smaller than the fine structure constant α ($\alpha = 1/137$, $\alpha_w = 1/29.5$), but the mediators of the weak force are massive and therefore their propagators are suppressed by $|p^2 - M^2|$ where M is either the mass of the W or Z boson. These few elements give rise to the rich phenomenology of the weak interaction for which we refer to the already mentioned bibliography.

The Fermi Lagrangian is the low energy limit of a gauge theory whom interaction part can be written as

$$\mathcal{L}_{int} = g\left(W_\mu^+ J^{+,\mu} + W_\mu^- J^{-,\mu}\right) + \bar{g} Z_\mu^0 J^{0,\mu} \tag{2.9}$$

where the charged currents enter symmetrically in the interaction Lagrangian and they couple only with the left components of the fields. They consist of two parts, one for the leptons and one for the hadrons, J^- can be written as:

$$J^{-,\mu} = J_l^{-,\mu} + J_h^{-,\mu} \tag{2.10}$$

$$= \frac{1}{\sqrt{2}} \left(\bar{e}_L \gamma^\mu \nu_{e,L} + \bar{\mu}_L \gamma^\mu \nu_{\mu,L} + \bar{\tau}_L \gamma^\mu \nu_{\tau,L} \right) \tag{2.11}$$

$$+ \frac{1}{\sqrt{2}} \left(\bar{d'}_L \gamma^\mu u_L + \bar{s'}_L \gamma^\mu c_L + \bar{b'}_L \gamma^\mu t_L \right) \tag{2.12}$$

while J^+ is the hermitian conjugate. Here (d', s', b') are linear combinations of the mass eigenstates of the down quarks (d, s, b) according to the Cabibbo-Kobayashi-Maskawa (CKM) matrix [23, 74].

Besides the charged current we have also a neutral current J^0 that can also be expressed in terms of leptonic and hadronic contributions, restricting only to the first family for simplicity, we can write:

$$J^{0,\mu} = J_l^{0,\mu} + J_h^{0,\mu} \tag{2.13}$$

$$= a_1 \bar{\nu}_{e,L} \gamma^\mu \nu_{e,L} + a_2 \bar{e}_L \gamma^\mu e_L + a_3 \bar{e}_R \gamma^\mu e_R \tag{2.14}$$

$$+ b_1 \bar{u}_L \gamma^\mu u_L + b_2 \bar{u}_R \gamma^\mu u_R + b_3 \bar{d}_L \gamma^\mu d_L + b_4 \bar{d}_R \gamma^\mu d_R \tag{2.15}$$

where the coefficients a_i and b_i are predicted by the SM in terms of a single parameter called the Weinberg angle (θ_W).

The Weinberg angle not only defines the coefficients above but also the relative strength of the coupling in Eq. 2.9:

$$\bar{g} = \frac{g}{\cos \theta_W} \tag{2.16}$$

Thus we can describe the weak interactions using a $SU(2)_L$ gauge group, but there is an important difference compared to the other symmetries considered before: the weak force vector-bosons are massive, therefore there should be some mechanism that breaks the symmetry to explain their non-zero masses.

In the full SM framework electromagnetic and weak forces are combined into the $SU(2)_L \times U(1)_Y$ symmetry group, where $SU(2)$ couples only to left-handed particles and the $U(1)$ is a symmetry of the weak hypercharge. Higgs proposed a mechanism [67] that allows the weak mediators ($W^{\pm}$ and Z) to become massive and the photon massless. This mechanism, known as the Higgs mechanism, is based on the spontaneous symmetry breaking (SSB) and keeps the SM Lagrangian invariant under the $SU(2)_L \times U(1)_Y$ gauge transformations.

The Higgs Lagrangian can be written as

$$\mathcal{L}_H = \left(D^{\mu}\right)^{\dagger} \left(D_{\mu}\right) - V\left(\phi^{\dagger}\phi\right) - \frac{1}{4} F^a_{\mu\nu} F^{a\,\mu\nu} \tag{2.17}$$

where we choose the potential to be

$$V\left(\phi^{\dagger}\phi\right) = \frac{1}{2}\lambda^2 \left(\phi^{\dagger}\phi - \eta^2\right)^2 \tag{2.18}$$

and $D_{\mu} = \partial_{\mu} - igA^a_{\mu}T^a$ is the covariant derivative defined for the three $SU(2)$ generators $T^a = \sigma^a/2$ (where σ^a are the Pauli matrices).

The field ϕ is a complex doublet that has four real components

$$\phi = \begin{pmatrix} \frac{1}{\sqrt{2}}\left(\phi_1 + i\phi_2\right) \\ \frac{1}{\sqrt{2}}\left(\phi_3 + i\phi_4\right) \end{pmatrix} \tag{2.19}$$

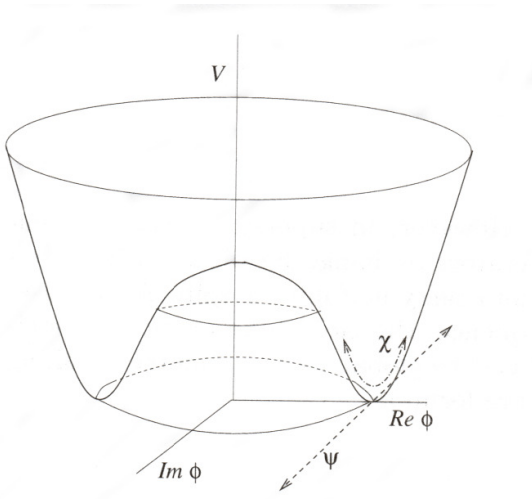


Figure 2.4: Representation of the Higgs potential 2.18, the orthogonal fluctuation χ and tangential fluctuation ψ are also shown.

that can anyway be reduced taking advantage of the $SU(2)$ invariance of the Lagrangian 2.17 to get the following representation:

$$\phi = \begin{pmatrix} 0 \\ \\ \eta + \frac{1}{\sqrt{2}}\chi \end{pmatrix} \tag{2.20}$$

Using $A_\mu^\pm = \frac{1}{\sqrt{2}}\left(A_\mu^1 \pm A_\mu^2\right)$, the first term of the right-hand side of Eq. 2.17 can be rewritten as

$$(D^\mu)^\dagger (D_\mu) = \frac{1}{2}\partial^\mu\chi\partial_\mu\chi + \frac{g^2}{4}\left(\eta + \frac{\chi}{\sqrt{2}}\right)^2 A^{3\,\mu}A_\mu^3 + \frac{g^2}{2}\left(\eta + \frac{\chi}{\sqrt{2}}\right)^2 A_\mu^+ A^{-\,\mu} \tag{2.21}$$

where we see that, apart the standard kinetic term of the field χ as well as coupling term between it and the gauge fields, all the three gauge fields acquire a mass $m_A = \frac{g\eta}{\sqrt{2}}$.

This process is known as spontaneous symmetry breaking, the key of this procedure is the choice of the potential given in Eq. 2.18, in order to fulfil an important condition necessary obtain the gauge fields masses. The condition needed is that the

potential has to have a continuous set of minima, as shown in Fig. 2.4, whereas a double-well potential (Fig. 2.5) would only create a ground state that is a superposition of the two discrete minima. In others words the two minima of Fig. 2.5 will be simply connected by tunnel effect and do not give rise to any SSB, while for the Higgs potential the fluctuations around a minima would either be along the ψ direction (see. Fig 2.4), and therefore does not cost energy, or along χ. In the latter case the displacement sees a rise of the potential and therefore it costs energy, χ would therefore be a massive fluctuation mode as opposed to ψ that is massless.

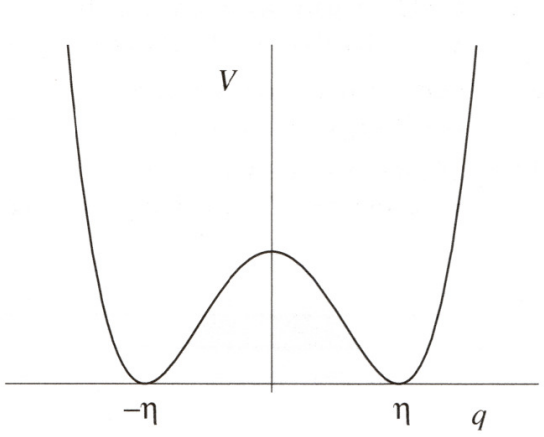


Figure 2.5: A double-well potential.

The same procedure can be applied so the $SU(2) \times U(1)$ symmetry of the SM, keeping in mind we now have another generator we can write the covariant derivative as:

$$D^{\mu} = \partial_{\mu} - igT^{a}A^{a}_{\mu} - ig'SB_{\mu} \tag{2.22}$$

and then compute $(D^{\mu})^{\dagger}(D_{\mu})$ as before and look for the mass term of the gauge fields, that are:

$$\frac{1}{4}\eta^{2}\left(gA^{3}_{\mu} - g'B_{\mu}\right)\left(gA^{3\,\mu} - g'B^{\mu}\right) + \frac{1}{2}g^{2}\eta^{2}A^{+}_{\mu}A^{-\,\mu} \tag{2.23}$$

Using the Weinber angle θ_W and the relations $\bar{g} = \sqrt{g^2 + g'^2}$, $g/\bar{g} = \cos\theta_W$, and $g'/\bar{g} = \sin\theta_W$, we can define:

$$Z^0_\mu \equiv A^3_\mu \cos\theta_W - B_\mu \sin\theta_W \tag{2.24}$$

so the Z boson gets a mass $m_Z = \bar{g}\eta/\sqrt{2}$ and, recognizing the $W^\pm$ fields as $W^\pm_\mu = A^\pm_\mu$, the W mass is given by $m_W = g\eta/\sqrt{2}$, with the ratio $\frac{m_W}{m_Z} = \cos\theta_W$. On the other hand the orthogonal field

$$A_\mu \equiv A^3_\mu \cos\theta_W + B_\mu \sin\theta_W \tag{2.25}$$

remains mass less and can be identified with the photon.

Of the initial four components of the Higgs potential, three have been absorbed as longitudinal components of the now massive Z and $W^\pm$ fields, and, in the SM, the fourth is expected to produce produce a neutral spin-0 particle called the Higgs boson. The discovery of a particle compatible with the Higgs boson was finally announced in 2012, and from the first studies it seems to have the expected properties predicted by the SM. This particle couples with the fermions causing them to acquire mass as well.

To summarize we have briefly described the SM basic symmetries $SU(3)_{QCD} \times SU(2)_L \times U(1)_Y$ and the Higgs mechanism, the interaction of the gauge bosons with fermions as well as their self interactions are described by this gauge theory.

2.1.4 Fermion masses

Once solved the mass generation problem in the boson sector we can now focus on the fermion sector. As just mentioned the coupling of the Higgs field with the fermions fields is responsible to generate the observed masses. This is achieved by

simply adding a Yukawa-coupling term of this form:

$$\mathcal{L}_Y = -g_f \left(\bar{\Psi}_R \phi^\dagger \Psi_L + \bar{\Psi}_L \phi \Psi_R \right) \tag{2.26}$$

where $\Psi_{L,\,R}$ are the chiral components of the considered fermion field and g_f is the corresponding coupling. Expanding the Higgs field according to 2.20, the Lagrangian 2.26 becomes:

$$\mathcal{L}_Y = -g_f \eta \bar{\Psi} \Psi \left(1 + \frac{\chi}{\eta\sqrt{2}} \right) \tag{2.27}$$

the mass of the fermion is therefore given by $m_f = g_f \eta$, where η is the vacuum expectation value of the Higgs field. The second term of the equation express the coupling between the fermion and the Higgs that is therefore proportional to the mass of the fermion itself.

All the fermions in the SM acquire mass via the Yukawa-coupling term except the three neutrinos, as we can see from Eq. 2.11 and 2.14 the neutrinos do not have the right component needed to build an invariant Yukawa Lagrangian as 2.26, therefore they as strictly massless in the SM.

2.2 Current limits of the Standard Model

The Standard Model has obtained remarkable achievements in understanding, describing, and predicting the fundamental structure of nature, in both precision and accuracy. Nonetheless there are still many open questions that do not have any answer in the SM.

We have already mentioned that gravity does not have a proper implementation in the SM framework, in fact there is no consolidate way to describe general relativity in

terms of quantum field theory. Still remaining in the domain of gravitation, we have to mention also astronomical observations such as Dark Matter [91] and Dark Energy [71]. The first concerns the observation of some sort of matter that seems to interact only gravitationally with ordinary matter, therefore is possible to observe it only via its effect on the ordinary visible matter, so the name Dark Matter. The second instead require to assume the general relativity, and again, from cosmological measurements, it is found that the universe is accelerating its expansion and therefore the presence of Dark Energy, an elusive form of energy that is responsible of this expansion, has been postulated.

The SM does not have any particle that can be a good candidate for the observed Dark Matter and of course cannot predict the Dark Energy. From the mentioned measurements it turns out that only about 5% of the universe is composed by ordinary matter (included in the SM), while 25% is believed to be Dark Matter, and the remaining 70% Dark Energy.

Along this cosmological problems there is also the evidence of the prevalence of matter compared to anti-matter. In the SM there are three fundamental symmetries: C (charge conjugation), P (parity), and T (time-reversal). These discrete symmetries, essentially, tell us that a certain process happens with the same probability between positive and negative particles, from left-to-right as right-to-left, and if a particle decay in time the products of the decay can form the mother particle reversing the time-line. If all of them are respected by all the interactions there is no way to justify the dominance of matter respect to anti-matter.

Fortunately the EWK interactions violate, even if by a small amount, all the symmetries above. In particular CP-violation [87] is a necessary condition to explain the predominance of matter and the baryogenesis. This violation of CP in the quark sector alone, anyway, is not enough to explain the abundance of matter and therefore a

similar mechanism of leptogenesis is also hypnotize in the lepton sector. This hypothesis is supported also by the recent measurements of neutrino oscillations that are discussed in the following section.

2.2.1 Neutrino oscillations

During the last twenty years neutrino physics has become more and more important, and has produced excellent measures and interesting results. The complete discussion of neutrino physics is behind the scope of this thesis and can be found in Refs. [96, 19], while some of the most interesting experimental results can be found in Ref. [90].

The neutrino oscillations, or some sort of, were first suggested in 1957 [85], but only with the first experimental evidence almost 40 years later, they became one of the main reason to expand the SM scope. In fact to explain the oscillations the neutrino cannot be massless as shown in Sec. 2.1.4, but they need to have a non-zero mass.

Essentially the observed neutrino oscillations are among flavor eigenstates $\nu_e \leftrightharpoons \nu_\mu \leftrightharpoons \nu_\tau \leftrightharpoons \nu_e$, that are also the way neutrinos interact and therefore are detected. On the other hand for the neutrinos propagation we shall consider the mass eigenstates, that we can denote ν_1, ν_2, ν_3. These different eigenstates are mixed according to the follow equation:

$$\nu_\ell = U_{\ell i}\, \nu_i \tag{2.28}$$

where ℓ denotes flavor eigenstates and i mass eigenstates, $U_{\ell i}$ is a unitary matrix that relates the flavor eigenstate ℓ to the mass eigenstate i.

The matrix U can be expressed in different ways, we choose the most simple one with a minimal number of parameters. In the case of Dirac neutrinos we have three

mixing angles θ_{12}, θ_{23}, θ_{13} and one phase δ. Using standard rotations and with the following abbreviations $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$ we can write:

$$U = \begin{pmatrix} c_{13}c_{12} & c_{13}s_{12} & s_{13}e^{-i\delta} \\ -c_{23}s_{12} - s_{23}c_{12}s_{13}e^{i\delta} & c_{23}c_{12} - s_{23}s_{12}s_{13}e^{i\delta} & c_{13}s_{23} \\ s_{23}s12 - c_{23}c_{12}s_{13}e^{i\delta} & -s_{23}c_{12} - c_{23}s_{12}s_{13}e^{i\delta} & c_{13}c_{23} \end{pmatrix} \tag{2.29}$$

on the other hand, in the case of Majorana neutrinos there are at least three angles and three phases, for simplicity the mixing matrix is expressed as

$$U^M = US^M(\bar{\alpha}) \tag{2.30}$$

where U is the same Dirac matrix of Eq. 2.29 and S^M is a vector containing the two extra Majorana phases

$$S^M(\bar{\alpha}) = \begin{pmatrix} e^{i\bar{\alpha}_1} \\ e^{i\bar{\alpha}_2} \\ 1 \end{pmatrix} \tag{2.31}$$

Within his framework several assumptions can be made regarding the mass hierarchy of the three neutrinos or the various mixing angles, in this work only few important aspects of the oscillations will be underlined and the implications that their non-zero value have.

The time propagation of the neutrino mass eigenstate is generally expressed as

$$|\nu_i(t)\rangle = e^{-Ht}|\nu_i(0)\rangle \tag{2.32}$$

or as propagation of the three flavor eigenstates we can write

$$|\nu_\ell(t)\rangle = e^{-E_i t} U^*_{\ell i} |\nu_i(0)\rangle \tag{2.33}$$

Then we can also write the transition probability to find, at any time t, a neutrino $\nu_{\ell'}$ in a beam that started with only neutrino ν_ℓ at $t = 0$:

$$P\left(\nu_\ell \to \nu_{\ell'}\right) = \left| U_{\ell' i} e^{-i(E_i - E_j)t} U^*_{\ell i} \right|^2 \tag{2.34}$$

Assuming that $m_i \ll p$ we can write for the neutrinos $E_i - E_j \simeq \frac{\Delta m^2_{ij}}{2E}$, moreover in case of ultra-relativistic neutrinos $t \simeq L$ where L is the flying distance between production and detection points. Therefore we can write

$$P\left(\nu_\ell \to \nu_{\ell'}\right) = \left| \delta_{\ell'\ell} + U_{\ell' i} \left(e^{-i \frac{\Delta m^2_{ij}}{2E} L} - 1 \right) U^*_{\ell i} \right|^2 \tag{2.35}$$

where we made use of the unitary condition $U_{\ell' i} U^*_{\ell i} = \delta_{\ell'\ell}$

In Eq.2.35 we have already a clear picture of what the observation of neutrino oscillations means for the SM, in particular we can see that the only way to explain non-zero oscillations is that Δm_{ij} has to be different from zero. This implies that only one neutrino is allowed to be massless while the others must have mass. Moreover, we can point out that the probability to observe an oscillation depends on the distance L and the neutrino energy E. Thus experiments at different distances can look for appearance and disappearance of different neutrino flavors from the same beam, or the energy of the neutrino beam can be tuned accordingly to the process we want to investigate between two experiments at a given distance.

2.2.2 Neutrino masses

As already pointed out the original version of the SM does not allow the neutrinos to have mass and, therefore, some sort of extension is needed. As already pointed out in Sec. 2.1.4, the masses of all the other fermions are explained by the Yukawa-coupling of the fermion fields Ψ with the Higgs field ϕ, so it is straightforward to use the same method and add a right-handed singlet for the neutrino ν_R as has been already done for all the other fermions.

This description will raise masses for the neutrinos with the same Yukawa mechanism we saw for the other fermions, but some open questions still hold. In particular, we know that within the same generation the mass difference of the quarks and lepton is in general small, at the most one or two orders of magnitude. From experimental measurement we know that the electron neutrino has a mass of the order of the electron-volt, that is almost six orders of magnitude less than the electron mass. Such a difference could be explained with a different Yukawa couplings but it could also be sign of a different mechanism responsible for the neutrino mass.

2.3 Possible extensions

2.3.1 Dirac mass term

As just mentioned the simplest way to add a neutrino mass term in the SM Lagrangian is to use the Yukawa-coupling as with all the other fermions. In the most general case the mass term in the Lagrangian will be:

$$\mathcal{L}_\nu^D = -\bar{\nu}_{\ell' L}\, M^D_{\ell'\ell}\, \nu_{\ell R} + \text{h.c.} \tag{2.36}$$

where M^D is a complex 3×3 matrix the indices ℓ and ℓ' run over the flavor states.

M^D can be diagonalized using the unitary matrix U in Eq. 2.29 for the left-handed fields and the corresponding V unitary matrix for the right-handed fields:

$$M^D = U^\dagger \, m \, V \tag{2.37}$$

to write

$$\mathcal{L}_\nu^D = -m_i \bar{\nu}_i \nu_i \tag{2.38}$$

here ν_i indicates mass eigenstates.

This mass term has exactly the same for as Eq. 2.27 and also the same properties, in fact in this description the neutrinos are Dirac particles as all the other fermions.

2.3.2 Majorana mass term

This is not the only possibility to write a mass term for the neutrinos. Another open question about the neutrinos nature is if they are Dirac or Majorana particles. The fundamental between the two is that in one case the particle and antiparticle are different entities, while in the other they are equivalent. All the known fermions are Dirac particles since, for example, charge must be conserved so processes like $e^+ \nrightarrow e^-$ are forbidden. For the neutrino sector, instead, there is no known conserved quantum number, so Majorana neutrinos are a possibility.

If neutrinos are Majorana particles then $(\nu_{\ell L})^c = \nu_{\ell R}$, therefore we can write a mass term using only left-handed components

$$\mathcal{L}_\nu^M = -\frac{1}{2} \bar{\nu}_{\ell' L} \, M^M_{\ell' \ell} \, (\nu_{\ell L})^c + \text{h.c.} \tag{2.39}$$

again we can diagonalize M^M using

$$M^M = U\, m\, U^T \tag{2.40}$$

to get the following Lagrangian for the mass eigenstates

$$\mathcal{L}_\nu^M = -\frac{1}{2} m_i \bar{\nu}_i \nu_i \tag{2.41}$$

where $\nu_i = \nu_{iL} + (\nu_{iL})^c$ is the Majorana field and does not conserve lepton number.

2.3.3 Dirac and Majorana mass terms

We can generalized the neutrino Lagrangian to include both type of mass terms including active left-handed fields and sterile right-handed fields:

$$\mathcal{L}_\nu^{D+M} = -\frac{1}{2} \bar{\nu}_{\ell' L}\, M^M_{\ell' \ell L}\, (\nu_{\ell L})^c - \bar{\nu}_{\ell' L}\, M^D_{\ell' \ell}\, \nu_{\ell R} - \frac{1}{2} \overline{(\nu_{\ell' R})^c}\, M^M_{\ell' \ell R}\, \nu_{\ell R} + \text{h.c.} \tag{2.42}$$

or in a more compact form

$$\mathcal{L}_\nu^{D+M} = -\frac{1}{2} \bar{n}_L\, M^{D+M}\, (n_L)^c + \text{h.c.} \tag{2.43}$$

here

$$n_L = \begin{pmatrix} \nu_{\ell L} \\ \\ (\nu_{\ell R})^c \end{pmatrix} \tag{2.44}$$

and

$$M^{D+M} = \begin{pmatrix} M_L^M & M^D \\ \\ \left(M^D\right)^T & M_R^M \end{pmatrix} \tag{2.45}$$

is a symmetrical 6×6 matrix.

We can also define, similarly as we did before, a unitary 6×6 mixing matrix U to diagonalize M^{D+M} and define the mass eigenstates of the six left-handed flavor fields and six right-handed sterile fields, so the mass Lagrangian can be expressed in the familiar form:

$$\mathcal{L}_{\nu}^{D+M} = -\frac{1}{2}\sum_{i=1}^{6} m_i \bar{\nu}_i \nu_i \tag{2.46}$$

and the flavor eigenstates are:

$$\nu_{\ell L} = \sum_{i=1}^{6} U_{\ell i}\nu_{iL} \qquad\qquad (\nu_{\ell R})^c = \sum_{i=6}^{6} U_{\bar{\ell} i}\nu_{iL} \tag{2.47}$$

In this picture the Lagrangian, as for the Majorana neutrino alone, does not conserve the lepton number and therefore the resulting neutrinos can be seen as Majorana particles.

2.3.4 The heavy neutrino and the see-saw mechanism

The see-saw mechanism was first introduced at the end of the seventies, it is based on the Dirac-Majorana mass term and it gives a natural explanation of the small neutrino masses compared to the other fermions.

For simplicity we assume only the case of one generation so in the Lagrangian 2.43 we have

$$n_L = \begin{pmatrix} \nu_{eL} \\ (\nu_{eR})^c \end{pmatrix} \tag{2.48}$$

where only the electron neutrino appears, and

$$M^{D+M} = \begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix} \tag{2.49}$$

is a simple 2×2 matrix. The sum in the Lagrangian 2.46 now has only two indices for the two components of the considered electron neutrino field, hence the mass eigenstates are:

$$m_{1,2} = \frac{1}{2}\left(m_R + m_L\right) \mp \frac{1}{2}\sqrt{\left(m_R - m_L\right)^2 + 4m_D^2} \tag{2.50}$$

We will now formulate the main assumptions of the see-saw mechanism:

- there is no left-handed Majorana mass term, i.e. $m_L = 0$;
- the Dirac mass term is generated by the standard Yukawa-coupling, therefore m_D is of the order of the other fermion masses;
- the right-handed Majorana mass violates the lepton number at a scale much larger than the electroweak scale, i.e. $m_R \equiv M_R \gg m_D$.

It follows that the masses eigenstates are given by:

$$m_1 \simeq \frac{m_D^2}{M_R} \ll m_D \qquad\qquad m_2 \simeq M_R \gg m_D \tag{2.51}$$

therefore the weak eigenstates are:

$$\nu_L = i\nu_{1L} + \frac{m_D}{M_R}\nu_{2L} \qquad\qquad (\nu_R)^c = -i\frac{m_D}{M_R}\nu_{1L} + \nu_{2L} \tag{2.52}$$

In this simple picture the smallness of the neutrino masses is connected to the violation of the lepton number at some high scale, assuming $m_D \simeq m_e \simeq 0.5$ MeV and $m_1 \simeq 1$ eV, such scale is $M_R \simeq \frac{m_D^2}{m_1} \simeq 10^{11}$ GeV. Thus if the see-saw mechanism is realized in nature then:

- neutrinos are Majorana particles,

- neutrinos masses are smaller than leptons and quarks even if realized by the same mechanism,
- the see-saw partner of the known neutrino, the heavy Majorana neutrino, must exist.

This simple model has been widely considered to explain the small neutrino masses and also several extension have been developed, in the next few paragraphs we are going to discuss the one by which this study is ispired, for a complete overview we point to the Refs. [47, 81, 63, 94, 82]. We would like also to mention that CP-violation decay of heavy Majorana neutrinos in the early universe could also be a possible source of the barion asymmetry of the universe.

The model just illustrated, of course, cannot be observed at the LHC because of the high heavy neutrino mass and a lighter mass would change the known SM Lagrangian in contrast with some precise low energy EWK measurements, but this can be avoided if the coupling of the heavy neutrino N_R with the SM bosons is mitigated by some small mixing angles $V_{\ell N}$. In the latter case we can write the following interaction Lagrangians [48]:

$$\mathcal{L}_W = -\frac{g}{\sqrt{2}}\left(\bar{\ell}\gamma^\mu V_{\ell N} P_L N W_\mu + \bar{N}\gamma^\mu V^*_{\ell N} P_L \ell W^\dagger_\mu\right) \tag{2.53}$$

$$\mathcal{L}_Z = -\frac{g}{2c_W}\left(\bar{\nu}_\ell\gamma^\mu V_{\ell N} P_L N + \bar{N}\gamma^\mu V^*_{\ell N} P_L \nu_\ell\right) \tag{2.54}$$

$$\mathcal{L}_H = -\frac{g m_N}{2M_W}\left(n\bar{u}_\ell V_{\ell N} P_R N + \bar{N} V^*_{\ell N} P_L \nu_\ell\right) \tag{2.55}$$

where the heavy neutrino N can either be Dirac or Majorana and we use the notation $P_{L,R}$ for the projection operator already given in Eq. 2.5. We can, therefore, draw the three vertices in Fig. 2.6.

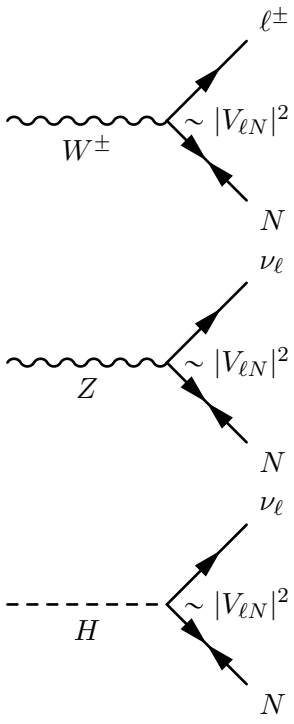


Figure 2.6: Coupling vertices of the heavy neutrino with the known SM bosons: W, Z and Higgs. Each vertex has a factor $|V_{\ell N}|^2$ that suppresses the mixing between SM particles and the N.

As we just mentioned, the heavy partners of the known neutrinos can be either Majorana o Dirac particles, as the light neutrinos. We have also shown in Sec. 2.3.3 that in the most general case the neutrinos will be Majorana particles and violate the lepton number. Hence, from now on, we will refer to the heavy partners as Majorana neutrinos, moreover for simplicity we consider only the lightest of them.

Given the Lagrangians 2.53, 2.54 and 2.55 we can also draw the Feynman diagrams for possible production and decay at the LHC. We will focus on the W production mode and decay because is the one with the larger cross-section, of course the N can be produced also via Z or H boson but, as we know, they have a lower cross-section. Among the production via W boson we have two different diagrams contributing: Fig. 2.7 shows the so-called fusion production, while Fig. 2.8 shows the resonant production. In both

diagrams we consider the lepton number violating decay, in which a same-sign lepton pair is produced in the final state. In principle Majorana neutrinos can also conserve the lepton number and therefore also an opposite-sign pair can be produced, but any final state with two opposite-sign leptons will have a large SM background due to Drell-Yan processes and therefore low significance for a possible discovery. Thus we consider in this work only the same-sign channel in order to keep a good background rejection and signal efficiency.

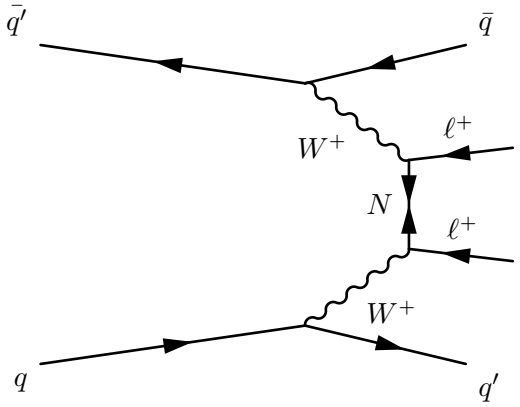


Figure 2.7: Feynman diagram for like-sign dilepton signature via WW fusion in hadronic collisions.

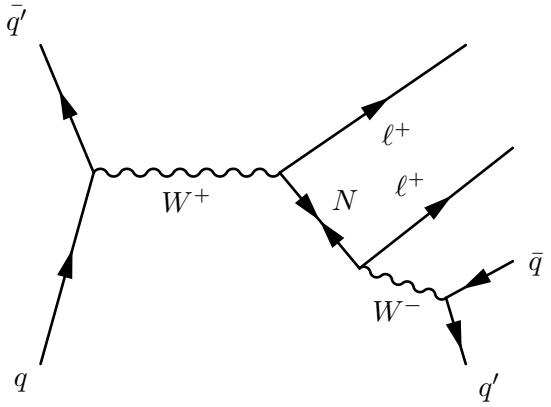


Figure 2.8: Feynman diagram for like-sign dilepton signature via resonant W production in hadronic collisions.

Even if the two processes have similar signatures the first one in Fig. 2.7 has a cross-section that is suppressed by a factor $|V_{\ell_1 N} V_{\ell_2 N}|^2$, while the resonant production shown in Fig. 2.8 has a cross-section proportional to $\frac{|V_{\ell_1 N} V_{\ell_2 N}|^2}{\sum_i |V_{\ell_i N}|^2}$. Thus is the resonant production the one with the larger cross-section and branching ratio.

Chapter 3

LHC and CMS

3.1 The LHC collider

The Large Hadron Collider (LHC) [55] is presently the highest energy collider in the World. Its design started more than twenty years ago and, in of 2012, the accelerator is operating at a center-of-mass energy of 8 TeV. The final goal is to achieve collisions at $\sqrt{s} = 13 - 14$ TeV in 2015.

The start-up luminosity in 2010 was 10^{32} cm^{-2}s^{-1} end reached 5×10^{33} cm^{-2}s^{-1} at the end of the 2011 data taking period. In 2012 the luminosity kept increasing reaching about 8×10^{33} cm^{-2}s^{-1}. The goal is to reach the full design luminosity of 10^{34} cm^{-2}s^{-1} in 2014 after the so called Long Shut-down 1 (LS1) and, after major upgrades during other shut-down periods, pass that value and go as high as 5×10^{34} cm^{-2}s^{-1} for the high-luminosity LHC phase (HL-LHC).

The current energy and luminosity are limited by the magnet performance after the incident in September 2008 in which a magnet quench caused the lost of cryogenic conditions and mechanical damage to 53 magnets; thereafter the energy was reduced to 3.5 TeV per beam during the 2011 run and increased to 4 TeV per beam in 2012.

The design performance of the collider will be reached only in 2015 after a long shut-down, during which several improvements are planned at machine and detector level. The final energy is expected to be 6.5-7 TeV with 2808 bunches per beam and about 1.5×10^{11} protons per bunch. The bunch space will be further decreased from the 50 ns bunch spacing used during the first three years to the nominal value of 25 ns.

The location of the collider is in the same 27 km circumference tunnel used by the Large Electron Positron (LEP) collider that was capable of 100 GeV electron-positron collisions. The LEP program ended in 2000 after having reached its limits. LEP was limited by the synchrotron radiation emitted by the electrons and positrons given by:

$$W = \frac{e^2 v^3}{3\varepsilon_0 c^3 R} \left(\frac{E}{mc^2} \right)^4$$

where e is the charge, v is the velocity, R is the radius of curvature of the beam, E and m represent the energy and rest mass of the particle, c is the speed of light and ε_0 is the vacuum dielectric constant. In the relativistic approximation, $\beta \approx 1$, we can write:

$$W \propto \frac{E^4}{m^4 R}$$

The energy loss for protons is a factor 10^{12} smaller than for electrons and it is therefore feasible to achieve much higher energy in the same tunnel without synchrotron radiation losses. Rather than colliding protons and anti-protons as the Tevatron collider, the LHC collides protons in order to reach higher luminosities since they do not need do be produced and stored.

The disadvantage of this is that the design of the magnets is more complicated. In order to accelerate two particles with the same charge in opposite directions we need

two different vacuum tubes with two different magnetic fields as we can see in figure 3.1.

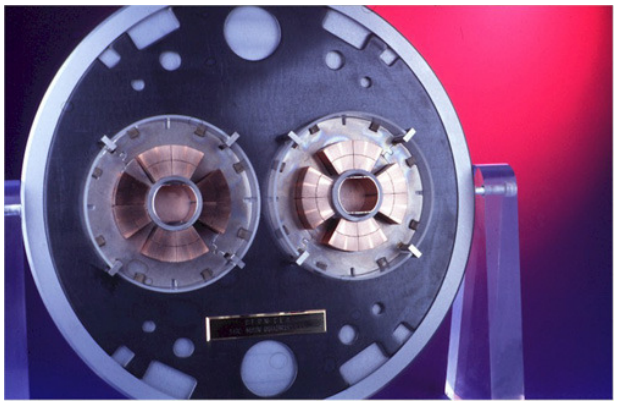


Figure 3.1: LHC section of the two vacuum tubes.

In addition to technical difficulties the choice to use protons, as other not point-like particles, gives more challenges to the physics analysis since the actual collision energy is not the full available energy of the proton beam but roughly 1/3 of it. In more details the actual production rate is related to the particle density function (P.D.F.) which gives the probability of a parton, either gluon or quark, to have a certain energy.

3.2 Main LHC components

The LHC is the 27 km circumference synchrotron that is the last stage in a series of accelerators used to produce the two proton beams. The protons are first produced from a tank of hydrogen and then accelerated through a linear accelerator (Linac) and then they are further boosted in three circular accelerators (the Booster, Proton Synchrotron, and Super Proton Synchrotron) as we can see in Figure 3.2.

These accelerators, operated with conventional magnets, provide an injection energy into the LHC of 450 GeV and, once there, the protons will be accelerated to the TeV scale via Radio-Frequency Cavities (RF) and are kept on the 27 km orbit

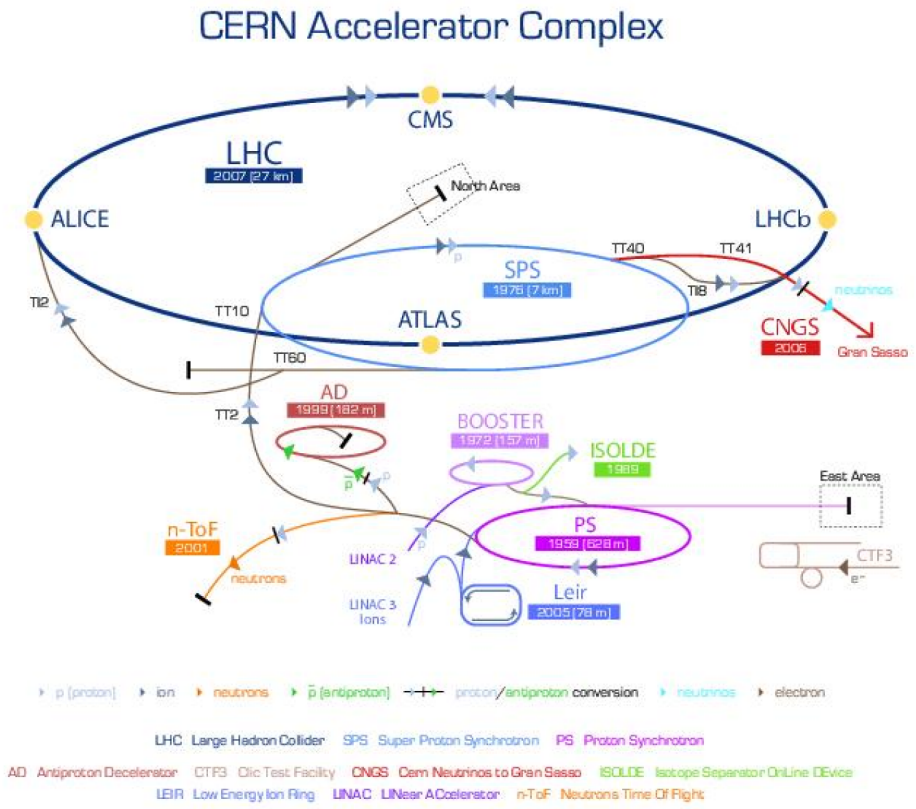


Figure 3.2: LHC injection system.

with superconducting magnets. Using the following equation $p = 0.3BR$, where B is the magnetic field in Tesla, R the radius in km and p the momentum in TeV, we can calculate the magnetic field needed to keep protons of 7 TeV on a fixed orbit. Assuming a constant curvature, the magnetic field that will be needed is about 5.5 T but the shape of the collider is closer to an octagon and therefore the magnetic field required at the curvature points is of the order of 8.4 T. A dipole magnet with this high field would be very energy consuming if built as a conventional magnet using copper wires, so the magnets are built using superconducting wires. However, the challenge in building superconducting magnets is to provide the cooling needed to operate them at 1.8 K, hence by 2008 LHC became the largest cryogenic facility in the World.

As well as dipoles to bend the protons there are also quadrupole, esapole and

octapole magnets to focus and squeeze the beams. A short summary of the LHC characteristics is provided in Table 3.1.

Collision energy	7+7 TeV
Relativistic factor γ	7461
Injection energy	450 GeV
Circumference	27 km
Dipoles magnetic field	8.4 T
Crossing points	4
Number of dipoles	1232
Number of quadrupoles	520
Number of esapoles	2464
Number of octopoles	1232
Luminosity	10^{34} cm^{-2}s^{-1}
Number if bunches	2835
Proton per bunch	1.5×10^{11}
Bunch space	25 ns
Beam crossing rate	40 MHz
p-p collisions per b.x.	20
Beam current	2×0.536 A
Stored energy	2×334 MJ
Beam section at interaction	15 μm^2
Beam lifetime	13 h

Table 3.1: Summary of LHC design characteristics.

The LHC provides p-p collisions to four experiments:

- **CMS** Compact Muon Solenoid
- **ATLAS** A ToroidaL ApparatuS
- **ALICE** A Large Ion Collider Experiment
- **LHC-B** LHC Beauty

CMS and ATLAS are the largest and called "general purpose" experiments, since they are designed to detect all the particles produced during the collisions (expect the elusive neutrinos) and study a wide variety of physics processes. The second two

are dedicated to particular physics studies: LHCb is designed to detect rare decays of the B-mesons and measure CP violation processes and ALICE studies the quark-gluon plasma produced during Pb-Pb collisions. The LHC is indeed capable of accelerating lead ions up to an energy of 2.76 TeV per nucleon. These interactions are useful to understand the properties of a really high energy density state referred to as the quark gluon plasma.

3.3 The CMS detector

CMS [28] is one of the two general purpose detectors. Its structure consists of a series of detectors, as we can see in figure 3.3. This design tries to maximize the particle detection capabilities of the experiment, realizing an almost full 4π radians coverage around the interaction point.

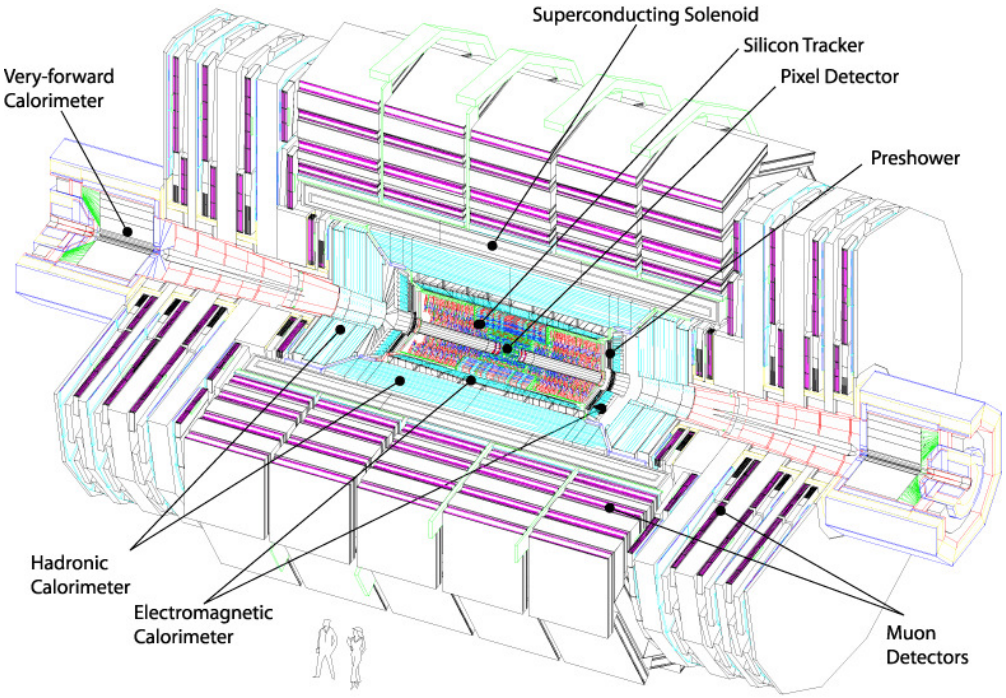


Figure 3.3: Full view of the CMS detector with internal structures.

As revealed by its name, the detector is based on a solenoidal magnetic field of 3.8 T and is designed to have excellent muon reconstruction capabilities. Muons are indeed an important signature in hadron colliders since they may be considered stable at high energy and they have a very long interaction length. These two properties make their reconstruction easier compared to other particles produced in hadronic showers or even electrons, since they escape the high multiplicity zone near the interaction point and can be measured in the outer region of the detector.

The general guidelines taken into account during the design of CMS can be summarized as:

1. an excellent system to detect muons and measure their transverse momentum;
2. as good an electromagnetic calorimeter possible compatible with 1. to measure electron and photons energies;
3. a tracking system able to measure tracks near the interaction point and with high quality vertex reconstruction in agreement with the two previous points;
4. a hadronic calorimeter with the maximum solid angle coverage possible in order to fully contain all the hadronic particles that participated in the event;
5. a solenoidal magnetic field needed to measure the transverse momentum of the produced particles.

Of course, another important parameter to consider was the cost of the each sub-detector and according to the location of each component, magnetic field and radiation damage had to be considered. All of this was realized in a detector of diameter 15 m and length 22 m that weights 12500 t mainly because of the iron yoke that surrounds the magnets to close the field lines of the solenoid.

In order to better describe the detector components it is useful to introduce here the coordinate system that will be used later on: the origin is at the interaction vertex, the z-axis is along the beam line, x points towards the center of the ring and y towards the surface. An alternative to the Cartesian system are the cylindrical coordinates that better suit the detector geometry. In this case z is still along the beam line, r is the distance from the axis and the angle ϕ is measured on the xy plane from the x axis. Another suitable system uses the polar coordinates in which the polar axis is the beam, r is the distance from the center, θ is the zenith angle and ϕ describes the azimuth. Instead of θ we often use the rapidity y defined as:

$$y = \frac{1}{2} \ln \left(\frac{|\overrightarrow{p}| + p_z}{|\overrightarrow{p}| - p_z} \right)$$

where $\overrightarrow{p}$ is the momentum of the particle and p_z is the projection along the z-axis, in the ultra-relativistic approximation ($|\overrightarrow{p}| >> m$) the rapidity is equal to the pseudo-rapidity η, defined by:

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

this new variable is really useful to describe the processes occurring inside the detector.

The CMS detector looks like a cylinder when closed, but its structure is much more complicated than it appears. The central part (*barrel*) is sliced into five wheels and at each end there are three disks of decreasing thickness from inside out. Each wheel and end-cap has a number, the wheel numeration starts from 0 on the central wheel and proceeds to ± 1 and ± 2 according to the positive or negative z-axis. The end-caps are numbered ± 1, ± 2 and ± 3.

If we consider a cross-section of the barrel in the xy plane we can divide it is

a central part inside the solenoid ($r = 0 - 3$ m) and an external part ($r = 3 - 7.5$ m). The inner part hosts the tracking system and the calorimetry while the outer part is occupied by the muon system and the return iron yoke. Outside the end-caps that close the magnetic field there are two additional calorimeters (*very forward calorimeter*) which cover the region down to $\eta = 5.3$. This enables detection of particles down to very low angles so that the total energy of the collision can be well measured.

In the following sections we will briefly discuss all the sub-detectors starting from the interaction point. The magnet, that is not a detector but since it has a fundamental role and will be treated first.

3.3.1 The magnet

As already mentioned CMS uses a solenoidal magnetic field that is generated by a superconducting magnet of diameter 6 m and length 12.5 m, generating a magnetic field of 3.8 T [31]. This field is used to measure the transverse momentum of the particles in the inner part of the detector. The external muon system instead uses the magnetic properties of the iron yoke that surrounds CMS, the magnetic field lines are forced into the iron of the end-caps and barrel and this provides an additional 2 T magnetic field that is used in conjunction with the muon chambers placed in between the iron yokes to independently measure the momentum of the muons (see field map in Fig. 3.5). The field is weaker due to the saturation of the iron, also the direction of the field is reversed going from the inside out as we can see in Fig. 3.4 (a).

The solenoid is cooled to 4.5 K to reach the superconducting phase and uses about 19 kA to produce its magnetic field. Due to the large current several hours are needed for its charge and discharge under normal conditions. This strong magnetic field

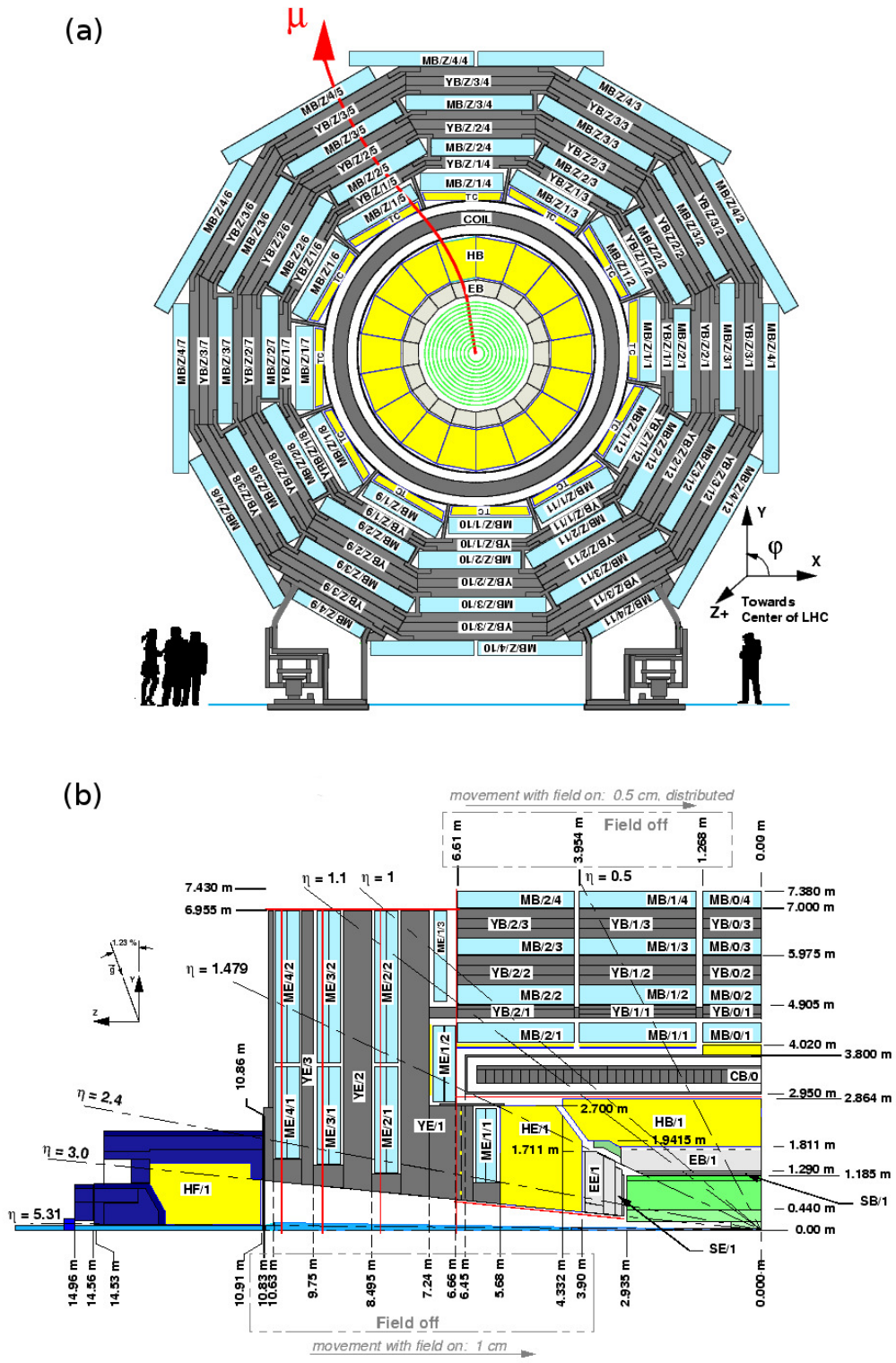


Figure 3.4: Transverse (a) and longitudinal (b) view of CMS.

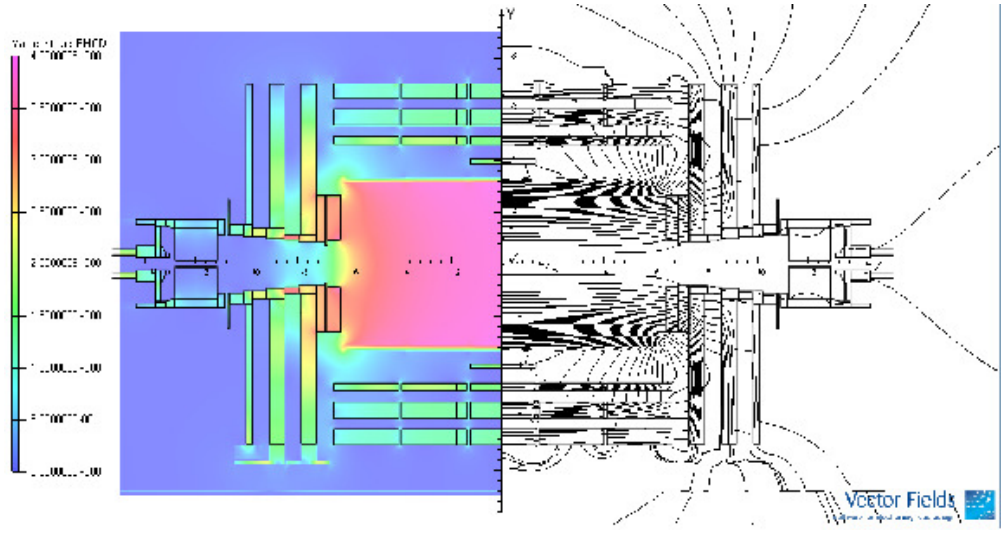


Figure 3.5: (Color) Values of $|B|$ (left) and field lines (right) predicted for a longitudinal section of the CMS detector by a magnetic field model at a central magnetic flux density of 3.8 T. Each field line represents a magnetic flux increment of 4 Wb. .

is a special feature of CMS and gives to the experiment not only its name but excellent capabilities of muon reconstruction and measurement of their transverse momentum (p_T).

3.3.2 The Tracker

The Tracker is located in the inner part of the detector in the 3.8 T magnetic field and closest to the interaction point so it will be exposed to large radiation doses [33].

Being the first detector that all the particles will go trough it has to be lightweight but at the same time radiation hard and has to offer good tracking capabilities in order to give precise p_T measurements, vertex reconstruction and rejection of track generated by secondary collision (also called pileup events).

To satisfy the previous requirements the central tracking system has been built using silicon sensors with two different technologies: *silicon pixels* for the inner part and *silicon strips* for the much larger larger outer part.

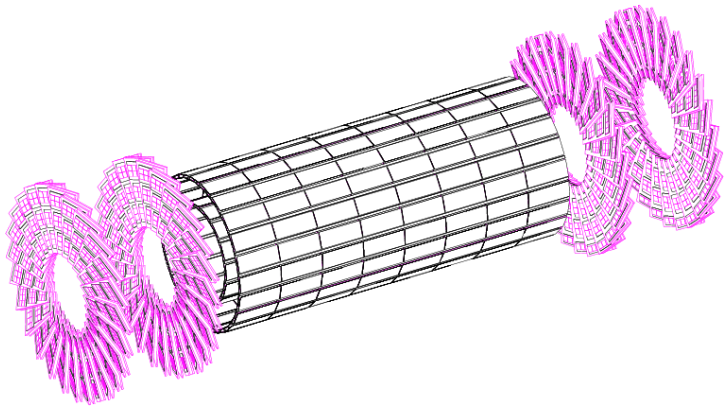


Figure 3.6: Pixel detector with the three central barrel layers and two end-caps on each side.

Due to its fine granularity and 3D hit reconstruction, the pixel detector (shown in Fig. 3.6) gives very good vertex reconstruction that is needed to tag *b-jets* and τ*-jets*. Moreover the pixel detector also provides the seeds used for track reconstruction and therefore it is also important for pileup rejection. Its closest location to the interaction point makes this sub-detector the most susceptible to radiation damage and its replacement it is foreseen after a few years of operation.

The outer strip tracker (Fig. 3.7) is a cylinder 5.8 m long and 2.5 m in diameter, that covers the pseudorapidity range $\eta < 2.5$. It is categorized into four main parts: 4 layers of the tracker inner barrel (TIB) closest to the pixel detector; 6 layers of the tracker outer barrel (TOB); 3 layers of the tracker inner disks (TID) and 9 layers of tracker end-caps (TEC) on each side.

The strip sensors will also suffer from radiation damage during the CMS lifetime and there is already a plan to replace it during a major upgrade before the high luminosity phase of LHC.

The performance of the Tracker is summarized as follows:

- Reconstruction efficiencies of above 98% for muons and 90% for electrons for $p_T >$

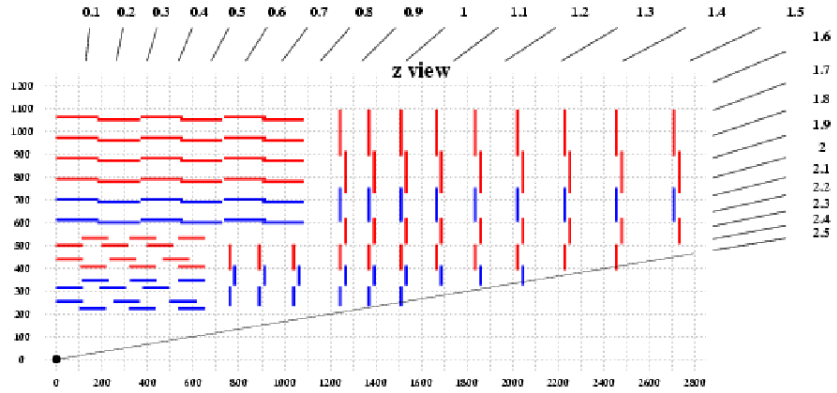


Figure 3.7: Longitudinal view of the tracker strips detector.

1GeV. The tracking efficiency for jets is over 95% for jets with p_T larger than 10 GeV and 85% for p_T above 1 GeV.

- Momentum resolution of $\frac{\delta p_T}{p_T} = (15 \cdot p_T \oplus 0.5)\%$, that degrades in the forward region to $\frac{\delta p_T}{p_T} = (60 \cdot p_T \oplus 0.5)\%$.

- Transverse and longitudinal impact parameter resolutions, necessary for vertex ID, of 35 μm and 75 μm respectively. Thanks to this good performance the b-tagging efficiency between 50 and 200 GeV is 50% and the miss-tag rate is 1%. In addition the good resolution also helps to identify e^+e^- pairs from photon conversion.

3.3.3 The Electromagnetic Calorimeter

Outside the Tracker is the Electromagnetic Calorimeter (ECAL) [29], used to detect photons and electrons. The sensitive parts of the detector are made of lead-tungstate crystals. This choice helped to build a very compact and hermetic calorimeter with good capabilities in measuring the energy of the particles and providing photon and electron identification.

Lead-tungstate ($PbWO_4$) is a dense material (8.278 g/cm^3) that increases the

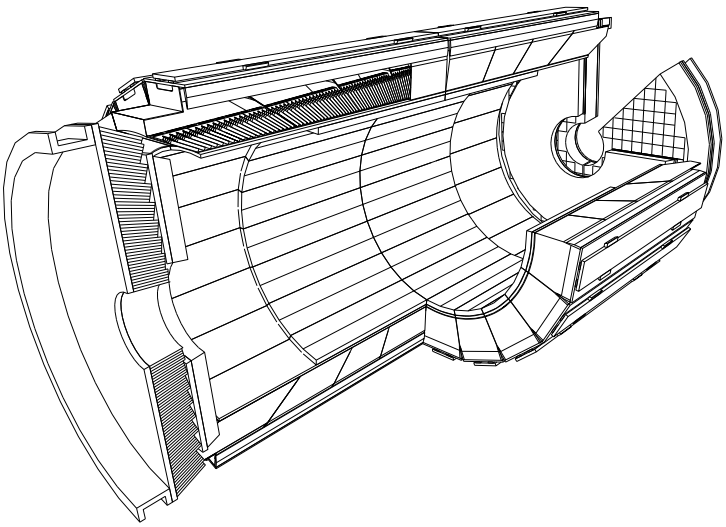


Figure 3.8: ECAL 3D view.

probability of a electromagnetic shower from the incident particles. The scintillation light provided in ($PbWO_4$) has a decay times of the same scale of the LHC bunch crossing (25 ns). The light produced by the active material is collected using avalanche photo-diodes (APD) designed to work inside strong magnetic fields. The ECAL is divided into three main parts: the barrel, endcaps, and pre-shower (Fig. 3.8), and consist of total of 76,832 crystals.

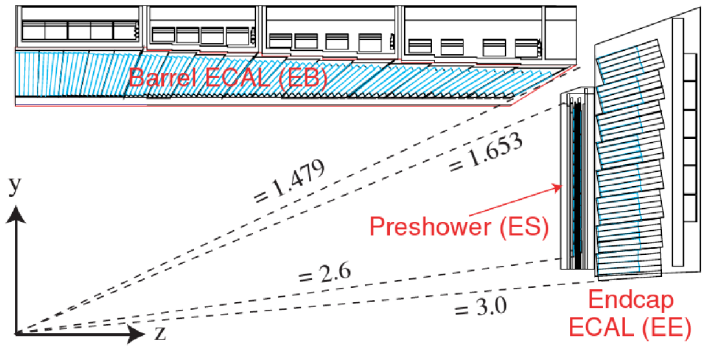


Figure 3.9: Longitudinal view of the ECAL and its components.

Most of the crystals (80%) are located in the ECAL barrel, which covers the region up to $|\eta| < 1.479$ and is at 1.29 m from the center of the detector. The Molière

radius and the radiation length for this material are 2.19 cm and 0.89 cm respectively, allowing a compact design and high granularity. The cross-sectioned area of each crystal is 21.8×21.8 mm^2 (0.0174×0.0174 $\Delta\phi \times \Delta\eta$) and the length is 230 mm (equivalent 25.8 radiation lengths).

The end-caps cover up to $|\eta| < 3$ but due to high radiation and pileup the precision of the measurements quickly degrades after $|\eta| < 2.6$. The crystals in the end-caps are 290 mm long. The end-caps are located at about 3 m from the interaction point in an area with large neutron flux ($\sim 7 \cdot 10^{14}$ cm^2). During the first 10 years of LHC operation they will receive about 5 kGy, and the APD choice was not practical due to high noise. Therefore the crystals here are equipped with vacuum photo-triodes (VPT). The transparency of the crystals is monitored constantly with lasers, since it degrades with radiation and its knowledge is important for the correct calibration of the detector.

As shown in Fig. 3.9 right in front of each end-cap there is a pre-shower detector. If necessary another pre-shower element will be added right before the barrel for the high luminosity phase of LHC. The current pre-shower covers $1.65 < |\eta| < 2.61$ and is composed of two lead plates of length 2 X_0 and 1 X_0 with silicon strips in between to improve the spatial resolution in the end-cap regions.

The energy resolution for electrons and photons between 25 and 500 GeV is parametrized by the following equation:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{a}{\sqrt{E}}\right)^2 + \left(\frac{\sigma_n}{E}\right)^2 + c^2$$

where a is a stochastic term that accounts for the limited photo statistics and shower containment, σ_n is a term due to noise, and c is a calibration constant that will be

monitored using $Z \rightarrow e^+e^-$ events and is less than 1%.

3.3.4 The Hadronic Calorimeter

After the ECAL and before the magnet there is the hadronic calorimeter (HCAL) [30]. Its goal is to contain and measure all the energy related to hadronic activity in the events and also provide a precise measurement of the missing energy ($E_{\mathrm{T}}^{\mathrm{miss}}$). In order to contain as much energy as possible the HCAL is split into three main components: two of which are inside the solenoid and are divided into HCAL barrel (HB) and HCAL end-cap (HE, two per side and covering up to $|\eta| < 3.0$); and one outside the solenoid called the forward calorimeter (HF that covers up $|\eta| < 5.3$).

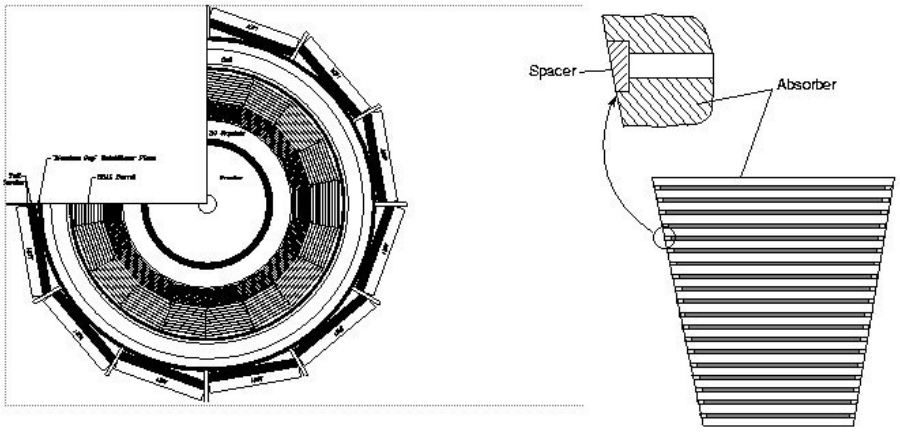


Figure 3.10: Transverse view of HCAL (left) and detailed view of one of the wedges (right).

The HB is located between $r_{int} = 1.806$ m and $r_{ext} = 2.950$ m. It is 8.6 m long and covers $|\eta| < 1.26$, as is possible to see in Fig. 3.10. It is divided into 18 wedges in ϕ and is subdivided into 16 in η (the resulting tower area is 0.084×0.084 $\Delta\phi \times \Delta\eta$). Each wedge is composed of 17 layers of absorber and active elements. Since this part is operated inside the magnetic field the absorber is non-magnetic brass (90% Cu and 10%

Zn) plus some trace of stainless steel for the innermost and outermost layers to provide more structural strength. The brass has a short interaction length ($\lambda_{Cu} = 15$ cm) and low atomic number ($Z_{Cu} = 29$), which are essential characteristics to keep the experiment relatively compact.

The layers of active material are composed of plastic scintillators 4 mm thick. The blue-violet light that they emit is collected by optical guides that also work as wavelength shifters to the green band. Some additional scintillators have been placed outside the solenoid (HO), directly before the muon chambers, to measure the most energetic jets that cannot be fully contained by the HB. Its geometry is constrained by the return yoke and the magnet itself is used as an absorber (which interaction length is described by $1.4/\sin\theta$).

Also within the return yoke is the HE. Its towers are built with the same technology used for the HB and their area is about the same for the region $|\eta| < 1.7$ and is $0.175 \times 0.175\ \Delta\phi \times \Delta\eta$ for $|\eta| > 1.7$.

Outside the magnetic field return yoke, at about 11 m from the interaction point, is the third part of the calorimeter called the HF since is the most forward part of the HCAL. It covers the range $3 < |\eta| < 5.3$ ($r = 0.12$ to 1.5 m). It has a similar segmentation as the rest of the HCAL with 36 wedges in ϕ and 12 segments in η. Since it is located outside the magnetic field it does not have the same non-ferromagnetic requirements as the part inside the solenoid. On, the other hand, it has to be radiation hard and therefore the sensing material used is quartz (fusion-silicon core and polymer hard-cladding). The quartz fibres emit Cherenkôv light detected by photomultipliers located in areas with less radiations exposure.

The energy resolution of the various parts of HCAL can be expressed as:

- barrel: $\frac{\sigma}{E} = \frac{65\%}{\sqrt{E}} \oplus 5\%$
- end-caps: $\frac{\sigma}{E} = \frac{83\%}{\sqrt{E}} \oplus 5\%$
- forward: $\frac{\sigma}{E} = \frac{100\%}{\sqrt{E}} \oplus 5\%$

3.3.5 The Muon System

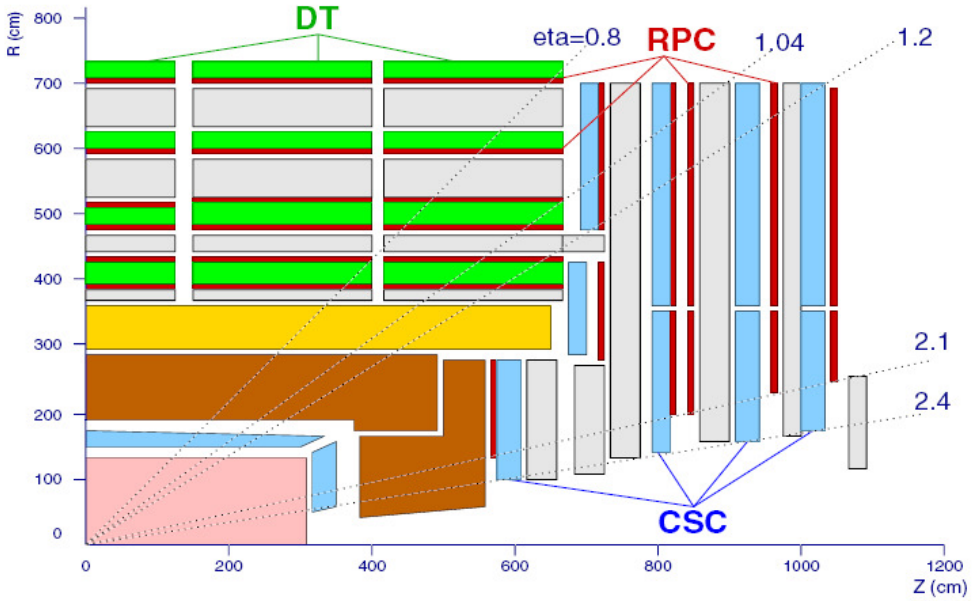


Figure 3.11: Transverse view of the CMS muon system, highlighted in different color DT (green), RPC (red) and CSC (blue).

The muon system [32] is the outermost part of the CMS detector and is divided into three main subsystems according to their location inside the detector and their purpose. The whole system has been designed to achieve the best transverse momentum measurement for muons in a wide range of energies, from a few GeV up to the TeV scale. As already mentioned, muons constitute a clear signature for important physics events so they are used to trigger these events; moreover they are used for Bunch Crossing (BX) identification.

All the muon system sub-detectors are located inside the magnetic field return

yoke (see figure 3.11). The iron of the yoke provides the necessary magnetic field for momentum measurements and screening from other backgrounds. As shown in Fig. 3.11 the system is divided into three different parts with each of them employing a different technology. The three subdetectors are the Drift Tube (DT) system located in the barrel, the Cathode Strip Chamber (CSC) system in the end-cap region and the Resistive Place Chamber (RPC) system. The different technologies used in barrel and end-caps is due to the different magnetic field conditions; the field is more homogeneous in the barrel and less so in the end-caps. The RPC are used mostly for triggering and redundancy purposes.

3.3.5.1 The DTs

The Drift tubes have been chosen to cover a large volume of the detector in a location with low radiation and a relatively small and homogeneous magnetic field. The location of the DT chambers is shown in Fig. 3.4. Their length is 2.5 m corresponds to the thickness of one of the CMS wheels. Each wheel is divided into 12 sectors covering a angle in ϕ of 30°. Each sector is divided into four layers, so there are four concentric cylinders named from inside out MB1, MB2, MB3, MB4. The thickness of the iron yoke between each station is variable to optimize the p_T measurements. Therefore, the dimension of the chambers increases from MB1 to MB4. In the last station not all the chambers have the same dimensions as is visible in Fig. 3.4 (a). From the same picture we can see that the different stations overlap one another to avoid dead regions.

The same principle was applied to design the internal structure of the chambers (see Fig. 3.12). Each chamber has three super-layers each of which has four layers (the only exception is MB4 that omits one of the super-layers). Of the three super-layers two (called ϕ super-layers) have tubes parallel to the z axis and are used to make

precise measurements on the $R\phi$ plane, while the other (called θ super-layer) has tubes orthogonal to the z axis to measure the position along z.

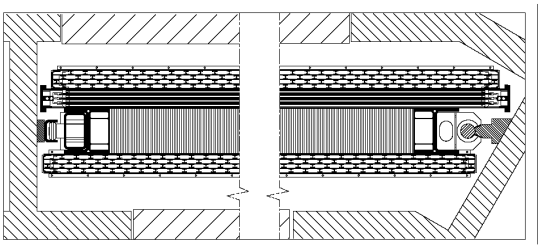


Figure 3.12: Section of a DT chamber in the $R\phi$ plane. The super-layer internal structure is clearly visible and within the ϕ super-layers the displacement of the four layers is visible.

The building block of the DT chamber is the drift tube. It has a rectangular section (shown in figure 3.13) of 42 × 13 mm. At its center there is a 50 μm anode wire while on the shortest side of the cell there is a strip of aluminium that forms the cathode. To achieve the most uniform possible electric field electrode strips are placed right above and below the anode wire. The voltage applied to the anode is +3600 V while the cathodes and electrodes are at −1800 V.

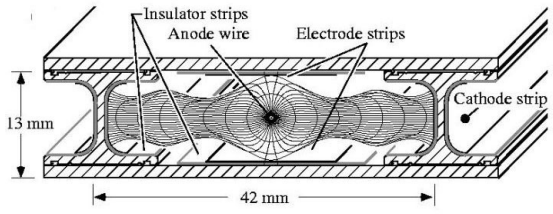


Figure 3.13: Section of a drift tube, showing the electric field generated inside the cell by the cathode strips and anode wire.

Inside the chambers the gas mixture is: 85% CO_2 and 15% Ar. This ensure good quenching properties and low inflammability. The drift velocity is about 55 μm/ns

and the spatial resolution 200 μm (depending on the magnetic field and the incidence angle). The efficiency is very high at 99.8%. The momentum resolution using all the DT stations is between 10% and 20% for muons with $p_T < 100$GeV.

3.3.5.2 The CSCs

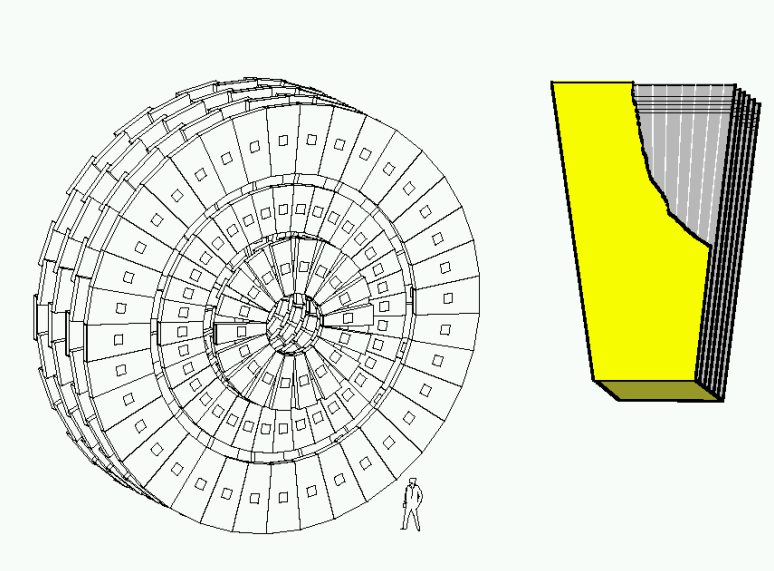


Figure 3.14: On the left is a drawing of the general layout of the CSC chambers in one of the end-caps. On the right, it a simple layout of one of the chambers with radial strips and transversal wires.

In the end-caps the magnetic field is not as homogeneous as in the barrel and therefore it was not possible to use the drift tubes described above. Instead, cathodic strip chambers with shorter drift paths are used. On the right of Fig. 3.14 there is a simple layout of one of the CSC chambers and on the left their location in one of the CMS end-caps is shown. Each end-cap has a substructure defined by three disks of iron magnetic return yoke. These layers provide the necessary bending power (from the induced magnetic field) plus structural support for the CSC chambers. As for the DT, the CSC stations are named from the inside out ME1, ME2, ME3 and ME4, as shown in Fig. 3.15. The last chambers, mounted on the external ring of the last disk (called

ME4/2), are not installed yet. They are currently being built and will be installed during the long shutdown between 2013 and 2015.

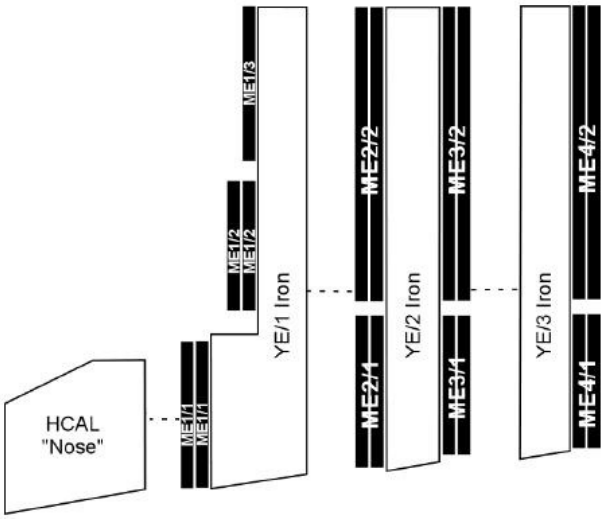


Figure 3.15: Section of one of the CMS end-caps along the θ plane, in white are shown the iron return yokes and in black the CSC chambers with their labelling (The ME4/2 chambers are not completely installed yet and the plan is to complete the system in the first LHC long shout-down).

The dimensions of the chambers varies from station to station with the largest, on the outermost part of the detector, about 3.3 m $\times$ 1.5 m. The ME1 chambers cover 10° in ϕ while all others cover 20°, they are both labelled anticlockwise from 1 to 18 (or 36 for ME1) with the first chamber starting from $\phi = -5°$ and the last finishing at $\phi = +15°$ (or $+5°$). Internally each chamber has six layers of anode wires and cathode strips. The chambers are filled with a mixture of Ar (30%), CO_2 (50%) ,and CF_4 (20%). Each signal is detected by both anode wires and cathode strips so both coordinates r and ϕ are measured.

3.3.5.3 The RPCs

The RPC chambers are used mainly to improve the timing performance of the muon system, but also for redundancy of the system. In the barrel there are six stations

of RPC, two chambers are placed right before and after MB1 and MB2 and one before MB3 and MB4. Four RPC disks are located in the end-caps along with the CSCs.

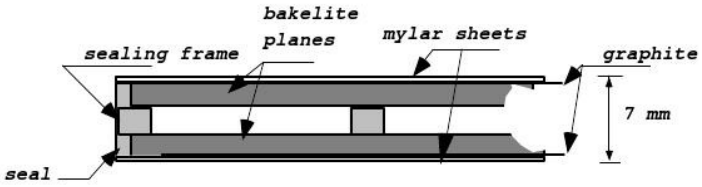


Figure 3.16: Section of the RPC chamber with the various elements.

Figure 3.16 shows a section of a chamber. The RPCs are built with two layers of Bakelite (high resistivity of the order of $10^{11}\ \Omega \cdot \mathrm{cm}$) two millimetres apart covered with graphite. Inside the chamber there is a mixture of freon ($C_2H_2F_4$) and butane (C_4H_10). High voltage is then applied between the layers and when a muon passes trough the electronic signal is read out by aluminium strips isolated from the graphite with Mylar.

The RPCs do not have good spatial resolution but have very fast response time of the order of few nanosecond that is really useful to the trigger system. Moreover, they do not need a complex readout system so they can be used to obtain a, fast, rough estimation of the muon p_T.

3.3.6 Trigger and Data Acquisition System (DAQ)

The trigger system [34, 35] is very important, since at full luminosity in LHC there will be more than 20 interactions per crossing and the bunch crossing will happen every 25 ns. With a frequency of $\sim 10^9$ Hz and $\sim 10^8$ channels to be read the data rate will be in the order of 100 TB/s, this huge amount has to be reduced to something of the order of 100 MB/s in order to be written on tape.

Therefore the target event rate is 100 Hz and in order to achieve it CMS uses a

Data Acquisition system with two trigger levels: the Level 1 Trigger (L1) and the High Level Trigger (HLT). These are discussed in the following sections.

3.3.6.1 The Level 1 Trigger

The level 1 trigger is the first step to suppress the high event rate and produce an output of 100 kHz which is the input capability of the HLT. This rate is set by the average time to read the information and the time for completion of the HLT algorithms.

The L1 trigger processes one event every 3.2 μs which means that with one crossing every 25 ns the events will need to be stored into a buffer (FIFO memory) for 128 bunch crossings (BX). To achieve good performance the L1 trigger has to have a quick response and therefore the reconstruction of the physics objects is only done partially, without all the calibration data (that will need a relatively long time to be read from the calibration database) and only the information from a few sub-detectors is used to make the L1 decision (see Fig. 3.17).

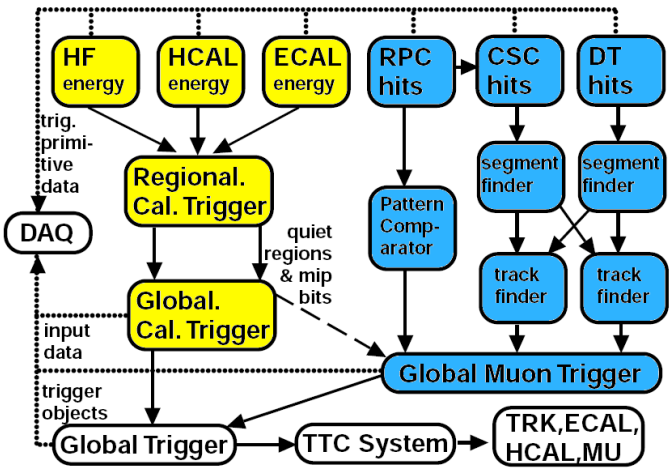


Figure 3.17: Logic flow diagram of the L1 trigger. Yellow represents calorimeter trigger path and blue represents muon trigger paths.

The four major components of the L1 trigger are the Calorimeter trigger, Muon

trigger, Global trigger, and TTC (Timing, Trigger and Control System).

- The calorimeter trigger includes both ECAL and HCAL (with HF). The tower energies are summarized to form a Regional Calorimeter Trigger (RTC). Here the candidate physics objects (such as photons, electrons, taus, and jets) are found and passed to the Global Calorimeter Trigger (GCT). This is the last step of the calorimeter trigger and here visible and missing energies are calculated, the four objects with highest energy are the passed to the Global Trigger.

- The muon system comprises three different sub-detectors already described (DT, CSC, and RPC) and all of them participate to form the L1 trigger. The tracks are built by the sub-system themselves and then sent to the Global Muon Trigger that selects the four highest p_T muons and sends them to the Global Trigger.

- Finally, the eight selected objects are processed in the Global Trigger where the final decision about the event is made. If the event is selected, it is passed to the next level.

- The TTC communicates with the front-end electronics of the various sub-detectors and distributes a precise 40 MHz bunch crossing clock and if the event is accepted it communicates to all sub-detectors the decision of the Global Trigger.

3.3.6.2 The HTL Trigger

After the L1 accept is distributed by the TTC, all the 512 front-end buffers send their data to the processing farm where the HLT algorithms are applied. The input rate is about 100 kHz and the goal is a reduction in rate by a factor 1000. The HLT has to give a quick response, but in this step the time available to make the decision is of the order of a second and therefore more information can be used to make the final

decision. In particular the full granularity of the calorimeters and the information from the tracker is used.

The HLT contains different paths that use single object, multiple objects of the same kind, or so called cross triggers, a mix of different physics objects. The HLT triggers also define the various datasets where the CMS data are stored and they are designed to have as small as possible overlap in order to save disk space. In particular, in this thesis the Single Muon and Double Muon (also known are SingleMu and DoubleMu) datasets are used, as well as the Electron datasets counterpart. These datasets store events triggered by single muon triggers and double muon triggers as well as some cross triggers such as muon plus jet triggers.

During 2011 the run conditions changed dramatically trough the year because of the increasing luminosity that went from $5 \cdot 10^{32}$ cm^{-2}s^{-1} to $5 \cdot 10^{33}$ cm^{-2}s^{-1}, therefore many different trigger paths have been deployed and changed every month. We will focus our attention only and the triggers and datasets used by this analysis

3.3.6.3 Data Acquisition (DAQ)

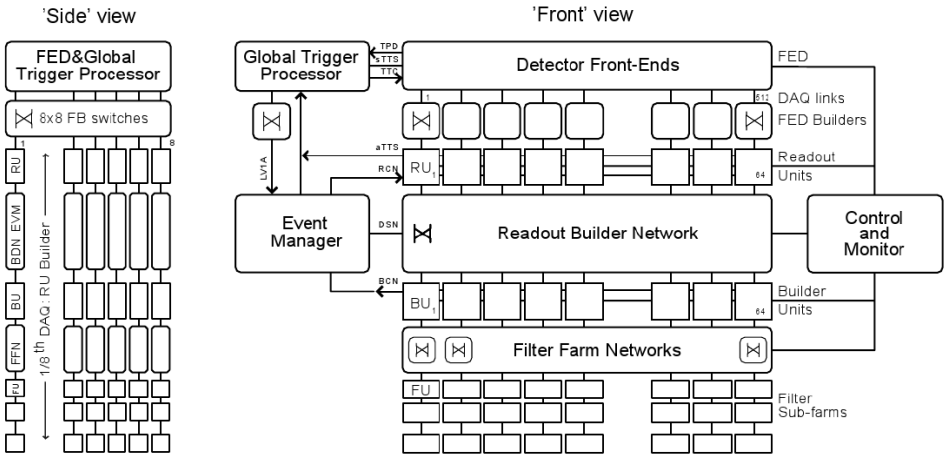


Figure 3.18: Diagram illustrating the architecture of the DAQ.

The DAQ not only contains the L1 and HLT trigger but also other vital components to manage the data flow and build the complete events. A summary of the DAQ structure is visible in figure 3.18 and the function of each unit is briefly described below.

The front-ends detector modules (FED) store the raw data waiting from the L1 decision to accept or reject the event.

The readout system reads the data from the FEDs.

The builder network is a switch-based network that connects the readout and the filter system.

The filter system includes builder units and a filter farm system. It runs the HLT algorithms and builds the event fragments into a full event.

The event manager controls the data flow from readout units to the filter farm and also synchronizes the DAQ with the L1.

The run control and monitor is responsible for the configuration, control and monitoring of all the elements. It also includes also the user interface.

Chapter 4

CMS software and reconstruction

In this chapter the software tools used for the thesis and needed for the CMS data analysis are illustrated.

4.1 Software framework

The CMS software (CMSSW) is the software framework used by the CMS collaboration for calibration, reconstruction, and data analysis purposes.

The CMSSW is a modular software framework written in C++ and Python. Its main element is the Event Data Model (EDM). As the name suggests the center of the framework is the "Event", which must be understood both in a physical way, as a physics object[1] recorded by the detector, and as a C++ object that contains all the interesting information. Each event is composed of different modules that cannot communicate with each other, but can only interact via the event itself, and that can read and add information to the event.

To use the CMSSW we need to provide a configuration file specifying the se-

[1] With physic object we indicate the correspondent C++ object, i.e. the muon object indicates both the C++ object that contains the information of the reconstructed muon as well as the particle with its associated track.

quence of the modules to be called and the parameters given to them. The different modules can be divided in the following categories:

Source: opens the external data which can be raw data from the detector acquisition system or more elaborate data in ROOT format [4]. Source of modules are also used to read simulated data from Monte Carlo or other databases for calibration, alignment, etc.

EDProducer: reads the input data and adds more information to the event. These new objects are saved in the event as RECOnstructed data objects (RECO). Therefore events can contain both raw data and RECO objects.

EDFilter: processes the data in one event and then returns a boolean indicating whether or not to proceed with the analysis of the data.

EDAnalyzer: a module that performs the analysis of the event without adding information to event itself but producing histograms and graphs as independent output.

EDLooper: used to control looping over the event data contained in the input source.

OutputModule: reads the data of an event and then produces an output file (in general in ROOT format) on an hard-disk.

Simulation

The event simulation process consists of two main parts, the event generation and the simulation of the detector. The first part is implemented in CMSSW via an interface that offers different choices of event generators including MadGraph [11],

PYTHIA [88] and ALPGEN [77], the three of which were used for the analysis described in this thesis.

The second part, detector simulation, is done using GEANT4 [27] that gives an accurate description of the CMS detector implementing detector behaviour, track selection mechanisms, readout electronics, etc. The detector behaviour is simulated using files that accurately describe the detector materials and using the known properties of the particle interactions with them. Then the standard reconstruction chain is used to build the tracks and also the behaviour of the readout electronics is simulated to take into account possible inefficiencies and data losses.

Reconstruction

Event reconstruction is defined as the process of building up the physics quantities from the information collected by the detector electronics.

The reconstruction is modularized by the EDProducer and is performed by building a series of independent objects. It can be classified into three major parts: local reconstruction; global reconstruction; and combined physics objects. For example, the local muon reconstruction produces Reconstructed Hits (RecHits) and Segments for a station, which are then combined across the subsystem (for instance all the DT chambers) and, finally, these are used to construct the final muon path implementing information from different sub-detectors.

Analysis tools

After the reconstruction is completed there are several tools, developed by different physics groups, that can be used to apply selection criteria to the final physics objects and eventually select a class of events with certain characteristics. The Physics

Objects Groups (POGs) are in charge of defining the quality requirements for each object in an event. These high-level requirements are implemented by the Physics Analysis Groups (PAGs) in the Physics Analysis Toolkit (PAT). The PAT is a common software package used in all the analysis to perform cleaning[2] and eliminate physics objects (for instance muons) reconstructed with poor quality.

4.2 Simulation

Simulations are useful to understand the detector characteristics and the analysis methods [89]. For the detailed description of the detector and its response GEANT is used, while the simulation of the actual physics process is handled by event generators. As already mentioned, there are various event generators available in high energy physics (HEP) and, according to the QCD factorization theorem [24], the cross section for a given process can be split into two components: the Parton Distribution Functions (P.D.F.s) and the Hard Scattering processes.

The total cross-section can be expressed as:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \hat{\sigma}_{ab \to X} \tag{4.1}$$

where σ_{AB} is the total cross-section for a given process, while $\hat{\sigma}_{ab \to X}$ is the cross-section for a particular hard scattering sub-process $ab \to X$. $f_{a/A}$ and $f_{b/B}$ are the P.D.F.s for each proton with momenta A and B, while x_a and x_b indicate the fraction of the proton momentum carried by the parton. The process $ab \to X$ can be expanded perturbatively

[2]Cleaning is the process used to remove objects that are mistakenly reconstructed twice or overlap each other, for instance a track reconstructed both as a muon and as an electron.

as a sum of the tree level cross-section $\hat{\sigma}_0$ plus higher order terms $\hat{\sigma}_1$, $\hat{\sigma}_2$, etc...

$$\hat{\sigma}_{ab\to X} = \hat{\sigma}_0 + \alpha_S(Q^2)\hat{\sigma}_1 + \alpha_S^2(Q^2)\hat{\sigma}_2 + ... \tag{4.2}$$

The particles resulting from the hard scattering sub-process are hadronized, and then passed to the GEANT detector simulation code.

Parton Distribution Function (P.D.F.)

The P.D.F.s are functions of x and Q^2. The Q^2 dependence is analytically determined from the DGLAP equations [10]. However the dependence on the longitudinal momentum fraction x needs to be determined from data of deep inelastic scattering (DIS) experiments, Drell-Yan (DY) processes, and jet production.

The quark distributions ($f_q(x, Q^2)$) are mainly determined by a global fit to data from DIS and DY processes at leading order (LO), while the gluon distributions ($f_g(x, Q^2)$) are determined from jet production measurements. This approach is only valid in the LO approximation. When next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) are included the global fit has to consider possible mixing of the two distributions and also $\alpha_s(Q^2)$. Currently two major groups provide semi-regular updates to the parton distribution functions: CTEQ [86] and MRST [79].

Currently the DIS and DY datasets allow a precise measurement of the quark distribution with less than 3% systematic error over a large momentum range. On other hand the gluon distribution has a larger uncertainty because is can only be determined directly from jet production at high x and indirectly from scaling violations in the quark distribution at low x.

Hard Scattering Process and Matrix Element (M.E.)

The perturbative expansion of the partonic cross section in Eq. 4.2 helps us to understand the actual formalism. After the identification of the partonic process, we calculate the leading-order matrix element $\hat{\sigma}_0$ and then all the possible initial state P.D.F.s must be considered. At this stage the choice of Q^2 has to be made and the the numerical integration over the x momentum is performed.

Unfortunately, the LO approximation is often not accurate enough, in particular at the LHC. Extra partonic processes may give a non negligible contribution and therefore NLO calculations are necessary to give an accurate description of the initial and final states. Of course there are many issues in extending the calculation beyond LO:

- Virtual and real radiation: adding to all diagrams an additional strong coupling (either quark or gluon). Virtual loops are corrections added within a Feynman diagram, while radiation gives rise to real radiated particles. Both these have to be considered to have a finite cross-section because they have opposite infra-red divergences.

- Scale dependence: an observable which is perturbatively expanded to the order α_s^n depends on the choice of factorization scale. This dependence is still present for the term α_s^{n+1} but the overall effect is generally reduced.

- k-factor: this is defined as the ratio of the NLO and LO cross-sections. This ratio can vary in different kinematic regions even for the same process. In practice this variation is usually small. It also depends on the P.D.F. used to calculate the LO and NLO cross-section as well as the phase space under consideration.

Parton Showering (P.S.)

The parton showering is used to simulate the evolution of the partons from a high energy scale to a soft scale like Λ_{QCD} where the hadron description can be applied.

The evolution uses the same DGLAP formalism mentioned above [10]. The solution of the equations is rewritten using the Sudakov form factor $\Delta(t, t_0)$, for the initial state parton and can be written as [24]:

$$\Delta(t, t_0) = \exp\left[-\int_{t_0}^{t} \frac{dt'}{t'} \int \frac{dz}{z} \frac{\alpha_S}{2\pi} P(z) \frac{f(x/z, t)}{f(t)}\right] \tag{4.3}$$

where t is the hard scale, t_0 is the cutoff scale and $P(z)$ is the splitting function for the branching under consideration. The Sudakov form factor for the final state parton is Eq. 4.3 without the last P.D.F. weighting term (i.e $f(x/z, t)/f(t)$).

As mentioned above, the purpose of the Sudakov form factors is to give the probability for a parton to evolve from an harder to a softer scale without emitting a parton harder than a fixed scale. These form factors depend on: parton type (quark or gluon), the momentum fraction x, the hard and cutoff scales for the process, and the resolution scale for the emission.

In CMSSW the parton showering is performing via the PYTHIA event generator.

ME-PS matching

The parton showering described above gives a good description of soft and collinear emissions, while the matrix element formalism provides the precise fixed order calculation of the hard sub-process. Both of these processes are needed to describe the simulated event, unfortunately the two regimes have overlapping phase-spaces. In order to avoid possible double counting in the overlap region two possible schemes can be

used, called CKKW [25] and the MLM [68]

The matching procedures of the two schemes follow a similar strategy:

1. All relevant cross-sections for a certain process are evaluated with increasing number of accompanying jets (i.e. $pp \to X + n\,\text{jets}$).

2. Hard parton event samples are produced with a probability proportional to the respective total differential cross-section, as determined from the matrix element corresponding to a specific kinematic configuration.

3. Each configuration is accepted or rejected according to a dynamical, kinematics-dependant probability that includes effects from running couplings and Sudakov effects. If the event is rejected, step 2 is repeated with a different parton sample.

4. The parton shower is performed with realistic conditions for each parton. In the MLM procedure this step is performed together with the previous one. In both cases the parton shower cannot produce final states with higher jet multiplicity that would belong to the realm of a different matrix element with extra jets. This avoids double counting of events.

The two merging procedures differ mainly: (a) the jet definition used in the matrix elements; (b) the way acceptance/rejection of the jet configurations is performed; and (c) in the details of the starting condition of the jet vetoing inside the parton showering.

Both algorithms are available within the CMSSW framework and are embedded in the event generators.

Detector Simulation and GEANT

After the simulation of the physics of the event under investigation, it is necessary to characterize the interaction of the generated particles with the detector material. In order to accomplish a reliable detector simulation it is necessary to describe it in terms of size, shape, and kind of material. The XML language is used to create the correct geometry for each subsystem in CMSSW, then the full detector simulation is performed using the GEANT software. Finally the standard CMS reconstruction software is used to complete the Monte Carlo event simulation. As well as the full simulation a fast parametrized version of the simulation called FASTSIM [2] is also available. This fast simulation is useful to speed up CPU intensive event simulations, such as events with high pileup.

4.3 Muon Reconstruction

We have already mentioned in the previous chapters some of the properties of the muons and their importance for the searches and measurements performed at LHC. Therefore, their reconstruction is of fundamental importance for CMS, and in this section we describe how this is performed at different levels following the CMSSW architecture.

4.3.1 Tracking and Tracker Muons

Muons can penetrate relatively large amount of material without losing much energy, and this property allows their reconstruction in two possible ways: inside out, and outside in. The first method starts from the inner tracker detector (pixel and strips) while the second starts from the muon chambers in the iron return yoke. Muons

reconstructed within the CMS Tracker are called "Tracker Muons" while the ones using the muon system outside the solenoid are called "Standalone Muons".

Tracker muons are produced starting from track seeds in the pixel detector, which are then propagated to the rest of the tracker volume. We briefly discuss the CMS tracking algorithm [57] in the following section. The procedure used in the standalone muon reconstruction follows the same general procedure.

The first step is to calculate the equations of motion for a charged particle in a magnetic field [20]. Assuming no electrical field, they can be written as:

$$\frac{d^2\vec{r}}{ds^2} = \frac{q}{p}\frac{d\vec{r}}{ds} \times \vec{B}(\vec{r}) \tag{4.4}$$

the most natural coordinate choice is $ds^2 = dx^2 + dy^2 + dz^2$. The curvature of the particle is given by the charge q over the momentum p times the direction $d\vec{r}/ds$ of the particle trajectory at a given position s. Thus the trajectory can be found (given the exact knowledge of the CMS magnetic field $\vec{B} = B\hat{z}$):

$$x(s) = x_0 + R_H \left[\cos(\phi_0 + hs\cos(\lambda/R_H) - \cos(\phi_0))\right] \tag{4.5}$$

$$y(s) = y_0 + R_H \left[\sin(\phi_0 + hs\cos(\lambda/R_H) - \sin(\phi_0))\right] \tag{4.6}$$

$$z(s) = z_0 + s\sin\lambda \tag{4.7}$$

$$R_H = \frac{p\cos\lambda}{qB} \tag{4.8}$$

where R_H is the radius of the helix, ϕ_0 the azimuthal angle of the reference point with respect to the helix axis and λ is the slope angle ($\arcsin(dz/ds)$).

These simple equations would be valid in vacuum but material effects must be

taken into account in reality. The major effect of a charge particle moving in a dense material is Coulomb scattering, this process is random and the net result is a Gaussian smearing of its path centered around the original trajectory. The energy loss due to this process is given by the Bethe-Block equation:

$$-\frac{dE}{dx} = \frac{4\pi}{m_e c^2} \cdot \frac{nz^2}{\beta^2} \left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 \cdot \left[\ln\left(\frac{2m_e c^2 \beta^2}{I \cdot (1-\beta^2)}\right) - \beta^2\right] \tag{4.9}$$

Equation 4.9 describes the energy loss per unit of length travelled in a material with mean ionization potential I and electron density n, that can be expressed as $N_A Z\rho/M_u$ where N_A is the Avogadro number, Z is the atomic number, ρ is the density, and M_u the Molar mass constant. These losses have to be taken into account by the tracking algorithm when performing the extrapolation of the tracks from one detector layer to the next.

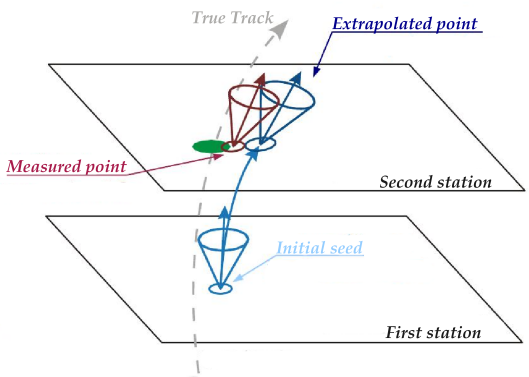


Figure 4.1: Example of the Kalman filter extrapolation procedure. The starting point is a seed located in the first layer, then the filter looks for possible hits in a second layer within a cone that takes into account errors and possible multiple scattering. Once the hit is found, the track parameters are updated and the procedure is repeated to the next layer.

The tracking algorithm consists of four steps: seeding, trajectory building, trajectory cleaning, and trajectory smoothing. The seeds are searched for within the

pixel detector for tracker muons, and in the inner muon chambers for standalone muons. A cone is constructed from the seed (as shown in Fig. 4.1), with the cone defined by $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$. Its aperture is calculated taking into account possible errors, multiple scattering, and, of course, the magnetic field.

The trajectory building takes as input the seeds and uses the Kalman filter [61] technique to fit the whole set of track parameters. The Kalman filter uses an iterative method so that the track parameters are updated at each iteration through the detector. Different propagators are present in the CMSSW framework: the analytic with material propagator, the Runge-Kutta propagator, and the stepping-helix propagator. The first one, the analytic propagator, is the simplest and assumes a uniform magnetic field; the material effects are taken into account only at the end point of the propagation. The Runge-Kutta propagator instead calculates possible material effects from one detector layer to the next and accounts for an inhomogeneous magnetic field. The last one, the stepping-helix propagator, proceeds with finite steps (with a helix length of 5 cm) as long as no material boundary, or magnetic field discontinuity is encountered, or the final point is reached. The track parameters and boundary conditions are updated at each interaction step.

The first two propagators are used in the smaller volume of the tracker, while for the much larger muon system the third one is used. Energy loss and multiple scattering are taken into account using the following fitted function for iron:

$$\frac{dE}{dx} = -\left(11.4 + 0.96|\ln 2.8p| + 0.033p(1 - p^{1/3})\right) \text{MeV/cm} \tag{4.10}$$

and approximate fractional values are used for non-iron material. The multiple scatter-

ing processes are described by:

$$\theta_0 = \frac{13.6\text{MeV}}{\beta cp}\sqrt{x/\chi_0}\,(1 + 0.38\ln(x/\chi_0)) \tag{4.11}$$

where θ_0 is the width of the scattered angular distribution and is related to the muon momentum and radiation length.

The final two steps are cleaning and smoothing. Cleaning is necessary to remove ambiguities among the possible candidates and facilitate the reconstruction of the tracks in successive steps. This is of fundamental importance in the tracker detector, where the tracking is performed in consecutive steps in order to keep the number of fake tracks as low as possible in each step and maximize the total efficiency. The smoothing uses the same Kalman fitting procedure but done backwards this time, to update the track parameters and make a consistency check of the fitted points.

4.3.2 Standalone muons

The local muon reconstruction is performed separately in each CMS muon sub-system: DT, CSC and RPC.

The standalone tracking information is collected mostly from the DT and CSC, while the RPC is mainly used for triggering and gives only approximate spatial measurements. Each DT cell provides a 1-D hit the is not unique because the reconstructed position of the hit depends on the drift time and therefore is symmetric with respect to the wire. This ambiguity is solved when a 2-D segment is formed within each superlayer; the hits from different layers are used to form a segment in the ϕ and the θ plane. At chamber level, 2-D segments are combined to form 3-dimensional segments (sometimes called 4-D segment), which are then used in fitting standalone muon tracks. A

CSC chambers have 6 layers, each containing strips and wires. A CSC hit is built by the intersection of a cluster of strips and a wire group. The local segment is formed by fitting at least three aligned hits from four different layers.

Using the local chamber segments as building blocks, the standalone muon track is made. The standalone fitting procedure consists of three main steps:

- The initial candidates are pairs of chamber muon segments. The reconstruction starts from the innermost layers of the muon system and collects the relevant segments in a region of interest within a cone ΔR given by the direction of the seeding segments and the associated errors. The seeding algorithm also calculates a first estimate of the muon p_T using the found segments and the equation $p_T \times \Delta\phi = \int Bds$, where $\Delta\phi$ is the difference in azimuthal angle between the two segments.

- The Kalman filter is used to fit all the segments and update the tracking information.

- Finally, cleaning is performed. Some of the tracks may share parts of the local chamber segments and to ensure track quality only the ones with the smallest χ^2 are kept.

4.3.3 Global muons

The global muon reconstruction is the highest level in the muon tracking, combining the information from the tracker and standalone muons. The matching starts from standalone muons. A region of interest is defined by the track parameters, and the matching algorithm loops over all possible candidates among the tracker tracks. The selection of a region of interest is necessary because of the high track multiplicity inside the tracker volume. The region of interest is defined by the following parameters:

- a cone in η and ϕ of the tracking region is estimated taking into account also the errors on these two parameters in the muon system;
- the p_T of the considered track has to be larger than 60% of the standalone muon p_T;
- the primary vertex or beam spot is used to define the origin of the track;
- a certain Δz and Δr spread with respect to the origin is allowed, where z is the longitudinal impact parameter and r is the transverse impact parameter.

After the region of interest is defined, the matching process takes place. The matching is performed by comparing the trajectory-state-to-surface of the considered tracks at the same reference plane. The reference plane is chosen by minimizing the covariant error matrix of the propagated track parameters. After propagating the two tracks to the same surface the pair with the best χ^2 is selected.

Once the matching process is done, all the hits (tracker and standalone) are refitted. This global fit is much faster than the previous fit because no pattern recognition is required. The algorithm minimizes only the χ^2 of the global muon, and for high energies only a subset of the standalone muon hits will be used to maintain a good momentum resolution for TeV muons.

Soft muons, typically below 6 GeV, do not have enough energy to leave any hits in the muon system and so do not have an associated standalone track. In order to keep a good efficiency for low p_T muons, these tracks are only reconstructed using the tracker tracks along with energy deposit information from the calorimetry. Of course, no global matching procedure is performed and the determination of the track parameter relies only on the tracker.

4.4 Jet Reconstruction

The cross-section for multi-jet production at the LHC is much higher than that of many of the interesting signal events. Therefore jet reconstruction is important to fully understand the backgrounds in most physics studies. In this thesis we are studying in particular events containing two muons and at least two jets. Even though the multi-jet background is suppressed by requiring a same-sign muon pair, good jet reconstruction is necessary to control and understand the residual background. Since a jet is a complex object consisting of the hadronization products of the hard parton scattering as well as some decay photons and leptons, the jet algorithm relies on the calorimetry (HCAL and ECAL) and the tracking information.

4.4.1 Calorimetric tower

The basic unit of the jet reconstruction is the calorimeter tower. In the ECAL the granularity of a cell is $\Delta\phi \times \Delta\eta = 0.0174 \times 0.0174$ compared to 0.087×0.087 in the HCAL. Since the HCAL has a larger granularity a tower is defined by the signal in a cell of the size of the HCAL granularity. The signal is collected for each readout cell with the requirement of passing an online zero-suppression threshold plus noise and pedestal thresholds. The electronic signal collected within a tower represents the energy deposit of that tower. The jet algorithm typically selects towers with at least $E_T > 0.5$ GeV or $E > 0.8$ GeV. These energy deposits are treated as massless particles in the reconstruction with energy determined by the measured deposit and direction given by the vector pointing from the primary vertex to the center of the tower.

4.4.2 Anti-k_T cone jet clustering algorithm

The are two main classes of jet clustering algorithms: the sequential recombination algorithm such as k_T [58] and Cambridge/Aachen [56], and the cone jet algorithms [53, 54]. The two differ mainly in the regularity of the boundaries of the resulting jets; these are sensitive to non-perturbative effects in the hadronization process and contamination of pileup events.

In this thesis we use the anti-k_T algorithm, which is also the standard algorithm used in the majority of the CMS analyses. However, other algorithms are also used for some CMS analyses and cross checks. In CMS there are two versions of cone jet algorithms: midpoint cone and siscone [59]. These have been studied for a long time; they are infra-red and collinear (IRC) safe, but soft radiation can result in irregularities in the boundaries of the final jets.

Algorithms, such as k_T, can simplify theoretical calculation and eliminate part of the effects due to underlying events. The equations used to construct the jets are:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \tag{4.12}$$

$$d_{iB} = k_{ti}^{2p} \tag{4.13}$$

where d_{ij} and d_{iB} are called resolution variables, they can be interpreted as the distance between the two entities i and j, and the distance between i and beam line B. k_{ti} is the transverse momentum of i, $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ with y_i the rapidity and ϕ the azimuthal angle, and R is the radius parameter of the jet typically set to be 0.5. p governs the power of the energy versus the geometrical scale Δ_{ij}.

The clustering algorithm loops over all the detector towers looking for the

"best" ij pairs that minimize d_{ij}. As the process goes on used towers are removed until none are left and all of them are clustered into jets. The parameter p distinguishes between three different types of recombination algorithm: for $p = 1$ we have the inclusive k_T jet algorithm; for $p = 0$ the Cambridge/Aachen algorithm; and for $p = -1$ the anti-k_T algorithm that is currently the official one for CMS and the one used in this work.

Taking the particular form of Eq. 4.12 for $p = -1$, we can see how the distance d_{ij} for soft jet constituents will be larger and therefore the clustering will be dominated by the hard ones. Taking as an example an event with a few well defined hard constituents from the hard interaction and some soft constituents from the underlying event, it is easy to see that the algorithm will preferentially cluster the energy deposit around the hard constituents rather than merge two soft ones.

If a second hard constituent is present the algorithm behavior is described below in three different cases:

- $\Delta_{12} > 2R$: the soft constituents are clustered around the two main components within a radius R, resulting in two perfect conical jets.

- $R < \Delta_{12} < 2R$: in this case the clustering process forms a conical jet around the harder component. If none of them prevails then both jets will not be conical but will have a boundary line b defined by $\Delta_{1b}/k_{t1} = \Delta_{2b}/k_{t2}$

- $\Delta 12 < R$: the two constituents will form a single jet.

The key feature of the anti-k_T algorithm is that the soft radiation does not modify the shape of the jets. The shape of the boundary is approximately modified by hard radiation but it is not affected by the presence of soft radiation due to pileup events.

4.4.3 Jet energy correction and jet energy scale

The reconstructed jets are affected by many factors that can results in an energy miss-measurement, such as:

- non-uniform and non linear detector response;
- electronics noise;
- pileup effects;
- unclustered energy deposits.

These effects are corrected using the following equation:

$$P_i^{corrected} = C(P_T^{raw}, \eta) P_i^{raw} \tag{4.14}$$

where the correction factor C is a function of the measured, yet uncorrected, momentum and η. The correction is done in three levels: Monte Carlo truth calibration, relative jet energy scale calibration and absolute jet energy scale calibration. In CMS the are three approaches to the jet energy reconstruction and calibration: calorimeter-based, "jets-plus-track", and "particle flow" (PF) [43]. All of them use the anti-k_T algorithm with the cone size fixed at 0.5. In this thesis the PF jets are adopted.

Monte Carlo truth calibration The MC simulation is used to study the detector response to a known process. In CMS the events are generated with PYTHIA then the same algorithms are used to build the so called "gen jet" (using generator level information) and "reco jets" (using the simulated detector response). The jets of these two families are matched using a $\Delta R < 0.25$ cone, and then the difference between the gen jet momentum p_T^{gen} and reco jet momentum p_T^{reco} is calculated. The ratio p_T^{reco}/p_T^{gen}

is used to calculate the correction in η and p_T^{gen} bins. This factor for particle flow jets with $p_T > 30$ GeV ranges from 1.1 in the barrel region to 1.2 in the forward region.

Relative jet energy scale To accurately calibrate the jet energy we cannot rely only on Monte Carlo simulations but also use a data-driven techniques called dijet balance method [38]. In principle in a dijet event the two jets should exactly balance each other, and using this property one can select a well defined jet ("tag") and calibrate the other jet ("probe") accordingly. The event selection is done by selecting events passing HLT with uncorrected average momentum of the two jets given by $p_T^{dijet} = (p_T^{tag} + p_T^{probe})/2$. The tag jet is required to be in the barrel because of the more uniform detector response there. The two jets must be back-to-back ($\Delta\phi_{ij} > 2.7$) and if a third jets is present its momentum should be lower than the average of the dijet pair ($p_T^{3^{rd}jet}/p_T^{dijet} < 0.2$) to avoid residual radiation contamination.

To calculate the correction, two variables are defined: the imbalance quantity

$$B = \frac{\left(p_T^{probe} - p_T^{tag}\right)}{p_T^{dijet}}, \tag{4.15}$$

and the relative response of the probe with respect to the tag jet

$$r = \frac{2 + \langle B \rangle}{2 - \langle B \rangle}. \tag{4.16}$$

Both quantities are binned in η and p_T^{dijet}. In the limit where p_T^{dijet} is small $r = \langle p_T^{probe} \rangle / \langle p_T^{tag} \rangle$. The relative correction for a given bin is defined by the inverse of the relative response or, eventually, as a second order polynomial of the logarithm of the

average momentum of the probe by fitting every pseudo rapidity bin:

$$Rel\left(\eta, \langle p_T^{probe} \rangle\right) = \frac{1}{r\left(\eta, \langle p_T^{probe} \rangle\right)} = a_0(\eta) + a_1(\eta) \ln p_T^{probe} + a_2 \ln^2 p_T^{probe} \tag{4.17}$$

Absolute jet energy scale To correct for possible systematics concerning the whole detector, another data-driven tagging technique is used. In this case the selected processes are the Compton scatterings $q+g \rightarrow q+\gamma$ and $q+q \rightarrow \gamma+g$ [39]. The momentum balance between the jet and the photon (measured by the ECAL) is used. The selection criteria to ensure clear back-to-back γ + jet events in a region with uniform detector response are:

- photon and jet must be in the central region defined as $|\eta| < 1.3$,
- momentum of the photon must be larger than 15 GeV,
- the photon candidate must be isolated in the HCAL, ECAL, and tracker, and it must have a shower shape consistent with a photon,
- back-to-back requirement of the two objects (i.e. $\Delta\phi_{\gamma,jet} > \pi - 0.2$),
- rejection of additional jets in the event if $p_T^{2^{nd}jet} > 0.2 \cdot p_T^{\gamma}$ or $p_T^{2^{nd}jet} > 5$ GeV.

For this calibration two different methods can be used, one uses the missing E_T projection fraction (MPF) and the other uses p_T balance. The MPF method is based on the fact that there is no real $E_{\mathrm{T}}^{\mathrm{miss}}$ in these events and the photon should perfectly balance the recoiling jet. This condition can be formalized with the following equations:

$$\vec{p_t}^{\gamma} + \vec{p_T}^{jet} = 0 \tag{4.18}$$

$$R_{\gamma} \cdot \vec{p_t}^{\gamma} + R_{recoil} \cdot \vec{p_T}^{jet} = -\vec{E}_{\mathrm{T}}^{\mathrm{miss}} \tag{4.19}$$

where R_γ and R_{recoil} are the detector response for the photon and the jet. R_γ is obtained independently using $Z \to e^+e^-$ or $\pi^0/\eta \to \gamma\gamma$ events, and therefore R_{recoil} can be derived as:

$$R_{recoil} = R_\gamma + \frac{\vec{E}_{\mathrm{T}}^{\mathrm{miss}} \cdot \vec{p_t}^{\,\gamma}}{|\vec{p_T}^{\,\gamma}|^2} \equiv R_{MPF} \tag{4.20}$$

This derivation is valid if there are not secondary jets or unclustered energy. In case more jets are present, the identity between the response of the leading jet and the R_{recoil} still holds if the other particles or unclustered energy have similar response, or are perpendicular to the photon axis.

The p_T balance method is similar to the one already discussed for the MC correction. The response is calculated as the ratio between the momentum of the jet and the momentum of the photon ($R = p_T^{jet}/p_T^\gamma$). This approach is more sensitive to systematics than the MPF method. Moreover the MPF method is well suited to PF jet because of the excellent PF $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution. Therefore the, MPF is the main method used by CMS to measure the energy resolution.

4.4.4 *b*-tagging

In many interesting physics processes (for instance $H \to b\bar{b}$) b quarks play an important role and their proper identification is therefore very important. The CMS pixel and tracker detectors were built with that purpose in mind and several b-tagging algorithms have been developed in CMS. They all use the long lifetime of the b and the high track multiplicity associated with its hadronization. The tracks are indeed the most important element of these algorithms and they have higher quality requirements than general tracks [37]:

- total number of silicon hits (strip + pixel) $\geqslant 8$,

- number of pixel hits $\geqslant 2$,
- transverse impact parameter $d_{xy} < 0.2$ cm,
- longitudinal impact parameter $d_z < 17$ cm,
- transverse momentum $p_T > 1$ GeV,
- $\chi^2/ndof$ of the track fit < 8,
- distance from the jet axis $\Delta R < 0.5$.

In the following we briefly describe some of the b-tagging algorithms used. The idea behind the taggers is to reconstruct secondary vertices where the b decays or to use the impact parameter (IP) of the resulting tracks [42].

The IP method uses the distance of closest approach to the primary vertex, also called impact parameter (IP). The IP is measured by the pixel detector, which has excellent three-dimensional capabilities. Another observable is the IP significance defined as the IP divided by its resolution (IP/σ_{IP}). In CMS there are two different taggers that use this methodology, the Track Counting High Efficiency (TCHE) and the Track Counting High Purity (TCHP). Both of them rank the IP significance of tracks in the jets and count the number of tracks (N) exceeding a given threshold; the TCHE fires if $N \geqslant 2$ while the TCHP requires $N \geqslant 3$.

The secondary vertex approach has its own category of dedicated taggers in CMS. The general algorithm is the same as the used to find the primary vertex. The Simple Secondary Vertex (SSV) tagger uses the significance of the three-dimensional flight distance as a discriminating variable. Also in this case we have the high efficiency (HE) and the hight purity (HP) version of the tagger; the first requires at least two tracks to be associated to the secondary vertex and the second at least three.

Other algorithms identify the b hadrons via their semileptonic decays, looking for electrons and muons associated with a jet. The discrimination is achieved using the p_T of the lepton compared to the jet and their impact parameter. Other jet probability algorithms combine information from all selected tracks in the jet.

4.5 Missing Transverse Energy

Although CMS was built to detect most of the known particles and cover the largest possible solid angle, there are weakly interacting particles (such as neutrinos) that can easily escape the detector without leaving any signature. However, these particles can be indirectly observed by the presence of missing transverse energy in the events. The conservation of momentum implies that

$$\sum_{i=1}^{N} \vec{p_T}^{\,i} = 0 \tag{4.21}$$

and, therefore, events with non zero sum must have a momentum imbalance due to elusive particles, resulting in the presence of missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$).

The principle used to measure the $E_{\mathrm{T}}^{\mathrm{miss}}$ [52] is simple to illustrate. The missing energy vector is calculated from the negative energy sum of the calorimeter deposits:

$$\vec{E_{\mathrm{T}}^{\mathrm{miss}}} = \left(\sum_{i=1}^{N} -E_i \sin\theta_i \cos\phi_i \right) \hat{x} + \left(\sum_{i=1}^{N} -E_i \sin\theta_i \sin\phi_i \right) \hat{y} = E_{\mathrm{T}}^{\mathrm{miss}\,x} \hat{x} + E_{\mathrm{T}}^{\mathrm{miss}\,y} \hat{y}. \tag{4.22}$$

Unfortunately at the LHC the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation is complicated by pileup effects, and in particular for CMS the bending effects of the magnetic field have to be taken in account, as well as the difference in detector response between the HCAL and ECAL to pions and photons. In CMS the $E_{\mathrm{T}}^{\mathrm{miss}}$ measured by the calorimeters is subject to two types

of corrections called Type I and Type II. The first corrects for effects due to muons and taus, while the second accounts for unclustered and out-of-cone energy [40].

Type I corrections are necessary to account for the different properties of the muons and the tau showers. Muons are minimum ionizing particles and leave only a small fraction of their energy in the calorimeter resulting in extra missing transverse energy. Nonetheless their momentum can be measured with good precision by the tracking and muon systems. Therefore the correction is applied by removing the muon contribution to the calorimeter energy and adding the muon momentum directly into the calculation of the missing transverse energy: $\vec{E}_{\mathrm{T}}^{\mathrm{miss}\,corr} = \vec{E}_{\mathrm{T}}^{\mathrm{miss}\,uncorr} - (\vec{p_T}^{\mu} - \vec{E}_{\mathrm{T}}^{\mathrm{miss}\,calo\mu})$. Also the τ leptons may cause miss-measurement of the $E_{\mathrm{T}}^{\mathrm{miss}}$. In this case the correction has to take into account the different topology of a τ decaying to hadrons compared to a normal jet. τ-jets have, in general, lower multiplicity compared with a standard jet with the same energy.

Type II corrections are done after type I correction and can be described by the following equation:

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}\,typeII} = \vec{E}_{\mathrm{T}}^{\mathrm{miss}\,typeI} + c \times \sum_{i}^{N} \left[\vec{p_{T_i}}^{IC7} - \vec{p_{T_i}}^{IC5} \right] \tag{4.23}$$

where $\vec{p_{T_i}}^{IC7}$ and $\vec{p_{T_i}}^{IC5}$ are the transverse momenta of jet i reconstructed using the jet cone size $\Delta R = 0.7$ and $\Delta R = 0.5$ respectively, and c can either be a constant or be a function of momentum and pseudo-rapidity.

In CMS the particle-flow reconstruction technique can be also used to reconstruct $E_{\mathrm{T}}^{\mathrm{miss}}$. The PF algorithm aims to provide a complete and unique description at the level of individually reconstructed particles, making use of all the information collected by the CMS sub-detectors. In principle the missing energy can be easily com-

puted within the PF framework adding the momentum of all the particles, as if they were MC truth particles, except that the PF reco objects have a certain efficiency, fake rate, and finite resolution. Of course, the reconstructed objects including muons, electrons, photons (either unconverted or converted), charged hadrons, as well as stable and unstable neutral hadrons must be properly calibrated to obtain a good $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution. So far the PF algorithm has proven to have superior performance in particular with jet reconstruction as well as $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution and we have decided to adopt it in this analysis.

4.6 Electrons and Photons

Since electrons and photons both produce showers in the ECAL their reconstruction is closely related, and relies strongly on the quality of the CMS electromagnetic calorimeter. About 94% of the energy of an incident photon or electron is deposited in a 3×3 ECAL group of cells and 97% in a 5×5. Nevertheless, the material present in front of the ECAL can result in bremsstrahlung and photon conversion, while the magnetic field causes spread in ϕ of the energy lost by the electron travelling trough the tracker.

The general reconstruction starts from the clustering of the energy released in the ECAL. Two super-clustering algorithms are used to identify and correctly reconstruct the showers. For the ECAL barrel the so called "hybrid" algorithm is used and for the end-caps the "island" algorithm is used. The concept of the clustering algorithms is similar to the one we have already described for the jet reconstruction; the algorithm starts with a seeding crystal with an energy above a certain threshold, then it merges the energy from nearby crystals into the super-cluster avoiding deposits from other particles and noise.

The ECAL fine granularity allows a precise measurement of the position of the particle that hit the crystal. This information is very useful, for example for photon conversion studies. The simplest method to compute the position x of the incident particle would be to weight the position x_i of the crystal by its energy; this, however, may bias the final measurement because of the exponentially decreasing behaviour of the energy with lateral distance. Therefore, the solution is to take the logarithm of the crystal energy and to use that as a weight: $W_i = W_0 + \ln E_i/E_{tot}$, where E_i is the energy of the crystal, E_{tot} is the total energy of the cluster, and W_0 is the smallest fractional energy that a crytal can contribute to the position measurement.

4.6.1 Electron reconstruction

The electron reconstruction makes use of both tracking and calorimeter informations, the relative abundance of informations allows different techniques to be adopted in order to maintain good efficiency and avoid possible fakes.

The electron reconstruction can either start from an ECAL super-cluster, as described in the previous subsection or from a tracker track; in the latter case the tracking procedure is similar to the one already described for muons. The ECAL seeded reconstruction is optimized for electrons with $E_T > 5$ GeV (mainly coming from Z and W decays), whereas for low p_T electrons that may not leave a good signal in the ECAL, the tracker is the most important source of information if not the only one.

The main difference for the tracking, when dealing with electrons as compared to muons, is the non Gaussian energy loss due to bremsstrahlung, This non Gaussian behaviour makes the Kalman filter very inaccurate and therefore, for the electrons, a Gaussian-Sum Filter (GSF) [36] is used instead that essentially uses a mixture of many Gaussians rather than only one.

After the tracking is done and the seeds have been chosen track-clustering matching is performed. In this phase of the reconstruction there are many criteria that must be met to ensure good quality of electrons.

- Energy-momentum matching between super-cluster and track $E/p < 3$.
- η, ϕ geometrical matching: $|\eta_{sc} - \eta_{trk}^{extrap}| < 0.1$ and $|\phi_{sc} - \phi_{trk}^{extrap}| < 0.1$, where sc indicates the super-cluster position, and η_{trk}^{extrap} with ϕ_{trk}^{extrap} are the coordinates of the extrapolated position of the track at the closest point to the super-cluster.
- The ratio between the energy deposit in the HCAL tower just behind the ECAL cluster that constitutes the electron candidate must be less than 0.2
- The combined electron energy between ECAL super-cluster and track is calculated according the following rules:
  - $|E/p - 1| < 2\sigma_{E/p}$: the mean of E and p weighted by their uncertainties is used,
  - $E/p > 1 + 2\sigma_{E/p}$: E alone is used,
  - $E/p > 1 - 2\sigma_{E/p}$ and $E > 15$ GeV: E alone is used,
  - $E/p > 1 - 2\sigma_{E/p}$ or $E < 15$ GeV: p alone is used.

Calibration procedures are used to control the energy response of the ECAL crystal as well as corrections due to the energy lost by the electrons while travelling trough the tracker material. Some of the calibrations are similar to the ones we have already discussed for the jet reconstruction such has p_T balance, but for the ECAL the decay of the Z boson is often used due to its well-known properties. Other techniques are particular to the ECAL such as the monitoring of the crystal transparency, since they become more and more opaque (and less responsive) with radiation exposure.

4.6.2 Photon reconstruction

Unlike the electron, the photon does not leave any signature inside the tracker and therefore the only information we can use is obtained by the ECAL.

As for the electrons the reconstruction starts from a seed provided by the ECAL. In this case a new discriminating variable is used called R9. This variable is defined to be the ratio between the energy of a 3×3 group centered around the crystal with the highest energy deposit divided by the total energy of the super-cluster. Since approximately 94% of the incident energy of a single photon is contained in 3×3 group of crystals, the photon candidates with R9 > 0.94 in the barrel and R9 > 0.95 in the end-caps are selected. In order to improve the purity of the photons the super-cluster is required not to match any possible track consistent with originating the primary vertex.

The conversion of a photon into an electron-positron pair is one of the biggest sources of efficiency loss in the photon measurement. These lost photons are recovered by reconstructing e^+e^- pairs, so electrons close together in the ECAL, and then using the tracker information to find a possible photon conversion vertex with a dedicated algorithm.

4.7 Particle Flow

We have already mentioned the particle-flow (PF) algorithm [41] and in this section we will illustrate some of its characteristics. The PF algorithm aims to reconstruct and identify all the stable particles (i.e. muons, electrons, photons, neutral and charged hadrons) resulting from the pp collision, with a thorough combination of all CMS sub-detector information towards an optimal determination of their direction, energy, and type.

The CMS detector has many characteristics that make it well suited for PF reconstruction: an excellent tracking system; a large and high field solenoidal magnet; and a fine granularity electromagnetic calorimeter.

The principle behind the PF algorithm is rather simple but its implementation is far from easy. The building blocks of the reconstruction are similar to those we have already described above: charged particles tracks, calorimeter clusters, and muon tracks. These building blocks must have a good efficiency and low number of fakes. In order to achieve these goals, dedicated algorithms for tracking and clustering have been developed. The information needs to be properly combined in order to form the final physics objects and for this purpose a linking algorithm is used. In the following sections we will briefly discuss the main parts of the PF reconstruction.

4.7.1 Iterative tracking

The tracking is of fundamental importance to PF reconstruction: for instance in a jet almost 2/3 of the energy is carried by charged particles and these particles can be detected by the tracker and their momentum measured with an excellent resolution, much better than is achievable by the calorimeter, especially for low momentum tracks. The hight efficiency requirement along with a negligible number of fake tracks is achieved using an iterative tracking algorithm: strict selection criteria are applied in the first step where the first high quality tracks are reconstructed with relatively low efficiency and almost no fakes, then all the hits used are removed to reduce the combinatorics and the track selection criteria are relaxed to recover efficiency and keep fake tracks under control. The iterative tracking consists of six steps with looser and looser track selection, the last two are used to reconstruct displaced vertices of long-lived particles such as the K^0. Tracks with p_T of the order of 150 MeV and an origin more than 50 cm away from

the primary vertex can be reconstructed, while keeping the fake rate of the order of a few per cent.

4.7.2 Calorimeter clustering

The calorimeter information is used to detect and measure neutral particles (photons and neutral hadrons) as well as separate the neutral particle deposits from the charged hadrons, to reconstruct and identify electrons with all accompanying Bremsstrahlung photons, and to improve the energy measurement of charged hadrons in the case the track parameters were not precisely determined (in particular for low quality or high-p_T tracks).

The algorithm is used independently in the different subsystems: the ECAL and HCAL barrel and end-caps, as well as the ES first and second layer. The first step is to identify the cluster seeds. For this purpose calorimeter cells above a certain threshold are used. Next, PF "topological-clusters" are grown starting from the seed and merging all contiguous cells with energy that is two standard deviations above electronic noise (from 80 MeV to 300 MeV for ECAL and about 800 MeV for HCAL). Multiple PF clusters may be included in a topological-cluster.

The position of a PF cluster and its energy are determined with an iterative procedure. The initial estimate is the position of the seed, while energy is shared proportionally (using a Gaussian estimate $\exp\left[-d_{ij}^2/R^2\right]$) among all the PF clusters. The position and energy are then recomputed using a similar weighing technique to the one used for the ECAL, i.e. using as a weight the logarithm of the energy deposit of each cluster. The process ends when the PF cluster positions do not change more than a small fraction of the spatial resolution.

For the ECAL the position of the cluster is corrected for the fact that the

cells do not point to the origin of the reference system; the so-called depth correction is applied to account for this effect. The center of the electromagnetic shower (see Figure 4.2) inside the crystals is parametrized with the formula $depth = a(b + \ln E)$.

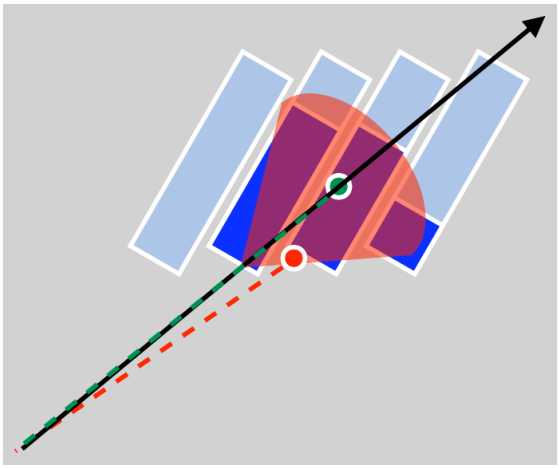


Figure 4.2: The impact on the ECAL surface is shown in red, the depth corrected position in green. The new position better matches the direction of the electron shower with respect to the primary vertex.

4.7.3 Link algorithm

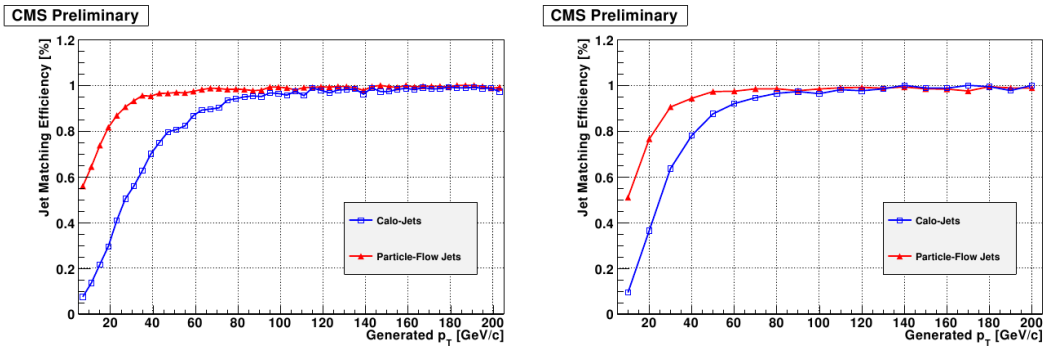
Once the reconstruction in all sub-detectors is done the physics objects must be built, and for this purpose a link algorithm is used. The charged tracks are extrapolated to the calorimeters: to the ES layers, to the ECAL reconstructed PF cluster position, and to the HCAL at a depth corresponding to one interaction length. If the extrapolated track position falls within the cluster boundaries then the two are linked together. The link algorithm also tries to collect emitted bremsstrahlung photons by extrapolating tangents to the electron track to the ECAL, using as a conversion vertex the intersection between the track and the tracker layers. A similar linking mechanism is used between calorimeter clusters, making sure that the cluster in the detector with higher granularity (PS or ECAL) is within the boundaries of the less granular detector (ECAL or HCAL).

The linking between tracker tracks and a muon track in the muon system is established using the standard χ^2, while this method has been abandoned for the clusters due to the poor resolution they may have.

The linking proceeds physics object by physics object, matching the relevant PF elements and removing them from the collection. The process ends when the PF elements collection is empty and all the final physics objects are constructed.

4.7.4 Particle reconstruction improvements

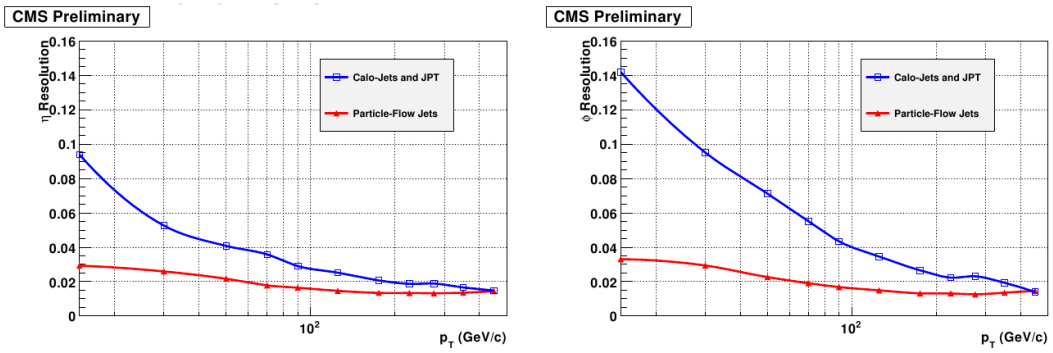
The particle-flow algorithm has superior performance (Fig. 4.3) compared to calorimeter only reconstruction of jets, in particular when dealing with low p_T jets that have a lot of soft particles that cannot be detected by the calorimeters but can be precisely measured in the tracker.



(a) $-1.5 < \eta < 1.5$ (b) $1.5 < |\eta| < 2.5$

Figure 4.3: Jet matching efficiency as obtained for calo-jets (open squares) and particle-flow jets (triangles) pointing to the barrel, $-1.5 < \eta < 1.5$ (a) and to the end-caps, $1.5 < |\eta| < 2.5$ (b), with a matching distance of 0.1.

The improved jet reconstruction and resolution (Fig. 4.4), as well as improved reconstruction of the other particles, also improves the $E_{\mathrm{T}}^{\mathrm{miss}}$ determination and this is very important for new physics searches.



(a) (b)

Figure 4.4: Jet η (a) and ϕ (b) resolutions as a function of p_T in the barrel.

Chapter 5

Event Selection

In this chapter we describe the methodology used for the signal event selection and also describe the data sample and Monte Carlo simulated events used in this analysis.

5.1 Data and MC samples

5.1.1 2011 collision data

This study is based on the complete 2011 dataset, comprising two running periods, 2011A and 2011B. The total integrated luminosity is 4.98 ± 0.11 fb^{-1} and the detailed list of the datasets used is listed Table 5.1 for the muon channel and in Table 5.2 for the electron channel.

We have used both "`SingleMu`" and "`DoubleMu`" datasets for background studies and only the second for our search as the needed trigger informations are stored in both. On the other hand all the electrons triggers used in this analysis (listed in Sec 5.2.1) are stored into the "`DoubleElectron`" datasets.

The internal division of the "A" and "B" run ranges is done to take into account

Dataset	Run Range
/DoubleMu/Run2011A-May10ReReco-v1/AOD	160329 – 163869
/DoubleMu/Run2011A-PromptReco-v4/AOD	165071 – 167784
/DoubleMu/Run2011A-05Aug2011-v1/AOD	170053 – 172619
/DoubleMu/Run2011A-PromptReco-v6/AOD	172620 – 175770
/DoubleMu/Run2011B-PromptReco-v1/AOD	175832 – 180252
/SingleMu/Run2011A-May10ReReco-v1/AOD	160329 – 163869
/SingleMu/Run2011A-PromptReco-v4/AOD	165071 – 167151
/SingleMu/Run2011A-05Aug2011-v1/AOD	170053 – 172619
/SingleMu/Run2011A-PromptReco-v6/AOD	172620 – 175770
/SingleMu/Run2011B-PromptReco-v1/AOD	175832 – 180252

Table 5.1: Summary of 2011 muon datasets used.

Dataset	Run Range
/DoubleElectron/Run2011A-05Jul2011ReReco-ECAL-v1/AOD	160329 – 168437
/DoubleElectron/Run2011A-05Aug2011-v1/AOD	170722 – 172619
/DoubleElectron/Run2011A-PromptReco-v6/AOD	172620 – 175770
/DoubleElectron/Run2011B-PromptReco-v1/AOD	175832 – 180252

Table 5.2: Summary of 2011 electron datasets used.

differences in the detector operating conditions. Datasets labelled "PromptReco" in Tables 5.1 and 5.2 indicates data that have been reconstructed after they where recorded by the CMS detector. Datasets labelled "May10ReReco" and "05Aug2011" are datasets that have been reconstructed twice, the second time to correct errors, recover integrated luminosity, or simply to make improvements.

5.1.2 Standard Model monte carlo samples

The Monte Carlo samples used for background estimation and validation of the data-driven background estimation methods are listed in Table 5.3[1] and in Tables 5.4 and 5.5[2], along the full dataset name, the corresponding integrated luminosity, and the best available cross-section. The cross-section is, in general, given by the NLO

[1]MC used for both or only the muon channel analysis with inclusive decay modes.
[2]MC samples with exclusive muon or electron decay modes.

theoretical prediction except samples binned by p_T (as the Z+jet) for which the cross-section given by generator is used. The two only exceptions are the the $t\bar{t}$ sample for which the CMS measured cross-section is used since has a smaller uncertainty, and the W+jets cross-section that is calculated at the NNLO.

Samples used to estimate the number of events from standard model processes include diboson production, $t\bar{t}W$ production, and same sign W pair production via double W-strahlung and double-parton scattering ($2 \times q\bar{q}' \rightarrow W$). The $t\bar{t}$, single-top, W+jets, and QCD samples are used to validate the fake background estimation method, while the Z+jet and W+jets samples are used to correct the fake rate measurement. The conditions of the LHC collisions changed dramatically during the year and the MC simulations tried to keep pace with the increasing luminosity and number of pileup events. Despite the effort to produce up-to-date Monte Carlo samples, not all of them had the correct pileup event distributions. Therefore the Monte Carlo are reweighed by the the number of interactions per crossing in data, so that the distributions in data and MC match exactly.

5.1.3 Majorana neutrino signal monte carlo

Heavy Majorana neutrino production and decay are simulated using the event generator described in Ref. [48] and implemented in ALPGEN [77]. The production process simulated is the resonant production of a Majorana neutrino as discussed in Sec. 2.3.4. The output of the ALPGEN generator was stored in the Les Houches 1.0 format [92]. These datasets were then unweighed using PYTHIA to generate events for each neutrino mass. The generated event files were interfaced with CMSSW where

Dataset	Dataset Name	Cross Section (pb)	$\int \mathcal{L}dt$ (pb^{-1})
W^+W^-	/WW_TuneZ2_7TeV_pythia6_tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM	43	9.83×10^4
WZ	/WZ_TuneZ2_7TeV_pythia6_tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM	18.2	2.34×10^5
ZZ	/ZZ_TuneZ2_7TeV_pythia6_tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM	5.9	7.10×10^5
W^+W^+	/PhysicsProcesses_WplusWplus_42x_v1	0.165	8.11×10^5
W^-W^-	/PhysicsProcesses_WminusWminus_42x_v1	0.0551	3.18×10^6
dp $W^\pm W^\pm$	/DoublePartonWWFastSim_CMSSW425PUv1	0.38	1.06×10^6
$t\bar{t}$	/TTJets_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM	154	2.40×10^4
$t\bar{t}$	/TTJets_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v2/AODSIM	154	3.87×10^5
t (tW^-)	/T_TuneZ2_tW-channel-DR_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM	5.3	1.54×10^5
$\bar{t}$ ($\bar{t}W^+$)	/Tbar_TuneZ2_tW-channel-DR_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v2/AODSIM	5.3	1.53×10^5
t (s-channel)	/T_TuneZ2_s-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM	2.72	9.56×10^4
$\bar{t}$ (s-channel)	/Tbar_TuneZ2_s-channel_7TeV-powheg-tauola/ /Fall11-PU_S6_START42_V14B-v1/AODSIM	1.49	9.26×10^4
t (t-channel)	/T_TuneZ2_t-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM	42.6	9.16×10^4
$\bar{t}$ (t-channel)	/Tbar_TuneZ2_t-channel_7TeV-powheg-tauola/ /Fall11-PU_S6_START42_V14B-v1/AODSIM	22	8.84×10^4
$t\bar{t}W$	/TTbarInclWIncl_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM	0.153	2.29×10^6
W+jets	/WJetsToLNu_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM	3.13×10^4	2.6×10^3

Table 5.3: Summary of inclusive standard model monte carlo datasets used: dp $W^\pm W^\pm$ indicates the double-parton scattering ($2 \times q\bar{q}' \to W$). The samples W^+W^+, W^-W^-, and dp $W^\pm W^\pm$ were privately produced and validated.

Dataset	Dataset Name	Cross Section (pb)	$\int \mathcal{L} dt$ (pb^{-1})
$W\gamma$	/WGstarToLNu2Mu_TuneZ2_7TeV-madgraph-tauola//Summer11-PU_S4_START42_V11-v1/AODSIM	1.60	6.23×10^4
Z + jets	/ZJetToMuMu_Pt-0to15_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	4281	48.9
Z + jets	/ZJetToMuMu_Pt-15to20_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	145.1	1.50×10^3
Z + jets	/ZJetToMuMu_Pt-20to30_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	130.5	1.26×10^3
Z + jets	/ZJetToMuMu_Pt-30to50_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	84.0	1.85×10^3
Z + jets	/ZJetToMuMu_Pt-50to80_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	32.2	3.31×10^3
Z + jets	/ZJetToMuMu_Pt-80to120_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	9.98	1.10×10^4
QCD	/QCD_Pt-20_MuEnrichedPt-15_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1/AODSIM	2.97×10^8	296
QCD	/QCD_Pt-15to20_MuPt5Enriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1/AODSIM	5.79×10^8	1.51
QCD	/QCD_Pt-20to30_MuPt5Enriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1/AODSIM	2.36×10^8	7.51
QCD	/QCD_Pt-30to50_MuPt5Enriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1/AODSIM	5.31×10^7	18.3
QCD	/QCD_Pt-50to80_MuPt5Enriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1/AODSIM	6.35×10^6	77.3
QCD	/QCD_Pt-80to120_MuPt5Enriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1/AODSIM	7.85×10^5	291
QCD	/QCD_Pt-120to150_MuPt5Enriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1/AODSIM	9.30×10^4	1772
QCD	/QCD_Pt-150_MuPt5Enriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1/AODSIM	4.76×10^4	1409

Table 5.4: Summary of standard model monte carlo datasets used for the muon channel.

Dataset	Dataset Name	Cross Section (pb)	$\int \mathcal{L} dt$ (pb^{-1})
$W\gamma$	/WGstarToLNu2E_TuneZ2_7TeV-madgraph-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM	5.546	4.56×10^4
Z + jets	/ZJetToEE_Pt-0to15_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	4280	46.7
Z + jets	/ZJetToEE_Pt-15to20_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	145	1379
Z + jets	/ZJetToEE_Pt-20to30_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	131	1145
Z + jets	/ZJetToEE_Pt-30to50_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	84.0	1786
Z + jets	/ZJetToEE_Pt-50to80_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	32.3	3096
Z + jets	/ZJetToEE_Pt-80to120_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	9.99	1.00×10^4
Z + jets	/ZJetToEE_Pt-120to170_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	2.74	3.65×10^4
Z + jets	/ZJetToEE_Pt-170to230_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	0.722	1.39×10^5
Z + jets	/ZJetToEE_Pt-230to300_TuneZ2_7TeV_pythia6/Summer11-PU_S3_START42_V11-v2/AODSIM	0.194	5.15×10^5
QCD	/QCD_Pt-15to30_TuneZ2_7TeV_pythia6/Fall11-PU_S6_START42_V14B-v3/AODSIM	8.16×10^8	1.30×10^{-2}
QCD	/QCD_Pt-30to50_TuneZ2_7TeV_pythia6/Fall11-PU_S6_START42_V14B-v3/AODSIM	5.31×10^7	1.24×10^{-1}
QCD	/QCD_Pt-50to80_TuneZ2_7TeV_pythia6/Fall11-PU_S6_START42_V14B-v2/AODSIM	6.36×10^6	9.36×10^{-1}
QCD	/QCD_Pt-80to120_TuneZ2_7TeV_pythia6/Fall11-PU_S6_START42_V14B-v3/AODSIM	7.84×10^5	8.40
QCD	/QCD_Pt-120to170_TuneZ2_7TeV_pythia6/Fall11-PU_S6_START42_V14B-v1/AODSIM	1.15×10^5	51.6
QCD	/QCD_Pt-170to300_TuneZ2_7TeV_pythia6/Fall11-PU_S6_START42_V14B-v1/AODSIM	2.43×10^4	256
QCD	/QCD_Pt-20to30_EMEnriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1/AODSIM	2.36×10^8	14.0
QCD	/QCD_Pt-30to80_EMEnriched_TuneZ2_7TeV-pythia/Fall11-PU_S6_START42_V14B-v1/AODSIM	5.94×10^7	19.3
QCD	/QCD_Pt-80to170_EMEnriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v2/AODSIM	8.09×10^5	56.0
QCD	/QCD_Pt-170to250_EMEnriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v2/AODSIM	2.21×10^4	919
QCD	/QCD_Pt-250to350_EMEnriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v2/AODSIM	2.90×10^3	8152
QCD	/QCD_Pt-350_EMEnriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v2/AODSIM	520	54530

Table 5.5: Summary of standard model monte carlo datasets with only the electron decay mode.

parton showering, vertex smearing, GEANT detector simulation, digitization of simulated electronics signal, and reconstruction were performed. The Monte Carlo samples are reweighed to ensure the proper simulation of the number of interactions per crossing. The list of official production Monte Carlo samples with the corresponding cross-section as a function of Majorana neutrino mass, using a Majorana neutrino lepton coupling value $S_{N\ell} = |V_{N\ell}|^2 = 1$, are shown in Table 5.6 for the $\mu\mu$ decay and in Table 5.7 for the ee decay.

5.2 Signal Selection

As mentioned in Chapter 1, and shown again in Figure 5.1, the signal is characterized by two leptons and an accompanying W boson. Since we consider the heavy neutrino to be a Majorana particle, both opposite-sign and same-sign lepton pairs are produced. Because the SM background is much lower in the same-sign channel we have opted to concentrate on this particular decay mode and our selection criteria are optimized to detect same-sign lepton pairs. Along with the two leptons a W boson is also produced, and we look for signatures in which the W boson decays hadronically to two jets.

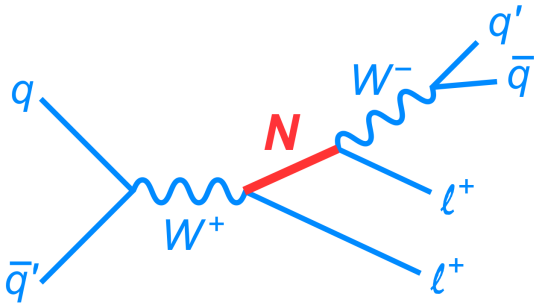


Figure 5.1: Faynman diagram of the Majorana neutrino (in red) production and decay.

Mass (GeV)	Dataset Name	Cross Section (pb)	$\int \mathcal{L} dt$ (pb^{-1})
50	/MajoranaNeutrinoToEE_M-50_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	866.4 ± 1.1	57.70
60	/MajoranaNeutrinoToEE_M-60_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	492.5 ± 0.7	101.52
70	/MajoranaNeutrinoToEE_M-70_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	171.3 ± 0.3	291.81
75	/MajoranaNeutrinoToEE_M-75_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	64.70 ± 0.10	768.11
80	/MajoranaNeutrinoToEE_M-80_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	15.21 ± 0.02	3286.9
85	/MajoranaNeutrinoToEE_M-85_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	$(8242 \pm 6) \times 10^{-3}$	6066.0
90	/MajoranaNeutrinoToEE_M-90_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	$(5601 \pm 2) \times 10^{-3}$	8604.5
95	/MajoranaNeutrinoToEE_M-95_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	$(3915 \pm 1) \times 10^{-3}$	12770
100	/MajoranaNeutrinoToEE_M-100_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	$(2787 \pm 1) \times 10^{-3}$	17938
105	/MajoranaNeutrinoToEE_M-105_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	$(2067.7 \pm 0.6) \times 10^{-3}$	24179
110	/MajoranaNeutrinoToEE_M-110_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	$(1590.2 \pm 0.5) \times 10^{-3}$	31250
120	/MajoranaNeutrinoToEE_M-120_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	$(1017.6 \pm 0.3) \times 10^{-3}$	49126
130	/MajoranaNeutrinoToEE_M-130_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	$(690.6 \pm 0.2) \times 10^{-3}$	69791
150	/MajoranaNeutrinoToEE_M-150_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	$(354.8 \pm 0.1) \times 10^{-3}$	140900
170	/MajoranaNeutrinoToEE_M-170_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	$(203.93 \pm 0.06) \times 10^{-3}$	236336
190	/MajoranaNeutrinoToEE_M-190_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	$(126.57 \pm 0.04) \times 10^{-3}$	395030
210	/MajoranaNeutrinoToEE_M-210_7TeV_Alpgen/Fall11-PU_S6_START42_V14B/AODSIM	$(83.011 \pm 0.024) \times 10^{-3}$	569804

Table 5.6: Summary of Majorana neutrino signal Monte Carlo samples. The samples are generated with $|V_{\mu N}|^2 = 1$. The calculated cross-section as a function of Majorana neutrino mass is also shown. Note that during the CMS official MC production the filenames were mislabeled as EE instead of MuMu and since this is not possible to correct in DBS the EE names here refer to MuMu samples.

Mass (GeV)	Dataset Name	Cross Section (pb)	$\int \mathcal{L} dt$ (pb^{-1})
50	/MajoranaNeutrinoToMuMu_M-50_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	866.4 ± 1.1	57.705
60	/MajoranaNeutrinoToMuMu_M-60_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	492.3 ± 0.6	101.56
70	/MajoranaNeutrinoToMuMu_M-70_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	171.7 ± 0.2	271.86
75	/MajoranaNeutrinoToMuMu_M-75_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	64.72 ± 0.08	767.92
80	/MajoranaNeutrinoToMuMu_M-80_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	15.21 ± 0.02	3287.3
85	/MajoranaNeutrinoToMuMu_M-85_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	$(8243 \pm 5) \times 10^{-3}$	6065.4
90	/MajoranaNeutrinoToMuMu_M-90_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	$(5608 \pm 6) \times 10^{-3}$	8915.9
95	/MajoranaNeutrinoToMuMu_M-95_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	$(3919 \pm 3) \times 10^{-3}$	12679
100	/MajoranaNeutrinoToMuMu_M-100_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	$(2795 \pm 3) \times 10^{-3}$	17886
105	/MajoranaNeutrinoToMuMu_M-105_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	$(2073 \pm 2) \times 10^{-3}$	23245
110	/MajoranaNeutrinoToMuMu_M-110_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	$(1594 \pm 2) \times 10^{-3}$	31354
120	/MajoranaNeutrinoToMuMu_M-120_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	$(1018.5 \pm 0.7) \times 10^{-3}$	49086
130	/MajoranaNeutrinoToMuMu_M-130_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	$(690.8 \pm 0.4) \times 10^{-3}$	72376
150	/MajoranaNeutrinoToMuMu_M-150_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	$(355.0 \pm 0.2) \times 10^{-3}$	140812
170	/MajoranaNeutrinoToMuMu_M-170_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	$(204.0 \pm 0.1 \times 10^{-3}$	255059
190	/MajoranaNeutrinoToMuMu_M-190_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	$(126.64 \pm 0.07) \times 10^{-3}$	392403
210	/MajoranaNeutrinoToMuMu_M-210_7TeV_Alpgen/Fall11-P_S6_START42_V14B-v1/AODSIM	$(83.05 \pm 0.05) \times 10^{-3}$	601999

Table 5.7: Summary of Majorana neutrino signal Monte Carlo samples. The samples are generated with $|V_{eN}|^2 = 1$. The calculated cross-section as a function of Majorana neutrino mass is also shown. Note that during the CMS official MC production the filenames were mislabeled as MuMu instead of EE and since this is not possible to correct in DBS the MuMu names here refer to EE samples.

Therefore the signal we are looking for consists of two same-sign leptons with two jets and no significant missing energy. The p_T of the decay particles depends on the mass of the neutrino that is produced, and for Majorana neutrino mass (m_N) below the W boson mass, the particles are relatively soft. Hence our selection cannot cut hard on p_T without loosing sensitivity at low Majorana neutrino mass.

5.2.1 Trigger

Given the event topology described above, the obvious trigger choice is to use the lowest available unprescaled double-muon and double-electron triggers. During the 2011 data taking period the LHC instantaneous luminosity increased rapidly and the various trigger paths had to be complaint with it.

For the muon channel we used three different triggers through the year:

- `HLT_DoubleMu7` for runs 160404 to 163869,
- `HLT_Mu13_Mu8` for runs 165071 to 178380,
- `HLT_Mu17_Mu8` for runs 178424 to 180252.

The corresponding integrated luminosities for these triggers are 0.22 fb^{-1}, 3.86 fb^{-1}, and 0.89 fb^{-1}, respectively.

For the electron channel only two triggers were needed:

- `HLT_Ele17_CaloIdL_CaloIsoVL_Ele8_CaloIdL_CaloIsoVL`
  for runs 160329 to 168437 counting the first 1.2 fb^{-1},
- `HLT_Ele17_ID_Ele8_ID` for all remaining runs,
  where `ID` stands for `CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL`.

Besides the triggers for the signal selection, other triggers were used for background studies, in particular for the determination of the fake rate in QCD enhanced collision data. For this purpose two possible approaches are possible: one is to use lepton triggers and then suppress electroweak processes, the other is to use hadronic triggers and then select events with leptons. We have decided to use the first approach and therefore use, for the muons, the following triggers:

- `HLT_Mu8`,
- `HLT_Mu8_Jet40`,
- `HLT_Mu15`.

These triggers were chosen with the purpose to match as close as possible the triggers used in the signal selection, and at the same time to provide reasonable statistics for the computation of the fake rate. The `HLT_Mu8_Jet40`, as we will shortly see, is the perfect match for our purposes since it has a good, low, muon p_T cut and includes a jet that is of fundamental importance for the fake rate measurement.

For the electron case the trigger choice is driven also by the different ID used in different periods of the year, therefore we have used the following triggers:

- `HLT_Ele8_CaloIdL_CaloIsoVL` for runs 160329 to 168437,
- `HLT_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL` for all remaining runs.

We then apply the fake rate calculated with a certain ID only to the corresponding signal trigger. The meaning of the electron IDs is summarized in Table 5.8.

Name	Criterion
`CaloIdL`	$H/E < 0.15(0.10)$ $\sigma_{i\eta i\eta} < 0.014(0.035)$
`CaloIdT`	$H/E < 0.05(0.05)$ $\sigma_{i\eta i\eta} < 0.011(0.031)$
`TrkIdVL`	$\lvert\Delta\eta\rvert < 0.01(0.01)$ $\Delta\phi < 0.15(0.10)$
`TrkIdT`	$\lvert\Delta\eta\rvert < 0.008(0.008)$ $\Delta\phi < 0.07(0.05)$
`CaloIsoVL`	ECalIso$/E_T < 0.2(0.2)$ HCalIso$/E_T < 0.2(0.2)$
`CaloIsoT`	ECalIso$/E_T < 0.15(0.075)$ HCalIso$/E_T < 0.15(0.075)$
`CaloIsoVT`	ECalIso$/E_T < 0.05(0.05)$ HCalIso$/E_T < 0.05(0.05)$
`TrkIsoVL`	TrkIso$/E_T < 0.2(0.2)$
`TrkIsoT`	TrkIso$/E_T < 0.15(0.075)$
`TrkIsoVT`	TrkIso$/E_T < 0.05(0.05)$

Table 5.8: Summary of requirements applied to electrons in the triggers used for this analysis. The selection requirements are given for electrons in the barrel (endcap). L=Loose, VL=Very loose, T=Tight, VT=Very Tight. H/E is the relative hadronic activity (energy deposited in the HCAL divided total energy), $\sigma_{i\eta i\eta}$ is the lateral shower shape, $\Delta\eta$ and $\Delta\phi$ are respectively the track-cluster matching parameters in the η and ϕ directions. The latter are the relative isolation deposits of the Tracker, ECAL and HCAL (for more details see Sec. 5.2.3).

Trigger efficiency

To determine the efficiency of the dilepton triggers we have used the standard trag-and-probe procedure described in Ref. [45] and used in the CMS $H \to WW \to 2\ell+2\nu$ search. Our search uses the same signal triggers used in the $H \to WW \to 2\ell+2\nu$ search and has a lot of similarities in the muon and electron selection. Using a data driven tag-and-probe method they measure the efficiencies for muons above 10 GeV in $Z \to \ell^+\ell^-$ ($\ell = e, \mu$) events. They measure the double muon efficiency to be in the range of $(92-97)\%$ depending on muon p_T and between $(99-100)\%$ for the double electron efficiency. This agrees with the efficiency we see in the simulated signal datasets, and we apply a trigger efficiency of $(96 \pm 2)\%$.

5.2.2 Primary vertex selection

Events are required to have at least one good primary vertex to ensure they are good collision events. The primary vertex must have more than 4 degrees of freedom in the fit. The z position of the vertex must be within 24 cm of the nominal detector center, and the transverse position must be within 2 cm of the beam spot. The vertex with the largest summed squared-p_T of the associated tracks that fulfils the above requirements is chosen as the primary vertex.

5.2.3 Muon selection

To ensure the good quality of muon candidates several requirements are put in place, most of them are standardized in CMS by the Particle Object Group (POG), while others are specific to this analysis. As already mentioned in Section 4.3, muons can be reconstructed either as standalone, tracker, or global muons, and good muons are often reconstructed in all three categories. We require them to be both Global and Tracker muons to be sure their quality is good. The other quality criteria are listed below:

- $\chi^2/ndof < 10$ of the fitted global track,
- transverse impact parameter d_0 relative to the primary vertex < 100 μm,
- longitudinal d_z within 0.1 cm of the primary vertex,
- number of hits in the tracker $\geqslant 11$,
- number of hits in the muon system $\geqslant 1$,
- $p_T > 10$ GeV
- $|\eta| < 2.4$

- veto for muons with ECAL deposits $\geqslant 4$ GeV
- veto for muons with HCAL deposits $\geqslant 6$ GeV
- relative isolation (RelIso) < 0.1

It is worth to mention few words about the relative isolation variable. This variable is defined by the equation:

$$RelIso = \frac{Iso^{tracker} + Iso^{ECAL} + Iso^{HCAL}}{p_T} \tag{5.1}$$

where p_T is the measured muon momentum, and the three Iso components indicate the energy deposited in a cone of size $\Delta R < 0.3$ as illustrated in Fig. 5.2. The tracker isolation is the p_T sum of all the tracks inside the cone.

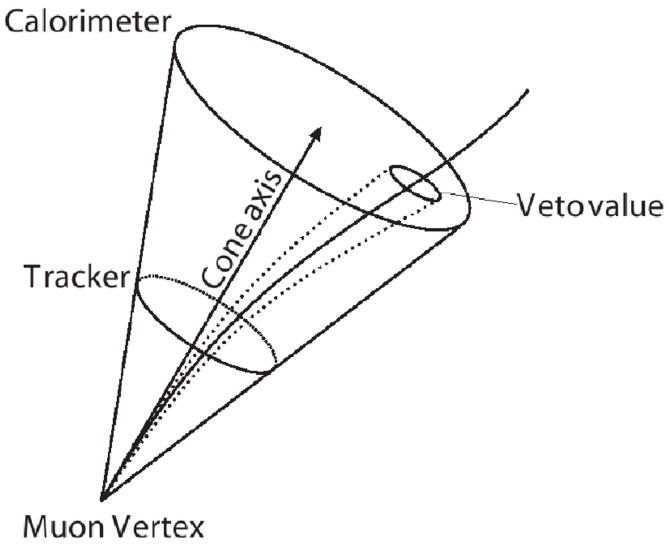


Figure 5.2: Chematic of the muon isolation cone.

The presence of pileup will increase the average energy in the detector and bias the RelIso value. To correct of this effect we use the so called "fastjet correction" or "ρ correction". The idea behind this correction is to subtract the energy contamination

due to pileup from the various energy deposits. In order to apply the correction Eq. 5.1 is modified to the following:

$$RelIso = \frac{Iso^{tracker}_{corrected} + Iso^{ECAL}_{corrected} + Iso^{HCAL}_{corrected}}{p_T} \tag{5.2}$$

where:

$$Iso^{tracker}_{corrected} = \Big(Iso^{tracker} - \rho A_{tracker}(\eta)\Big) \tag{5.3}$$

$$Iso^{ECAL}_{corrected} = \big(Iso^{ECAL} - \rho A_{ECAL}(\eta)\big) \tag{5.4}$$

$$Iso^{HCAL}_{corrected} = \big(Iso^{HCAL} - \rho A_{HCAL}(\eta)\big) \tag{5.5}$$

The quantity ρ is the energy density in the event as measured by the fastjet algorithm. This energy is multiplied by an effective area term that varies for each sub-detector. The values of the effective areas for the various sub-detectors are listed in Table 5.9. These values are empirically computed studying the distributions of the three detector

Region	Tracker	ECAL	HCAL
$\lvert\eta\rvert < 0.5$	0	0.091	0.029
$0.5 < \lvert\eta\rvert < 1.0$	0	0.077	0.027
$1.0 < \lvert\eta\rvert < 1.5$	0	0.055	0.036
$1.5 < \lvert\eta\rvert < 2.0$	0	0.034	0.040
$2.0 < \lvert\eta\rvert < 2.4$	0	0.046	0.054

Table 5.9: List of the effective areas used in the muon isolation correction for different sub-detectors as a function of η.

isolation variables versus the number of reconstructed vertices (n_{vtx}). Ideally they should not change, even if the number of vertices increases, but this is not the case; in fact, the amount of energy in the calorimeter increases with pileup and the energy inside the muon cone increase with pileup, and the effective area is derived from the slope

of these distributions. Once the subtraction is performed the isolation variables are approximately independent of n_{vtx} and, therefore, robust versus pileup.

Muons that satisfy the requirements above are defined as "prompt" muons, most of them are real muons that come from the decay of a Z or W boson, as well as muons from heavy Majorana signal events. On the other hand, some muons may fall in this category even if they are not from the mentioned electroweak processes; these muons are in general called "fake muons". For instance, fake muons include muons from b quarks (that are real muons but are not well isolated), or jets or hadrons reconstructed as muons. In Section 6.2 we will discuss in detail how this contamination of non-prompt muons in the prompt muon sample is measured.

5.2.4 Electron selection

Offline requirements for the the selection of electron candidates uses information from the tracker and calorimeters. We select GSF electrons requiring $p_T > 10$ GeV and $|\eta| < 2.4$ as for the muons. Electron candidates with $1.442 < |\eta| < 1.556$ in the gap between the ECAL barrel and end-cap are rejected. The transverse and longitudinal impact parameters with respect to the primary vertex must satisfy $d_0 < 0.1$ mm, and $d_z < 1$ mm, as for the muons.

Electrons are required to be ECAL seeded with no missing hits. Soft electrons with $p_T < 20$ GeV are required to have fbrem > 0.15, or super cluster with $|\eta| < 1$ and $E/p > 0.95$.

Electron candidates must pass the conversion rejection and simple cut-based identification developed by he Vector Boson Task Force at the 80% efficiency working point (VBTF80), the selection criteria of this identification are summarized in Table 5.10. Isolation criteria are excluded because we need to relax them for back-

ground studies. In addition, we require electrons to be isolated from muon candidates by $\Delta R > 0.4$.

	barrel	end-cap
Conversion rejection		
missing hits	0	0
Dist	0.2	0.2
$\Delta \cot \theta$	0.2	0.2
Electron ID		
$\sigma_{i\eta i\eta}$	0.01	0.03
$\Delta\phi$	0.06	0.03
$\Delta\eta$	0.004	0.007
H/E	0.04	0.025

Table 5.10: Cut values for the VBTF80 electron identification. Isolation requirements are excluded since we need looser selection for background studies. The conversion rejection has requirements on the number of allowed Tracker missing hits, and the distance between tangent points the $r-\phi$ plane (Dist) and the $\Delta \cot \theta$ in the $r-z$ plane.

Further requirements are made on electron charge consistency to reduce the probability of mismeasurement of the electron charge. For GSF electrons, the charge is calculated by three different methods: GSF track fitting, CTF track fitting, and pixel-supercluster comparison. We require that the three methods give the same result for the charge sign of the electron candidate. This requirement reduces the background from $Z \rightarrow e^+e^-$ events where one of the charges is mismeasured.

The relative isolation for the electron is computed, as for the muon, using Equation 5.2. The technique used for the pileup correction is the same but with different effective areas, appropriate to electrons and listed in Table 5.11.

Region	Tracker	ECAL	HCAL
barrel ($\|\eta\| < 1.442$)	0	0.078	0.026
end-cap ($1.556 < \|\eta\| < 2.4$)	0	0.040	0.072

Table 5.11: List of the effective areas used in the muon isolation correction for different sub-detectors as a function of η.

5.2.5 Jet selection

All events are required to have two jets with $p_T > 20$ GeV and $|\eta| < 2.5$, which are well separated from lepton candidates ($\Delta R_{\ell-jet} > 0.4$). Jets are reconstructed from calorimeter and tracker information using the particle flow algorithm, with all the corrections applied, and using the anti-k_T clustering algorithm with $R = 0.5$ as in the standard CMS prescription.

5.2.6 Final selection requirements

Events must satisfy the requirement of having two same-sign leptons, one of which must have $p_T > 20$ GeV, and two jets. Moreover, to further suppress residual background we require:

- di-lepton invariant mass $m_{ll} > 5$ GeV,
- for the electron channel, events with a third electron are vetoed to reduce the photon conversion background;
- for the muon channel, events with a third opposite sign muon that combines with one of the candidate muons to from an invariant mass within $76 - 106$ GeV are excluded.

Finally, the events are required to have missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ less than 50 GeV. This requirement suppresses backgrounds such a $t\bar{t}$ and W+jets, which have significant real $E_{\mathrm{T}}^{\mathrm{miss}}$ in the final state. Events with $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV are used as a control sample to cross check the background prediction.

We define two sets of selection cuts, one with all the requirements listed above ("loose selection") and one with the additional requirement that the jet transverse momentum of the two jets should be $p_T > 30$ GeV ("tight selection").

5.3 Selection Efficiency

5.3.1 Offline selection efficiency

The efficiency to select two isolated muons with $p_T > 10, 20$ GeV as described above, is given in Tables 5.12 and 5.13 for events passing the loose and tight selection, respectively.

In the region between the heavy Majorana neutrino masses of 70 to 100 GeV the efficiency of finding two muons passing our selection cuts drops considerably. This is due to the nature of the production process, where the heavy neutrino is either producing or being produced from an on-shell W boson. When the neutrino mass is very near to the W boson mass the muon momentum drops below our thresholds and many events do not pass our selection.

The overall efficiency to select signal events passing the muon and additional selection cuts is shown in Tables 5.14 and 5.15.

Generated	Total	% Events Accepted			
Mass Points	Events	Track Quality	η and p_T	Isolation	All
50 GeV	47954	50.2 ± 0.32	25.6 ± 0.23	39.6 ± 0.29	17.3 ± 0.19
70 GeV	47896	49.8 ± 0.30	8.68 ± 0.13	34.1 ± 0.27	5.22 ± 0.10
75 GeV	47903	44.3 ± 0.30	7.59 ± 0.13	30.2 ± 0.25	4.78 ± 0.10
80 GeV	46950	44.7 ± 0.31	12.9 ± 0.20	30.3 ± 0.26	6.12 ± 0.15
85 GeV	12988	47.8 ± 0.45	14.1 ± 0.24	31.3 ± 0.36	7.93 ± 0.18
90 GeV	23995	56.6 ± 0.49	20.9 ± 0.30	37.0 ± 0.39	10.7 ± 0.21
95 GeV	12998	60.0 ± 0.68	32.6 ± 0.50	40.9 ± 0.56	18.3 ± 0.37
100 GeV	17040	61.7 ± 0.60	41.7 ± 0.49	45.2 ± 0.51	25.3 ± 0.39
105 GeV	14090	63.9 ± 0.67	49.7 ± 0.59	49.1 ± 0.67	31.7 ± 0.47
110 GeV	16855	65.4 ± 0.62	56.4 ± 0.58	51.7 ± 0.55	36.8 ± 0.46
130 GeV	15043	69.3 ± 0.68	66.3 ± 0.66	58.6 ± 0.62	47.0 ± 0.56
150 GeV	8999	72.4 ± 0.90	72.0 ± 0.89	63.4 ± 0.84	52.1 ± 0.76
170 GeV	13494	75.0 ± 0.75	75.5 ± 0.75	65.7 ± 0.70	54.6 ± 0.64
190 GeV	13494	76.9 ± 0.76	78.4 ± 0.76	67.6 ± 0.71	56.6 ± 0.65
210 GeV	11566	78.6 ± 0.82	80.6 ± 0.83	68.9 ± 0.77	58.1 ± 0.71

Table 5.12: Muon selection and individual cut efficiencies for events passing the loose selection. Errors are statistical only.

Generated	Total	% Events Accepted			
Mass Points	Events	Track Quality	η and p_T	Isolation	All
50 GeV	49997	50.2 ± 0.32	25.8 ± 0.23	44.5 ± 0.30	16.8 ± 0.19
70 GeV	49999	50.1 ± 0.32	8.90 ± 0.13	39.5 ± 0.28	5.15 ± 0.10
75 GeV	49700	44.4 ± 0.30	7.84 ± 0.13	35.8 ± 0.27	4.71 ± 0.098
80 GeV	49994	45.2 ± 0.30	13.1 ± 0.19	36.1 ± 0.28	6.02 ± 0.15
85 GeV	49997	48.1 ± 0.31	14.1 ± 0.17	37.5 ± 0.28	7.46 ± 0.12
90 GeV	48192	57.1 ± 0.35	20.8 ± 0.21	42.8 ± 0.30	10.3 ± 0.15
95 GeV	49993	60.6 ± 0.35	33.0 ± 0.26	46.8 ± 0.31	18.2 ± 0.19
100 GeV	49994	58.3 ± 0.34	42.1 ± 0.29	38.6 ± 0.28	25.1 ± 0.23
105 GeV	49995	63.9 ± 0.36	49.9 ± 0.32	53.4 ± 0.33	30.8 ± 0.25
110 GeV	49694	65.4 ± 0.36	56.1 ± 0.34	56.1 ± 0.34	36.2 ± 0.27
130 GeV	48197	69.5 ± 0.38	66.8 ± 0.28	63.2 ± 0.36	46.7 ± 0.31
150 GeV	49997	72.5 ± 0.38	71.8 ± 0.38	67.1 ± 0.37	51.4 ± 0.32
170 GeV	48196	75.4 ± 0.40	75.7 ± 0.40	69.7 ± 0.38	54.7 ± 0.34
190 GeV	49999	77.3 ± 0.39	78.4 ± 0.40	71.4 ± 0.38	56.7 ± 0.34
210 GeV	47300	79.2 ± 0.41	81.0 ± 0.41	73.3 ± 0.39	58.4 ± 0.35

Table 5.13: Muon selection and individual cut efficiencies for events passing the tight selection. Errors are statistical only.

Generated	Total	% Events Accepted			Events
Mass Points	Events	All Muon	Jets	All Cuts	for 4.98 fb^{-1}
50 GeV	47954	17.3 ± 0.19	22.0 ± 0.21	2.18 ± 0.07	93582 ± 2660
70 GeV	47896	5.22 ± 0.10	31.5 ± 0.26	1.33 ± 0.06	11296 ± 428
75 GeV	47903	4.78 ± 0.10	35.6 ± 0.27	1.49 ± 0.06	4788 ± 171
80 GeV	46950	6.12 ± 0.15	44.7 ± 0.31	2.89 ± 0.10	2181 ± 48.4
85 GeV	12988	7.93 ± 0.18	55.2 ± 0.48	3.14 ± 0.12	1282 ± 43.4
90 GeV	23995	10.7 ± 0.21	58.6 ± 0.49	5.30 ± 0.16	1474 ± 39.2
95 GeV	12998	18.3 ± 0.37	61.6 ± 0.69	10.3 ± 0.29	1995 ± 52.1
100 GeV	17040	25.3 ± 0.39	63.1 ± 0.61	14.4 ± 0.30	1990 ± 38.4
105 GeV	14090	31.7 ± 0.47	63.5 ± 0.67	17.9 ± 0.37	1839 ± 34.8
110 GeV	16855	36.8 ± 0.46	64.2 ± 0.62	20.9 ± 0.37	1647 ± 26.5
130 GeV	15043	47.0 ± 0.56	67.7 ± 0.67	27.7 ± 0.45	949 ± 13.9
150 GeV	8999	52.1 ± 0.76	70.2 ± 0.88	31.2 ± 0.62	548 ± 9.9
170 GeV	13494	54.6 ± 0.64	73.0 ± 0.74	33.9 ± 0.53	343 ± 4.8
190 GeV	13494	56.6 ± 0.65	74.7 ± 0.74	35.8 ± 0.54	225 ± 3.1
210 GeV	11566	58.1 ± 0.71	76.8 ± 0.81	37.5 ± 0.60	154 ± 2.2

Table 5.14: Muon channel event selection efficiencies for the loose selection. Errors are statistical only.

For the electron channel, the event selection efficiencies show the same be-

Generated Mass Points	Total Events	% Events Accepted All Muon	Jets	All Cuts	Events for $4.98\,\text{fb}^{-1}$
50 GeV	49997	16.8 ± 0.19	6.20 ± 0.11	0.432 ± 0.026	18340 ± 1104
70 GeV	49999	5.15 ± 0.10	9.80 ± 0.14	0.323 ± 0.022	2709 ± 185
75 GeV	49700	4.71 ± 0.098	12.6 ± 0.16	0.429 ± 0.027	1360 ± 85.6
80 GeV	49994	6.02 ± 0.15	20.1 ± 0.20	1.07 ± 0.050	798 ± 37.3
85 GeV	49997	7.46 ± 0.12	30.7 ± 0.25	1.20 ± 0.046	483 ± 18.6
90 GeV	48192	10.3 ± 0.15	34.7 ± 0.27	2.60 ± 0.071	714 ± 19.5
95 GeV	49993	18.2 ± 0.19	36.9 ± 0.27	5.40 ± 0.10	1036 ± 19.2
100 GeV	49994	25.1 ± 0.23	39.1 ± 0.28	7.87 ± 0.12	1074 ± 16.4
105 GeV	49995	30.8 ± 0.25	40.8 ± 0.29	10.0 ± 0.14	1013 ± 14.2
110 GeV	49694	36.2 ± 0.27	42.4 ± 0.29	11.9 ± 0.15	926 ± 11.7
130 GeV	48197	46.7 ± 0.31	46.9 ± 0.31	17.0 ± 0.18	576 ± 6.09
150 GeV	49997	51.4 ± 0.32	51.3 ± 0.32	20.7 ± 0.19	359 ± 3.03
170 GeV	48196	54.7 ± 0.34	55.2 ± 0.34	23.3 ± 0.21	233 ± 2.10
190 GeV	49999	56.7 ± 0.34	59.2 ± 0.34	26.3 ± 0.22	163 ± 1.36
210 GeV	47300	58.4 ± 0.35	62.2 ± 0.36	28.8 ± 0.24	117 ± 0.98

Table 5.15: Muon channel event selection efficiencies for the tight selection. Errors are statistical only.

haviour. The efficiencies for the signal MC as a function of Majorana neutrino mass are shown in Table 5.16 for the loose selection and in Table 5.17 for the tight selection.

5.3.2 Data-MC scale factor

The accuracy of the simulation of the signal events can be examined via a tag-and-probe method on Z boson decays to $\ell^+\ell^-$. We follow this method which has been used successfully in previous CMS analyses [45, 13, 14, 15].

Using single muon triggered data we require that the tag muon passes our full muon selection criteria, while the other muon (the probe muon) must pass:

- $p_T > 5$ GeV
- $|\eta| < 2.4$
- be both a Tracker and Global muon type

Mass (GeV)	e p_T, η(%)	jet p_T, η(%)	e Quality (%)	Jet Quality (%)	Other (%)	All (%)
50	28.77 ± 0.27	73.80 ± 0.94	39.93 ± 0.73	13.05 ± 0.59	81.19 ± 5.16	0.90 ± 0.04
60	27.39 ± 0.26	67.12 ± 0.91	35.20 ± 0.72	19.07 ± 0.84	64.34 ± 4.14	0.79 ± 0.04
70	12.63 ± 0.17	74.89 ± 1.45	28.29 ± 0.88	26.79 ± 1.60	69.01 ± 5.73	0.49 ± 0.03
80	20.01 ± 0.22	90.98 ± 1.32	37.88 ± 0.76	28.97 ± 1.04	88.39 ± 4.08	1.77 ± 0.06
90	23.64 ± 0.24	94.95 ± 1.25	28.83 ± 0.58	42.65 ± 1.37	89.42 ± 3.50	2.47 ± 0.07
100	44.45 ± 0.36	95.64 ± 0.92	39.55 ± 0.51	46.67 ± 0.90	89.60 ± 2.08	7.03 ± 0.12
110	52.94 ± 0.41	97.87 ± 0.87	47.46 ± 0.53	46.12 ± 0.75	96.60 ± 1.86	10.95 ± 0.16
120	57.61 ± 0.48	98.80 ± 0.93	52.86 ± 0.60	46.09 ± 0.75	98.98 ± 1.90	13.72 ± 0.20
130	60.56 ± 0.47	99.10 ± 0.87	56.62 ± 0.58	46.78 ± 0.68	99.41 ± 1.69	15.80 ± 0.21
140	63.07 ± 0.51	99.22 ± 0.89	58.24 ± 0.61	47.72 ± 0.70	99.67 ± 1.69	17.34 ± 0.23
150	65.38 ± 0.48	99.40 ± 0.80	61.02 ± 0.56	48.19 ± 0.62	99.74 ± 1.48	19.06 ± 0.22
160	66.83 ± 0.52	99.48 ± 0.86	61.78 ± 0.61	48.76 ± 0.66	99.83 ± 1.57	19.99 ± 0.24
170	68.44 ± 0.50	99.56 ± 0.79	63.26 ± 0.57	48.88 ± 0.60	99.79 ± 1.42	21.02 ± 0.23
180	69.54 ± 0.52	99.61 ± 0.81	64.59 ± 0.60	49.80 ± 0.62	99.82 ± 1.43	22.24 ± 0.25
190	70.94 ± 0.52	99.64 ± 0.80	65.88 ± 0.59	50.12 ± 0.61	99.88 ± 1.39	23.31 ± 0.26
200	71.80 ± 0.534	499.68 ± 0.80	66.04 ± 0.59	50.74 ± 0.61	99.83 ± 1.38	23.94 ± 0.26

Table 5.16: Electron channel event selection efficiencies as a function of mass determined from the Majorana signal Monte Carlo for the loose event selection. The column labeled "Other" represents all selection requirement not listed in the preceding columns. Uncertainties are statistical only.

Mass (GeV)	e p_T, η(%)	jet p_T, η(%)	e Quality (%)	Jet Quality (%)	Other (%)	All (%)
50	27.93 ± 0.27	37.88 ± 0.61	35.48 ± 0.95	12.78 ± 0.88	82.79 ± 7.94	0.40 ± 0.03
60	26.94 ± 0.26	32.89 ± 0.57	32.32 ± 0.98	19.55 ± 1.28	68.41 ± 6.42	0.38 ± 0.03
70	12.17 ± 0.17	47.08 ± 1.10	27.71 ± 1.15	20.29 ± 1.81	74.02 ± 9.26	0.24 ± 0.02
75	9.374 ± 0.14	61.87 ± 1.47	33.61 ± 1.25	19.83 ± 1.57	84.24 ± 8.99	0.33 ± 0.03
80	19.68 ± 0.22	71.51 ± 1.12	39.15 ± 0.88	20.71 ± 0.95	91.08 ± 5.52	1.04 ± 0.05
85	16.45 ± 0.20	80.68 ± 1.33	29.33 ± 0.76	24.52 ± 1.25	90.87 ± 6.03	0.87 ± 0.04
90	22.58 ± 0.24	84.41 ± 1.17	27.37 ± 0.60	32.77 ± 1.29	89.45 ± 4.45	1.53 ± 0.06
95	35.14 ± 0.31	83.13 ± 0.93	33.10 ± 0.55	35.40 ± 1.00	88.47 ± 3.13	3.03 ± 0.08
100	41.93 ± 0.34	85.27 ± 0.87	39.25 ± 0.55	36.84 ± 0.85	90.50 ± 2.58	4.68 ± 0.10
105	47.66 ± 0.38	88.16 ± 0.85	44.49 ± 0.56	35.72 ± 0.73	94.51 ± 2.39	6.31 ± 0.12
110	50.35 ± 0.39	91.17 ± 0.83	47.37 ± 0.55	34.73 ± 0.66	96.99 ± 2.25	7.32 ± 0.13
120	55.01 ± 0.41	94.93 ± 0.82	53.11 ± 0.56	35.49 ± 0.59	98.99 ± 2.00	9.74 ± 0.15
130	58.12 ± 0.43	96.48 ± 0.81	55.96 ± 0.56	36.84 ± 0.57	99.33 ± 1.85	11.48 ± 0.16
150	61.81 ± 0.45	97.99 ± 0.79	60.14 ± 0.56	37.71 ± 0.53	99.77 ± 1.70	13.71 ± 0.18
170	65.66 ± 0.47	98.63 ± 0.77	63.23 ± 0.56	40.46 ± 0.53	99.81 ± 1.55	16.54 ± 0.20
190	68.80 ± 0.48	99.00 ± 0.76	65.30 ± 0.56	42.20 ± 0.52	99.75 ± 1.46	18.72 ± 0.21
210	69.50 ± 0.49	99.30 ± 0.75	67.34 ± 0.57	44.39 ± 0.53	99.83 ± 1.39	20.60 ± 0.22

Table 5.17: Electron channel event selection efficiencies as a function of mass determined from the Majorana signal Monte Carlo for the tight event selection. The column labeled "Other" represents all selection requirement not listed in the preceding columns. Uncertainties are statistical only.

The tag-probe pair is required to have a mass that lies within the same mass window which we use to eliminate Z background events (71-111 GeV). Thus our tag-and-probe sample gives us an independent control sample. We measure the efficiency of the muon track quality cuts and isolation cuts separately.

A detailed description of the method used to extract the efficiencies is discussed in References [45, 15]. In summary, the efficiency is $\epsilon = TP/(TP + TF)$ where TP is the number of probes that pass the selection cuts being studied (isolation or track quality) and TF is the number that fail. These numbers are found from fitting the tag-probe muon pair mass distributions with a signal model taken from simulation and an exponential to describe the background. In the kinematic range of the muons produced by our signal we find that the muons are well modelled by the simulation as seen in Table 5.18, for the electron channel refer to Table 5.19. A systematic error of 1% on the muon identification and isolation is estimated by comparing these results to the same calculation made by counting the number of di-muon tag probe pairs passing the selection criteria whose masses lies within 71 and 111 GeV. The background contribution can be estimated by looking at the same-sign pairs in the same mass window. There is an additional systematic uncertainty on the efficiency of finding isolated muons with low momentum; however, the effect is too small to be seen in our results.

Criteria	p_T range (GeV/c)	Data	MC	Data/MC
Iso	$10 - 20$	0.736 ± 0.003	0.753 ± 0.006	0.978 ± 0.009
	> 20	0.893 ± 0.001	0.903 ± 0.001	0.990 ± 0.002
ID	$10 - 20$	0.987 ± 0.001	0.989 ± 0.001	0.998 ± 0.002
	> 20	0.997 ± 0.004	0.990 ± 0.007	0.993 ± 0.008
Total	$10 - 20$			0.975 ± 0.009
	> 20			0.983 ± 0.008

Table 5.18: Muon identification and isolation efficiencies for both simulation and data measured with the tag and probe method.

Criteria	p_T range (GeV/c)	Data	MC	Data/MC
Iso	$10 - 20$	0.884 ± 0.012	0.927 ± 0.009	1.000 ± 0.002
	> 20	0.938 ± 0.001	0.980 ± 0.001	1.003 ± 0.001
ID	$10 - 20$	0.624 ± 0.015	0.673 ± 0.013	0.921 ± 0.002
	> 20	0.786 ± 0.003	0.825 ± 0.003	0.950 ± 0.001
Total	$10 - 20$			0.877 ± 0.046
	> 20			0.957 ± 0.016

Table 5.19: Electron identification and isolation efficiencies for both simulation and data measured with the tag and probe method.

Chapter 6

Background Estimation

In this chapter we explain the techniques used to estimated the background. The background can be split into two main categories: events with real same-sign lepton pairs and events with one or more misidentified leptons. In our description we will briefly illustrate the MCs used for the prediction of the real same-sign events and then explain in details the methodology we followed for the data-driven fake background prediction. Particular emphasis will be put on the muon channel since the method used for the electron channel is the same. For the electron channel the background from charge miss-measurement has to be considered while it will be shown to be negligible for the muon channel.

6.1 Backgrounds due to SM sources of same-sign events

Standard Model sources of same-sign lepton pairs have a relatively small cross-section and we use Monte Carlo to predict them. These sources originate primarily from WZ, same-sign W production via double W-strahlung, ZZ, $W\gamma$, $t\bar{t}W$ and $W^{\pm}W^{\pm}$ production via double parton scattering ($2 \times q\bar{q}' \rightarrow W$).

The MC samples used are listed in Tables 5.3, 5.4 and 5.5 along with the corresponding cross-sections. The larger, by cross-section, is the WZ process with 18.2 pb while all the others have much smaller cross-sections. Hence it is possible to achieve small statistical uncertainty on the MC with a reasonable sample size.

As already mentioned, all the MC samples had to be weighted to match the distribution of the number of vertices we saw in data (see Fig 6.1). In particular MC

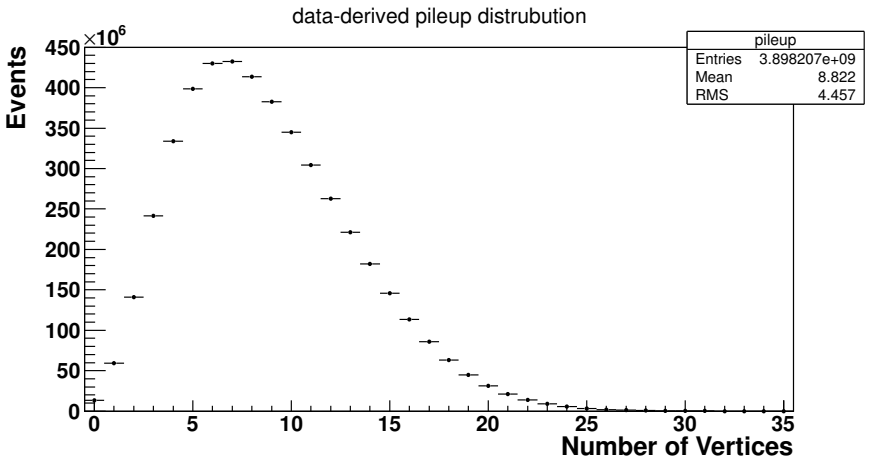


Figure 6.1: Distribution of the number of reconstructed vertices in 2011 data. This represents the number of underlying events as well, since the number of vertices is equal to the number of underlying events times 0.7 that is the efficiency to reconstruct a good vertex.

samples produced using the pileup distribution present during Summer 2011 data taking period (belonging to the so called Summer11 campaign) do not really represent the actual number of interactions per crossing, since most of the 2011 data was collected towards the end of the year with higher pileup (see Fig 6.2). Starting from the MCs produced during the Fall 2011 we can see a better matching, even without reweighing, as shown in Fig. 6.3. Anyway we have applied the same reweighing technique to all MC samples in order to obtain a good matching.

As mentioned in Sec. 5.3.2 we have measured the Data-MC scale factor and

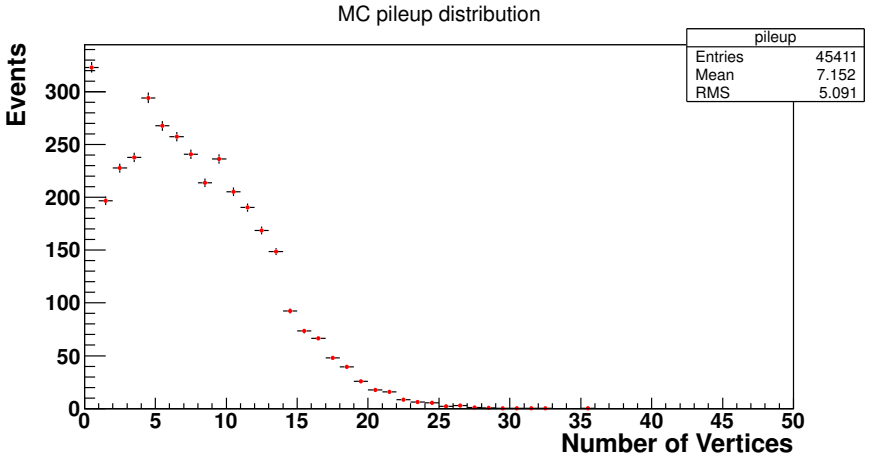


Figure 6.2: Distribution of the number of vertices in a $W\gamma$ MC sample produced during the Summer11 campaign. The distribution does not match the one in Fig 6.1.

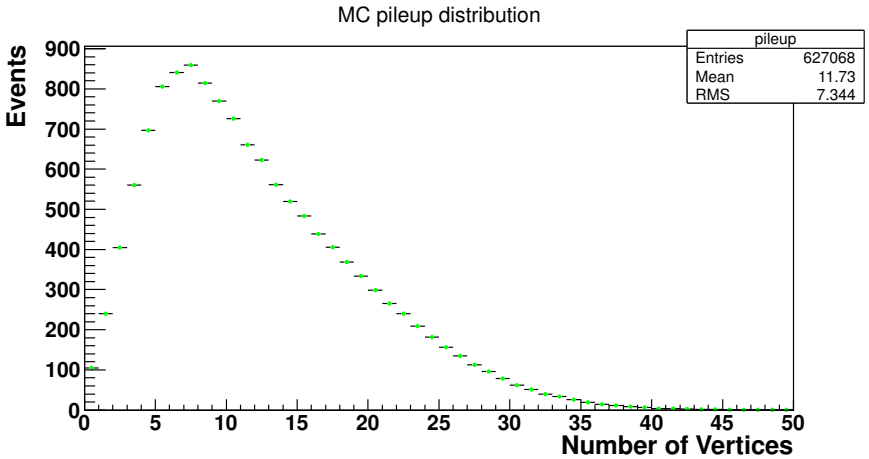


Figure 6.3: Distribution of the number of vertices in a WZ MC sample produced during the Fall11 campaign. The distribution has a good level of agreement with the final distribution observed in data as shown in Fig 6.1.

the trigger efficiency. We have decided to measure it directly from data rather than rely on the simulated triggers because of the changing conditions through the year, and the difficulties the MC had to exactly match the running conditions as we just saw in the number of vertices distributions.

6.2 Data-driven background prediction for fake muons

The main background for this analysis consists in "non-prompt" muons reconstructed as "prompt" muons. As already mentioned we define prompt muons as those produced by the decay of W or Z bosons, as well as muons produced by the heavy Majorana neutrino. These muons are well isolated and energetic enough to escape the detector, and point back to the primary vertex. Non-prompt, or fake muons, originate from jets and are typically not well isolated. These muons can be genuine muons from b-jets or muons from decays in flight, for example. These muons can contaminate our signal search region. The detailed description of the method used to predict this background is discussed in the following section.

This type of background cannot be accurately predicted using only Monte Carlo samples because the processes involved have large cross-sections with small fake muon probabilities. Furthermore, it is difficult to model the jets fragmentation process to the desired level of accuracy. Among the larger contributors to the fake background we can mention $t\bar{t}$ (with a cross-section of 154 pb), single-top processes (with cross-section from 1.2 to 35.7 pb), W+jets ($\sigma = 3.13 \times 10^4$ pb), and QCD multijet production (cross-section lager than 10^9 pb).

6.2.1 The fake rate method

To estimate the background due to fake muons we use the so called "fake rate method", which basically consists of measuring the fake rate (or tight-to-loose ratio) in QCD multijet enhanced events and then using it to weight events passing the nominal selection except for the muons that are divided into two classes:

- tight muons are defined using the signal selection (as described in Sec. 5.2.3),

- loose muons are selected using similar criteria as the tight muons but one or more variables are relaxed (generally $RelIso$, d_0, or $\chi^2/ndof$ of the global muon fit).

The weighting procedure is derived as follows:

$$N_{tot} = N_{pp} + N_{pl} + N_{ff} = N_{nn} + N_{n\bar{n}} + N_{\bar{n}\bar{n}} \tag{6.1}$$

$$N_{\bar{n}\bar{n}} = (1 - FR)^2 \, N_{ff} \tag{6.2}$$

$$N_{n\bar{n}} = (1 - FR) \, N_{pf} + 2 \, FR \, (1 - FR) \, N_{ff} \tag{6.3}$$

$$N_{nn} = N_{pp} + FR \, N_{pf} + FR^2 \, N_{ff} \tag{6.4}$$

where for simplicity we are assuming an average fake rate FR. N_{pp}, N_{pf}, and N_{ff} are the true number of events with two prompt, one prompt and one fake, and two fake muons respectively; N_{nn} is the number of events where both muons pass the tight selection, $N_{n\bar{n}}$ is the number of events in which one muon passes the tight selection and the other one fails the tight selection but passes the loose selection, and $N_{\bar{n}\bar{n}}$ is the number of events in which both muons fail the tight selection but pass the loose selection. Note that the factor of 2 in the second term on the right hand side of Eq. 6.3 is explained by considering that each of the two loose muons can fake a tight one; in the case that the two muons have different fake rates it would read "$FR_i(1-FR_j) + FR_j(1-FR_i)$". We can invert these equations and find the actual number of prompt-prompt events given the measured $N_{n\bar{n}}$ and $N_{\bar{n}\bar{n}}$ quantities and the fake rate FR:

$$N_{pp} = \sum_{i,j} \frac{FR_i\ FR_j}{(1-FR_i)(1-FR_j)} N^{ij}_{\bar{n}\bar{n}} + \sum_{i,j} \frac{FR_i}{(1-FR_i)} \left[N^{ij}_{n\bar{n}} - \frac{FR_i(1-FR_j) + FR_j(1-FR_i)}{(1-FR_i)(1-FR_j)} N^{ij}_{\bar{n}\bar{n}} \right] \tag{6.5}$$

The indices i and j label the two muons p_T and η, the summation is indeed performed over four indices independently for each muon momentum and pseudorapidity. The FR is itself binned in p_T and η in order to achieve a better description of the detector response at different angles and to different muon energies.

Eq. 6.5 represents an evolution of the standard equations used for this kind of background prediction technique. In general the following equation has been used:

$$N_{pp} = \sum_{i,j} \frac{FR_i\ FR_j}{(1-FR_i)(1-FR_j)} N^{ij}_{\bar{n}\bar{n}} + \sum_{i} \frac{FR_i}{(1-FR_i)} N^{i}_{n\bar{n}} - 2 \sum_{i,j} \frac{FR_i\ FR_j}{(1-FR_i)(1-FR_j)} N^{ij}_{\bar{n}\bar{n}} \tag{6.6}$$

The first term of the first line of both Eq. 6.5 and 6.6 represents the number of "double fake" events, i.e. the number of events in which two loose muons fake an event with two tight muons. The second line represents the number of "single fake" events, i.e. events in which only one muon is a fake. Of course, the number of events of type $n\bar{n}$ is subject to contamination from $\bar{n}\bar{n}$ events that can also yield events with one tight and one loose muon (as well as two tight muons). Therefore, taking the raw number of $N_{n\bar{n}}$ events will give an over-prediction of the expected prompt-prompt event counts. To avoid this in Eq. 6.6 the total number of predicted single fake events is corrected by subtracting twice the amount of double fake events (notice the factor 2 coming from Eq. 6.3). This sort

of correction looses some information since it averages over the p_T and η of the muons, and implicitly assumes that the fake rates are roughly constant in p_T and η and/or the amount of double fake events is small.

In equation 6.5, the correction is applied independently for each p_T and η bin (the i and j indices each represent the two variables p_T and η for each muon). In other words, the total count of events $N_{n\bar{n}}$ at high η and low p_T, for example, will be corrected by the predicted "double-to-single" fake events[1] at high η and low p_T and not by an average over all momenta and pseudo-rapidities. After we obtain the corrected number then we weight it properly to obtain the expected single fake prediction, using Eq. 6.5.

6.2.2 Fake rate event selection

Starting from the nominal selection criteria mentioned in Sec. 5.2.3 we first need to define the criteria used to defined loose muons. In general fake muons are defined to have large *RelIso*, transverse impact parameter, and $\chi^2/ndof$, hence it seems predictable to relax these three variables for the loose object definition.

Some early Monte Carlo studies were done using different settings to establish which variables to relax and the allowed values. In particular we have changed the following:

- $RelIso < 0.4$ (nominal is 0.1),
- $d_0 < 2$ mm (nominal is 0.2 mm),
- $\chi^2/ndof < 50$ (nominal is 10).

Using this selection we can define two sets of muons the "tight", with all muons passing the nominal cuts in Sec. 5.2.3, and the "loose", that includes muons passing the

[1] By double-to-single we indicate events with two loose muons that fake events with one prompt and one loose muons

relaxed selection. The second group of muons includes mainly misidentified of fake muons.

Another simpler selection was done relaxing only one variable:

- $RelIso < 0.4$

We exploited different choices for the loose cuts to obtain the best possible description of the background and we will illustrate them in the following sections, after discussing in detail the selection used to measured the fake rates.

Using an event selection meant to maximize the QCD event selection, we measured the tight-to-loose ratio (T/L), or fake rate (FR), defined as the ratio of events passing the tight muon selection divided by the events passing the loose muon selection.

The selection of the events plays an important role in the FR measurement on which this method is based. As we have mentioned we have decided to use the `SingleMu` triggers and datasets to compute the fake rate. This choice helps in getting more statistics in the event selection at the price of higher electroweak contamination. In fact we are interested in measuring the FR in QCD events and not in W or Z decays (for which the triggers are defined) which contain prompt muons by definition.

Muons and jets are defined using the basic requirements listed in Sec. 5.2, then we select events with $E_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV and $M_T < 20$ GeV to suppress $W \to \mu\nu$ events. We also exclude events with di-muon invariant mass within 20 GeV of the Z boson mass, and to further suppress electroweak backgrounds we require events to have only one muon.

Besides the muon selection also the jet selection is very important as we will see. The baseline jet selection requires at least one good jet with $p_T > 40$ GeV and well separated from the muon by at least $\Delta R > 1.0$ (referred to as the "away-side jet").

The jet requirement is the best handle we have to select the kind of interaction we want to map. The "fakable object", i.e. the jet that produces the fake muon, has different characteristics according to the process that it is generated from. In particular from Fig. 6.4 we can immediately see, as expected, that the spectrum of $t\bar{t}$ events is harder than QCD.

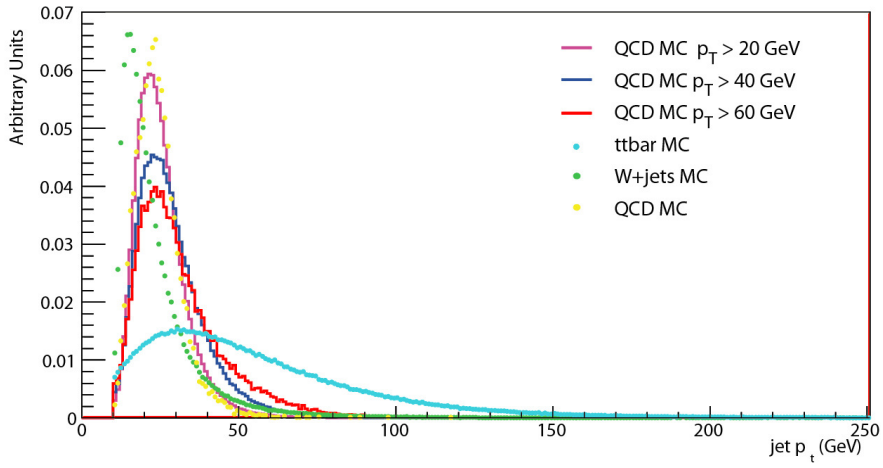


Figure 6.4: Distribution of the p_T of the fakable objects for different MC samples: cyan dots for $t\bar{t}$, green W+jets and yellow QCD. The pink, blue and red solid lines are the distributions obtained cutting at 20, 40 and 60 GeV respectively in an independent QCD used to calculate the FR.

Our signal selection includes two jets of at least 20 GeV (or 30 GeV) and given that $t\bar{t}$ is expected to be a small part of the total background, our final choice of the cut is decided to be $p_T > 40$ GeV. We have performed studies to check the stability of the final background prediction versus this parameter, and assigned a corresponding systematic uncertainty to this choice (see Sec 7.2).

In order to take into account different detector responses and different physics processes, and eventually achieve the most accurate possible background prediction, we have calculated the FR in bins of η and p_T. Unfortunately, even with all the mentioned cuts in the fakes rate selection, the number of events from the decay of W and Z

bosons still represents a significant fraction of the total number of events at high energy, where the QCD multijet spectrum rapidly falls and the $W \to \mu\nu$ peak starts to appear. Therefore we have made the choice to cut the FR at 35 GeV in muon p_T and assume a flat FR for $p_T > 35$ GeV.

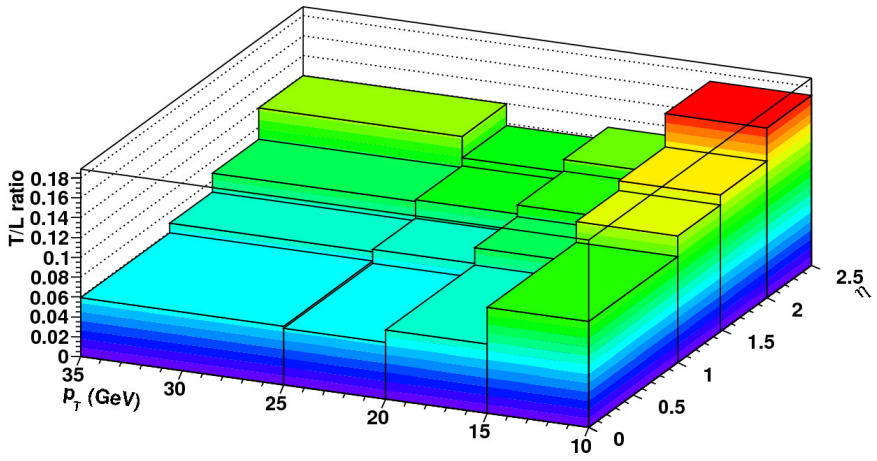


Figure 6.5: Muon fake rates obtained from data as a function of η and p_T.

Nonetheless, we correct the fake rates for possible contamination from electroweak processes. To do this we run the same FR selection on W and Z Monte Carlo, then weight the passing events by the effective luminosity of the triggers used in data. These events are then subtracted from the numerator and denominator of the T/L ratio before calculating the final ratio. The correction reduces the fake rate values by about 2% in the low p_T bins and up to 20% in the high p_T bins, where the contamination from electroweak processes is larger.

6.2.3 Calibration of the background prediction

The first step was to validate our background prediction using MC samples. We started using the $t\bar{t}$ sample since it has larger statistics. We have used a QCD μ enriched sample to calculate the FR using different selections for the loose object and

the minimum p_T of the away jet.

We have first relaxed all the three variables for defining the loose muons selection ($RelIso$, d_0 and $\chi^2/ndof$) applying the cut on the jet at 40 GeV. Then we have also tried relaxing only $RelIso$ in order to study possible bias and better understand our background description.

Once we obtained the FR as a function of p_T and η we applied the background prediction method to the $t\bar{t}$ MC to check if the predicted and expected results agree. In order to maintain high statical significance of this test we do not apply the $E_{\mathrm{T}}^{\mathrm{miss}} <$ 50 GeV cut used in the final selection. From this closure test we obtained the following results:

- prediction with three relaxed variables: 159 ± 5,
- prediction with $RelIso < 0.4$: 152 ± 5,

while the number of MC truth events is 116. Because the single variable relaxation gives better results we have decided to define our loose selection using only the relative isolation: this has the disadvantage to not perfectly describe backgrounds from displaced vertices with isolated leptons (such as rare $W\gamma^* \to \ell\nu\ell\ell$), but our background consists mostly of not isolated muons, so this is not an issue but they have to be included among the backgrounds predicted using Monte Carlo simulations. The choice of relaxing one variable only is also useful to better investigate systematics, understand, and eventually reduce them more effectively.

The first thing to notice is the large over-prediction of about 50%. To explain this disagreement we have to go back to Fig. 6.4. From the plot it is clear that the QCD spectrum is much softer than the $t\bar{t}$ and hence there is a bias when we calculate the fake rates from the QCD MC. Therefore, in principle, increasing the away-side jet

p_T requirement on the fake rates event selection from 40 to 60 GeV should reduce the discrepancy. In Table 6.2 we show the fake rate matrix using the standard jet p_T cut at 40 GeV, while in Tables 6.1 and 6.3 we show the resulting fake rates using a p_T cut of 20 and 60 GeV respectively.

$\lvert\eta\rvert \setminus p_T$	$10-15$ GeV	$15-20$ GeV	$20-25$ GeV	$25-35$ GeV
$0-1.0$	0.2314 ± 0.0009	0.1600 ± 0.0005	0.1285 ± 0.0008	0.1059 ± 0.0012
$1.0-1.479$	0.2583 ± 0.0013	0.1895 ± 0.0007	0.1591 ± 0.0013	0.1295 ± 0.0019
$1.479-2.0$	0.2743 ± 0.0015	0.2147 ± 0.0008	0.1864 ± 0.0015	0.1619 ± 0.0023
$2.0-2.5$	0.2911 ± 0.0023	0.2315 ± 0.0014	0.1942 ± 0.0024	0.1734 ± 0.0038

Table 6.1: Muon fake rates obtained from QCD MC with the away-side jet $p_T > 20$ GeV in bins of $|\eta|$ and p_T. Statistical error only.

$\lvert\eta\rvert \setminus p_T$	$10-15$ GeV	$15-20$ GeV	$20-25$ GeV	$25-35$ GeV
$0-1.0$	0.2160 ± 0.0023	0.1280 ± 0.0010	0.0915 ± 0.0014	0.0778 ± 0.0016
$1.0-1.479$	0.2405 ± 0.0035	0.1492 ± 0.0016	0.1193 ± 0.0022	0.1001 ± 0.0025
$1.479-2.0$	0.2551 ± 0.0038	0.1740 ± 0.0018	0.1353 ± 0.0024	0.1253 ± 0.0030
$2.0-2.5$	0.2674 ± 0.0057	0.1901 ± 0.0029	0.1478 ± 0.0040	0.1324 ± 0.0050

Table 6.2: Muon fake rates obtained from QCD MC with the away-side jet $p_T > 40$ GeV in bins of $|\eta|$ and p_T. Statistical error only.

$\lvert\eta\rvert \setminus p_T$	$10-15$ GeV	$15-20$ GeV	$20-25$ GeV	$25-35$ GeV
$0-1.0$	0.2076 ± 0.0056	0.1245 ± 0.0025	0.0822 ± 0.0031	0.0591 ± 0.0029
$1.0-1.479$	0.2301 ± 0.0084	0.1443 ± 0.0038	0.1004 ± 0.0047	0.0772 ± 0.0045
$1.479-2.0$	0.2574 ± 0.0094	0.1586 ± 0.0042	0.1237 ± 0.0054	0.0954 ± 0.0053
$2.0-2.5$	0.2425 ± 0.0135	0.1740 ± 0.0069	0.1337 ± 0.0090	0.1080 ± 0.0091

Table 6.3: Muon fake rates obtained from QCD MC with the away-side jet $p_T > 60$ GeV in bins of $|\eta|$ and p_T. Statistical error only.

Our intuition is confirmed by the fake rate results: cutting harder on the away jet reduces the FR and this will eventually reduce the final prediction. Considering Eq. 6.5 it is clear that in general, if the double-to-single correction is small, a reduction of FR results in a general reduction of the final prediction, and indeed the total prediction

drops from 152 ± 5 to 130 ± 6.

Of course, $t\bar{t}$ it is not the only process present in our background and we have to calibrate the technique using also QCD multijet and W+jets MCs, as well as real data. The main difference between $t\bar{t}$ and QCD multijet processes is the relative amount of single and double fakes; the first is dominated by single fakes while the second has only double fakes. In extreme cases the double-to-single correction can exceed the number of $N_{n\bar{n}}$ events and therefore the single fake prediction turns out to be negative. In order to avoid odd behaviours we have gauged the value of the loose cut on the $RelIso$ variable, looking for a value that ensures stability against small fluctuations and a non-negative single fake prediction.

Due to lack of sufficient MC statistics this calibration has been directly performed on data using a selection without $E_{\mathrm{T}}^{\mathrm{miss}}$ cut. Generally speaking we want the loose cut to be not too close to the tight cut in order to extrapolate over a not too small $RelIso$ range, and at the same time not too far away to avoid non linearities in the extrapolation. From our studies in data we have found the optimal $RelIso$ loose cut to be 0.8. This choice, together with the away jet p_T, affects the final background prediction and its effects will be considered among the systematic uncertainties.

Using some simple assumptions it is possible to calculate a functional form of the fake prediction starting from Eq. 6.5. We know that the T/L ratio decreases approximately in proportion to the *RelIso* of the loose object and is 1 when the loose object definition overlaps with the tight selection (at $RelIso = 0.1$). Therefore we can assume $FR \sim 0.1/x$ where x is $RelIso$. On the other hand, $N_{n\bar{n}}$ and $N_{\bar{n}\bar{n}}$ increase with $RelIso$, the first increases approximately linearly since one leg is fixed by the tight selection, while the second increases as $\sim x^2$ because possible combinations are added

inside the loose objects ensemble. Thus we can write Eq. 6.5 as:

$$N_{pp} = \frac{0.1}{1 - \frac{0.1}{x}} \times \left(1 - \frac{0.1}{1 - \frac{0.1}{x}}\right) \tag{6.7}$$

where we have assumed an average FR over p_T and η. Fig. 6.6 shows that the prediction changes rapidly when $RelIso \to 0.1$, so we use a definition of the loose object with a relatively large *RelIso* of 0.8 where N_{pp} is very stable, as we can also see in Fig. 6.7 where the derivative of Eq. 6.7 is plotted.

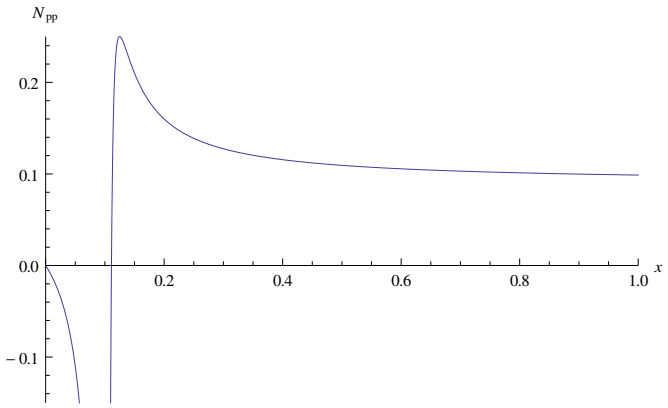


Figure 6.6: Plot of Eq.6.7, the x-axis represents *RelIso* and 0.1 is the cutting value of the signal selection, thus the observed discontinuity. The FR prediction is meaningless below it.

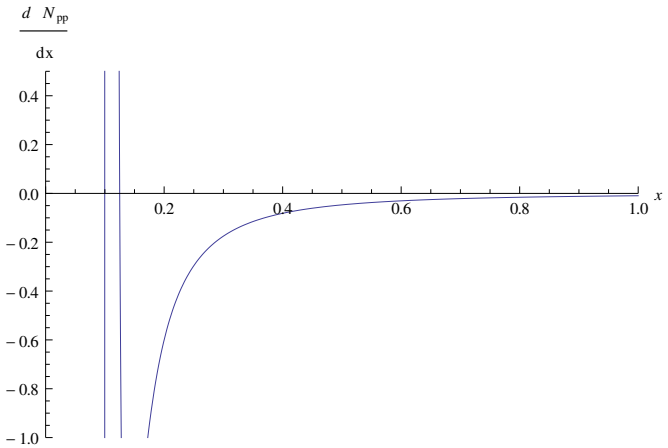


Figure 6.7: Plot of the derivative of Eq.6.7, the x-axis represents *RelIso*, as before 0.1 is the discontinuity due to the cutting value of the signal selection.

The final data fake rate matrix obtained from data, corrected for electroweak contamination, is shown in Table 6.4 and in Fig. 6.5. The selection was done using the away-side jet $p_T > 40$ GeV and loose object $RelIso < 0.8$. In Table 6.5 is the correspondent table for Monte Carlo QCD.

$\lvert\eta\rvert \setminus p_T$	$10 - 15$ GeV	$15 - 20$ GeV	$20 - 25$ GeV	$25 - 35$ GeV
$0 - 1.0$	0.1068 ± 0.0017	0.0698 ± 0.0010	0.0571 ± 0.0014	0.0593 ± 0.0017
$1.0 - 1.479$	0.1278 ± 0.0026	0.0874 ± 0.0016	0.0714 ± 0.0023	0.0690 ± 0.0028
$1.479 - 2.0$	0.1378 ± 0.0029	0.0992 ± 0.0019	0.0901 ± 0.0028	0.0882 ± 0.0035
$2.0 - 2.5$	0.1713 ± 0.0047	0.1107 ± 0.0033	0.0981 ± 0.0049	0.1196 ± 0.0068

Table 6.4: Muon fake rates obtained from data in bins of $|\eta|$ and p_T, using the away-side jet $p_T > 40$ GeV and loose object $RelIso < 0.8$

$\lvert\eta\rvert \setminus p_T$	$10 - 15$ GeV	$15 - 20$ GeV	$20 - 25$ GeV	$25 - 35$ GeV
$0 - 1.0$	0.0915 ± 0.0033	0.0545 ± 0.0005	0.0372 ± 0.0006	0.0298 ± 0.0007
$1.0 - 1.479$	0.0941 ± 0.0043	0.0689 ± 0.0008	0.0527 ± 0.0011	0.0444 ± 0.0013
$1.479 - 2.0$	0.1121 ± 0.0048	0.0829 ± 0.0010	0.0648 ± 0.0013	0.0596 ± 0.0017
$2.0 - 2.5$	0.1185 ± 0.0064	0.0898 ± 0.0016	0.0743 ± 0.0023	0.0645 ± 0.0030

Table 6.5: Muon fake rates obtained from MC in bins of $|\eta|$ and p_T, using the final selection with the away-side jet $p_T > 40$ GeV and loose object $RelIso < 0.8$.

We have also studied the stability of both efficiency of the signal selection and the tight-to-loose ratio vs the number of vertices. As we can see from Fig. 6.8, the pileup correction recovers some of the efficiency lost because of the increased number of underlying events (thus the energy in the event), as well as to obtain a more stable tight-to-loose ratio measurement.

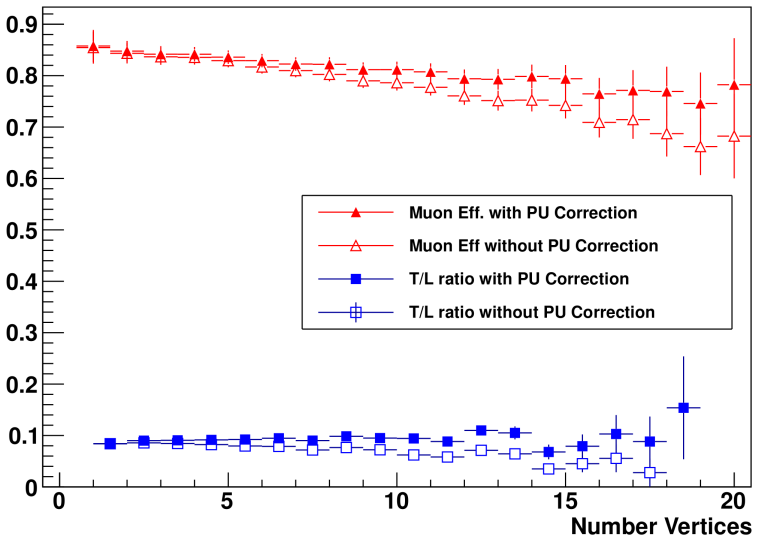


Figure 6.8: Efficiency (red triangles) and fake rate (blue squares) versus number of vertices with (solid markers) and without (empty markers) the correction for pileup applied.

6.2.4 Validation of the background prediction

6.2.4.1 Monte Carlo closure tests

Using the fake rates calculated in MC with the definition of $RelIso < 0.8$ for loose muons, we have performed a full set of Monte Carlo closure tests summarized in Table 6.6, these tests are performed with no $E_{\mathrm{T}}^{\mathrm{miss}} < 50$ GeV cut to increase statistics. The effect of the away-side jet p_T cut is shown for the $t\bar{t}$ closure test and, as already mentioned, the discrepancy between prediction and MC truth is reduced as the cut is increased. This effect is shown in the two histograms of Fig. 6.9 and Fig. 6.10 for jet $p_T > 40$ and 80 GeV respectively, which show the distributions of the invariant mass of the second highest p_T muon in the event and the two jets[2] . The shapes of the distributions agree reasonably well, while the overall prediction, as already explained,

[2]From MC studies, the choice of the second highest p_T muon is found to give better performance for reconstruction of the Majorana neutrino mass compared with taking either the highest p_T muon or randomly picking one of them.

is higher than the MC truth. In Fig.s 6.11 and 6.12 we show the distributions of the leading muon p_T.

The W+jets and QCD closure tests show good agreement. To increase the statistics for the QCD test we have performed an additional check. Since we do not expect any real single fakes to be present in the QCD MC samples, we know that the difference between the two terms in the square brackets of Eq. 6.5 must be zero. Thus we can count in MC the number of events with the one tight and one loose muons and predict it using the double-to-single correction. The results are listed on the last line of Table 6.6

Sample	p_T cut (GeV)	Predicted	MC truth	Ratio
$t\bar{t}$	40	182 ± 4	116	1.57 ± 0.03
$t\bar{t}$	60	156 ± 4	116	1.34 ± 0.03
$t\bar{t}$	80	143 ± 6	116	1.23 ± 0.05
W + jets	40	13.4 ± 1.1	12	1.12 ± 0.09
QCD	40	2.96 ± 0.72	3	0.98 ± 0.24
QCD double-to-single	40	112 ± 4	91	1.23 ± 0.04

Table 6.6: Predicted and observed fake backgrounds in $t\bar{t}$, W + jets, and QCD Monte Carlo samples for events passing the loose selection (signal event with jet $p_T > 20$ GeV). For the $t\bar{t}$ samples the effect of varying the cut on the p_T of the leading jet in the sample used to measure the fake rates is shown. Uncertainties are statistical only. The results are normalized to the total effective integrated luminosity of the Monte Carlo samples used.

We have performed the same set of tests for the tight signal selection with the two jets required to have $p_T > 30$ GeV. The results are summarized in Table 6.7. The slightly higher discrepancy is explained by the fact that now the event selection is done on "harder" signal events, while the fake rates were computed using the same selection with the 40 GeV cut on the away-side jet.

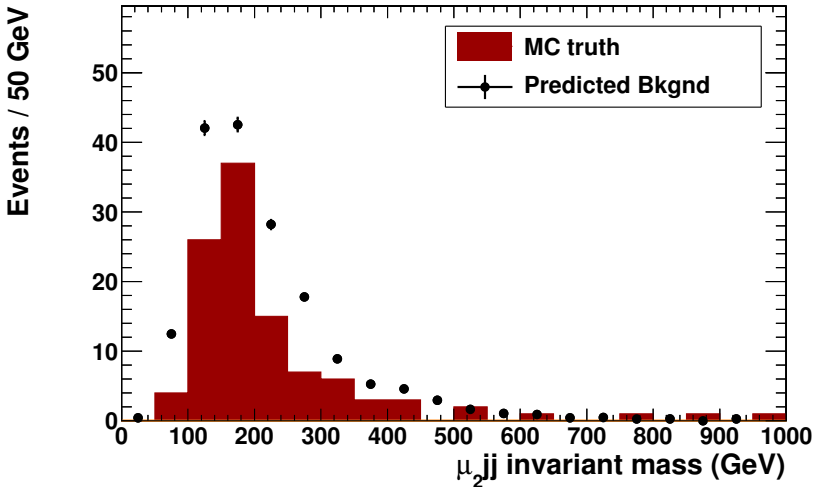


Figure 6.9: Closure test with $t\bar{t}$ MC with $p_T > 40$ GeV cut on the away jet. The invariant mass of the second leading muon with the two jets is shown, red histogram is the MC truth and the dots are the prediction with statistical error only.

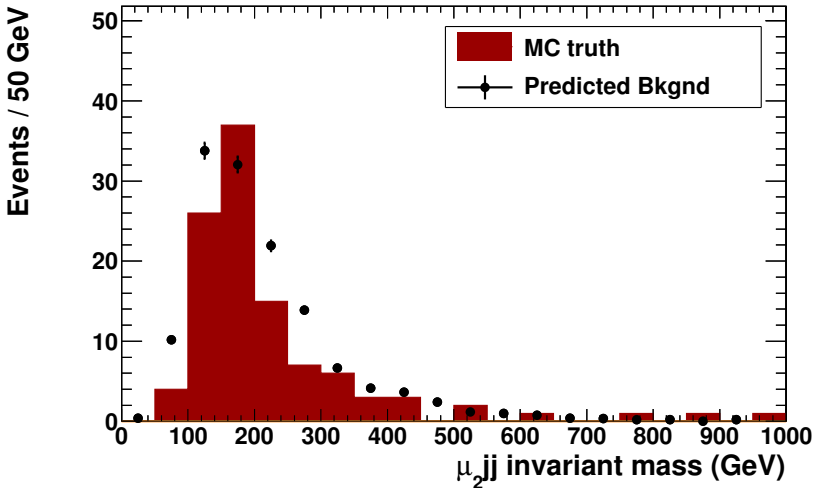


Figure 6.10: Closure test with $t\bar{t}$ MC with $p_T > 80$ GeV cut on the away jet. The invariant mass of the second leading muon with the two jets is shown, red histogram is the MC truth and the dots are the prediction with statistical error only.

The differences between the prediction and the MC truth are accounted for in the systematic uncertainty of 35% assigned to the overall background prediction (see Sec 7.2).

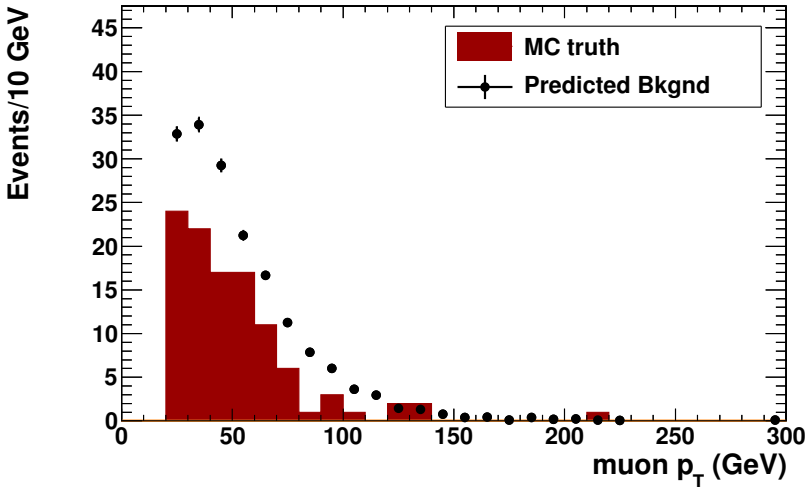


Figure 6.11: Closure test with $t\bar{t}$ MC with $p_T > 40$ GeV cut on the away jet. The momentum of the leading muon is shown, red histogram is the MC truth and the dots are the prediction with statistical error only.

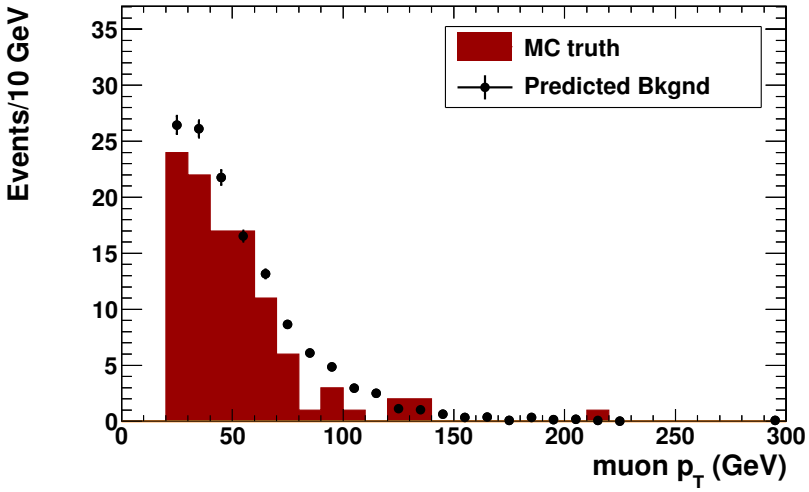


Figure 6.12: Closure test with $t\bar{t}$ MC with $p_T > 80$ GeV cut on the away jet. The momentum of the leading muon is shown, red histogram is the MC truth and the dots are the prediction with statistical error only.

6.2.4.2 Data closure test

Using the $E_{\mathrm{T}}^{\mathrm{miss}} < 50$ GeV cut in our signal selection we can define, in data, a control region for the background prediction technique. The only caveat is the different composition of the main backgrounds at high and low $E_{\mathrm{T}}^{\mathrm{miss}}$; the first is dominated by $t\bar{t}$ and W+jets, and the second by QCD on the other.

Sample	p_T cut (GeV)	Predicted	MC truth	Ratio
$t\bar{t}$	40	171 ± 4	107	1.59 ± 0.04
$t\bar{t}$	60	145 ± 4	107	1.36 ± 0.04
$t\bar{t}$	80	133 ± 6	107	1.25 ± 0.05
W + jets	40	5.98 ± 1.35	6	1.00 ± 0.23
QCD	40	2.23 ± 0.67	2	1.12 ± 0.34
QCD double-to-single	40	109 ± 4	80	1.36 ± 0.05

Table 6.7: Predicted and observed fake backgrounds in $t\bar{t}$, W + jets, and QCD Monte Carlo samples for events passing the tight selection (signal event with jet $p_T > 30$ GeV). For the $t\bar{t}$ samples the effect of varying the cut on the p_T of the leading jet in the sample used to measure the fake rates is shown. Uncertainties are statistical only. The results are normalized to the total effective integrated luminosity of the Monte Carlo samples used.

The results of this closure test in data using the loose event selection are shwon in Table 6.8 and Fig.s 6.13 - 6.16. The total background comprises the SM irreducible background estimated from MC (see Sec. 6.1) and misidentified (or fake) muon background estimated using the fake rate method described above and using the fake rates obtained from data listed in Table 6.4. The agreement between observation and prediction is remarkable.

Source	Events
Monte Carlo:	
WZ	5.85 ± 0.40 (stat.) ± 0.23 (syst.)
ZZ	0.39 ± 0.06 (stat.) ± 0.02 (syst.)
$W\gamma$	0.85 ± 0.31 (stat.) ± 0.09 (syst.)
$t\bar{t}W$	3.37 ± 0.10 (stat.) ± 1.69 (syst.)
W^+W^+	2.81 ± 0.13 (stat.) ± 1.41 (syst.)
W^-W^-	0.99 ± 0.04 (stat.) ± 0.50 (syst.)
dp $W^\pm W^\pm$	0.19 ± 0.03 (stat.) ± 0.10 (syst.)
Total Monte Carlo	14.47 ± 0.53 (stat.) ± 2.27 (syst.)
Data driven background estimate:	
Total fake background	56.82 ± 2.59 (stat.) ± 19.89 (syst.)
Total background	71.29 ± 2.65 (stat.) ± 20.02 (syst.)
Observed in data (4.98 fb^{-1})	63

Table 6.8: Results for the $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV control region for the loose selection. Observed event yields and estimated backgrounds with statistical and systematic uncertainties are shown.

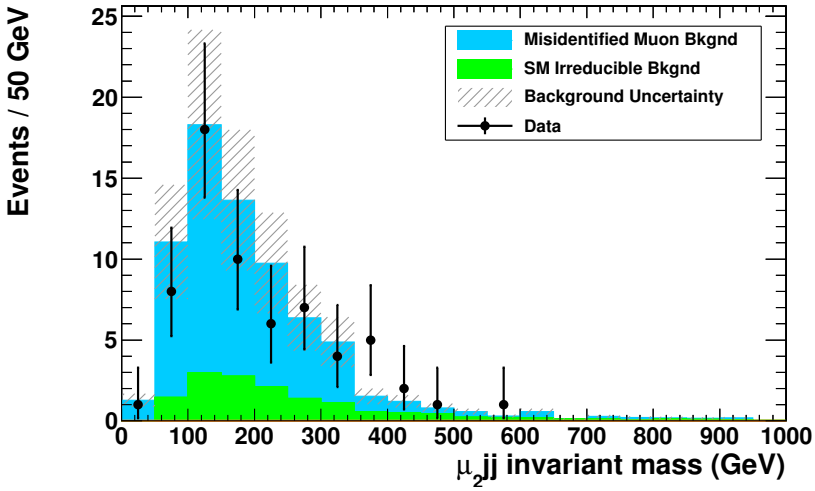


Figure 6.13: Closure test with data in the region defined by $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, the standard requirements are applied in the FR calculation (i.e. $p_T > 40$ GeV cut on the away jet and $RelIso < 0.8$ for the loose object) and the loose selection in signal selection (p_T(jet) > 20GeV). The invariant mass of the second leading muon with the two jets is shown. The black dots are data with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentified (fake) muon background prediction.

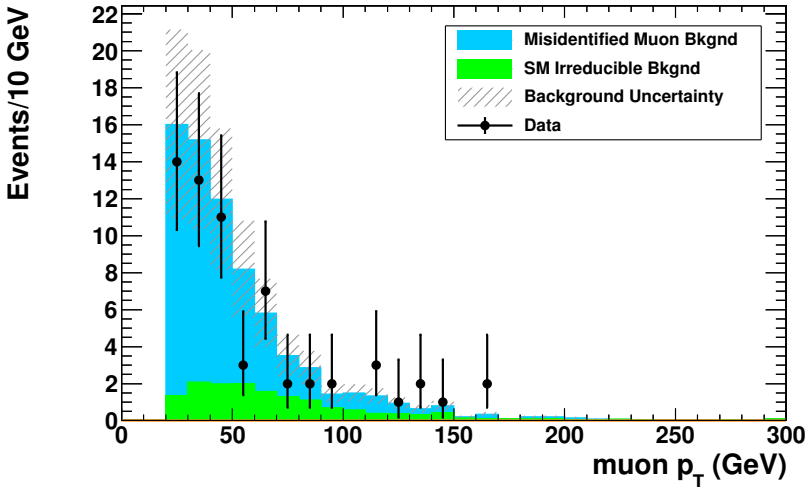


Figure 6.14: Closure test with data in the region defined by $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, the standard requirements are applied in the FR calculation (i.e. $p_T > 40$ GeV cut on the away jet and $RelIso < 0.8$ for the loose object) and the loose selection in signal selection (p_T(jet) > 20GeV). The leading muon p_T is shown. The black dots are data with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentified (fake) muon background prediction.

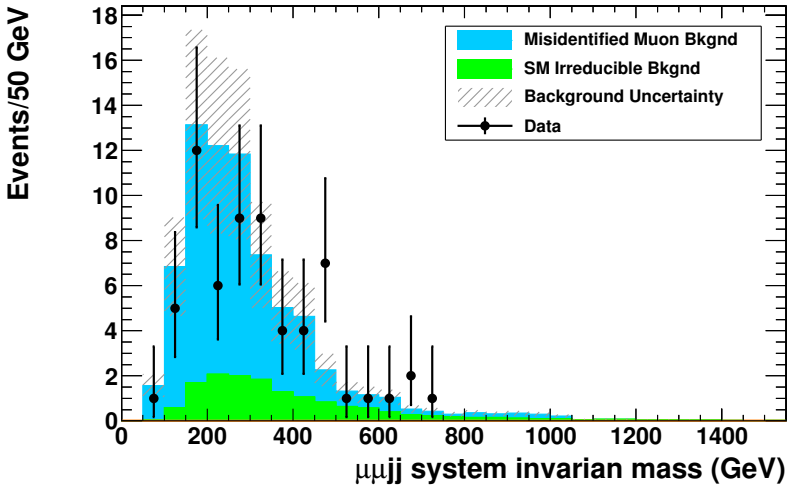


Figure 6.15: Closure test with data in the region defined by $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, the standard requirements are applied in the FR calculation (i.e. $p_T > 40$ GeV cut on the away jet and $RelIso < 0.8$ for the loose object) and the loose selection in signal selection (p_T(jet) > 20GeV). The invariant mass of the four selected particles is shown. The black dots are data with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentified (fake) muon background prediction.

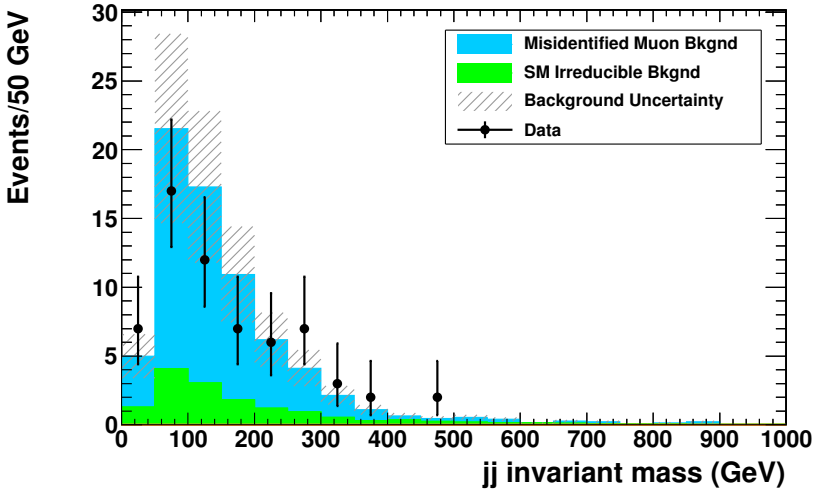


Figure 6.16: Closure test with data in the region defined by $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, the standard requirements are applied in the FR calculation (i.e. $p_T > 40$ GeV cut on the away jet and $RelIso < 0.8$ for the loose object) and the loose selection in signal selection (p_T(jet) > 20GeV). The invariant mass of the two selected jets is shown. The black dots are data with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentified (fake) muon background prediction.

The results for the same test but using the tight selection with the harder p_T cut on the jets ($p_T > 30$ GeV) are shown in Table 6.9, and Fig.s 6.17 - 6.20. Again, we see good agreement.

Source	Events
Monte Carlo:	
WZ	3.16 ± 0.30 (stat.) ± 0.13 (syst.)
ZZ	0.19 ± 0.04 (stat.) ± 0.01 (syst.)
$W\gamma$	0.31 ± 0.17 (stat.) ± 0.03 (syst.)
$t\bar{t}W$	3.19 ± 0.10 (stat.) ± 1.60 (syst.)
W^+W^+	2.57 ± 0.10 (stat.) ± 1.29 (syst.)
W^-W^-	0.84 ± 0.04 (stat.) ± 0.42 (syst.)
dp $W^\pm W^\pm$	0.08 ± 0.02 (stat.) ± 0.04 (syst.)
Total Monte Carlo	10.34 ± 0.37 (stat.) ± 2.10 (syst.)
Data driven background estimate:	
Total fake background	44.74 ± 2.00 (stat.) ± 15.66 (syst.)
Total background	55.08 ± 2.03 (stat.) ± 15.80 (syst.)
Observed in data (4.98 fb^{-1})	48

Table 6.9: Results for the $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV control region for the tight selection. Observed event yields and estimated backgrounds with statistical and systematic uncertainties are shown.

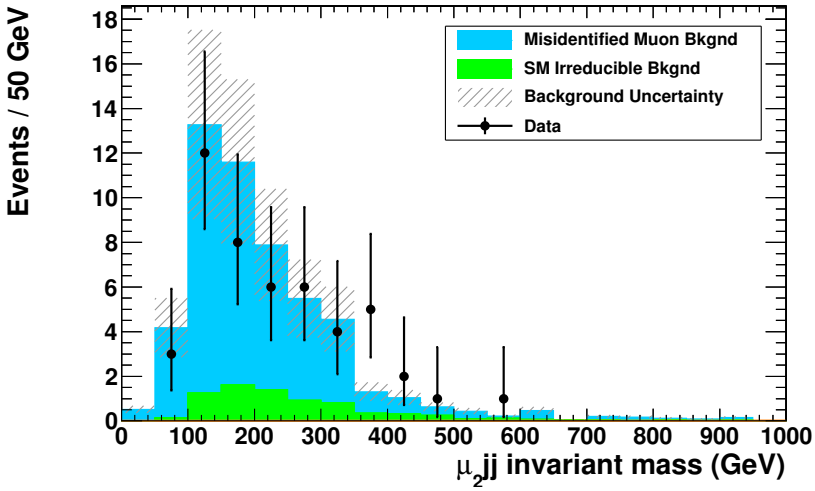


Figure 6.17: Closure test with data in the region defined by $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, the standard requirements are applied in the FR calculation (i.e. $p_T > 40$ GeV cut on the away jet and $RelIso < 0.8$ for the loose object) and the tight selection in signal selection (p_T(jet) > 30GeV). The invariant mass of the second leading muon with the two jets is shown. The black dots are points with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentification (fake) muon background prediction.

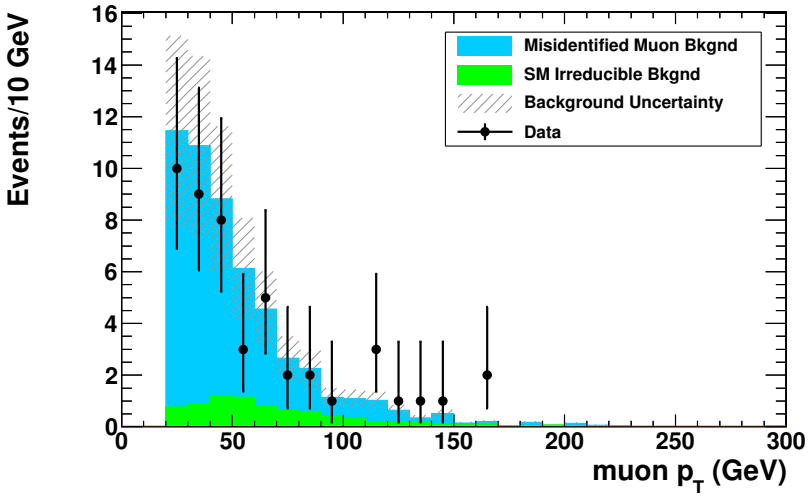


Figure 6.18: Closure test with data in the region defined by $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, the standard requirements are applied in the FR calculation (i.e. $p_T > 40$ GeV cut on the away jet and $RelIso < 0.8$ for the loose object) and the tight selection in signal selection (p_T(jet) > 30GeV). The leading muon p_T is shown. The black dots are points with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentification (fake) muon background prediction.

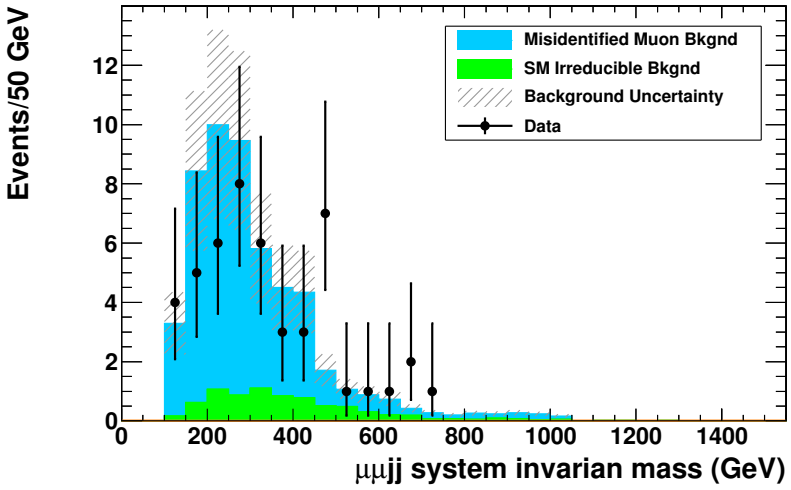


Figure 6.19: Closure test with data in the region defined by $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, the standard requirements are applied in the FR calculation (i.e. $p_T > 40$ GeV cut on the away jet and $RelIso < 0.8$ for the loose object) and the tight selection in signal selection (p_T(jet) > 30GeV). The invariant mass of the four selected particles is shown. The black dots are points with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentification (fake) muon background prediction.

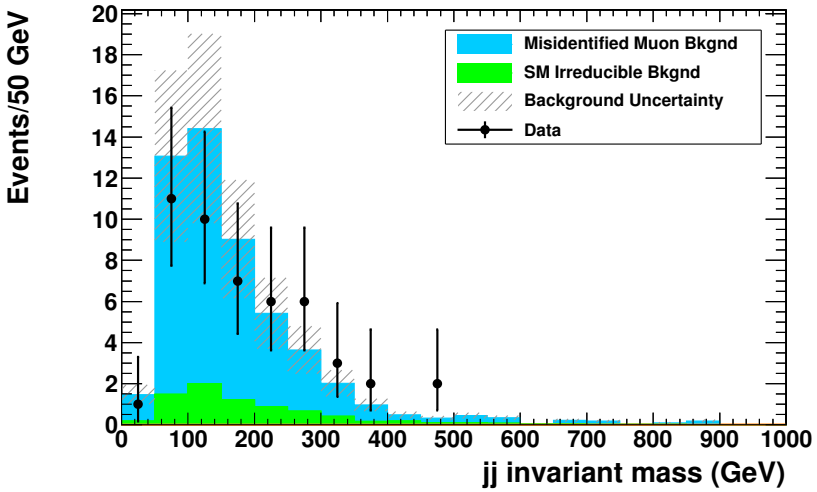


Figure 6.20: Closure test with data in the region defined by $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, the standard requirements are applied in the FR calculation (i.e. $p_T > 40$ GeV cut on the away jet and $RelIso < 0.8$ for the loose object) and the tight selection in signal selection (p_T(jet) > 30GeV). The invariant mass of the two selected jets is shown. The black dots are points with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentification (fake) muon background prediction.

6.3 Charge misidentification background

While the previous background estimates are applied in both channels the charge misidentification (missID) is specific to the electron channel. The muon charge is well determined in CMS due to the strong magnetic field and the excellent tracking system. Problems are foreseen only for highest p_T muons but most of the muons we consider are less than ~ 100 GeV. MC studies confirmed that the muon charge missID probability is well below 10^{-5}, roughly two orders of magnitude lower than that for electrons.

The electron charge sign measurement is performed in the high multiplicity environment of the Tracker and electrons are more subject to scattering and radiation emission. Thus, even if we require consistency of all three charge measurement methods (see Sec. 5.2.4), such a background cannot be neglected.

We have used a similar method to the one already used in Ref. [14], where the background is determined based on Monte Carlo simulation and a calculation of the charge misidentification rate from data. We perform the following four steps:

1. We calculate the electron charge misidentification probability from electron MC samples by matching the reco level electron to generator level truth electrons

2. We then test this misidentification probability using $Z \to e^+e^-$ control samples in real data. The $Z \to e^+e^-$ control region is defined by the requirements: (i) dielectron mass $76 < m_{ee} < 106$ GeV; (ii) $E_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV; (iii) transverse mass $m_T < 25$ GeV. We select both opposite sign and same sign events in the Z control region.

3. We apply the charge misidentification probability from Monte Carlo to the opposite

sign events in the Z control sample, and thereby obtain a prediction of the number of same sign events. We compare the number of predicted events to the number of observed same sign events. We check for consistency and use any discrepancy as the systematic uncertainty in the prediction.

4. We apply the charge misidentification probability to the opposite sign events in data that pass all the analysis event selection cuts (except for the same sign requirement) giving the final prediction of the number of background events from charge misidentification.

Figure 6.21 shows the charge misidentification rate from Monte Carlo. The average electron misidentification probability is found to be $(3.3 \pm 0.2) \times 10^{-4}$ in the ECAL barrel region ($|\eta| < 1.5$) and $(2.9 \pm 0.1) \times 10^{-3}$ in the ECAL endcap region ($1.5 < |\eta| < 2.5$) In the Z control sample we find 215,986 opposite sign events. After applying the charge misidentification probability to these opposite sign events, we obtain a prediction of 120.9 ± 10.3 same sign events, while the observed number of same sign events is 109. The agreement is good within the statistical uncertainty, and shows that the misidentification rate from Monte Carlo is consistent with that measured in data. We use the difference in the predicted and observed numbers, including uncertainties, to set a systematic uncertainty of 25% on the charge misidentification background.

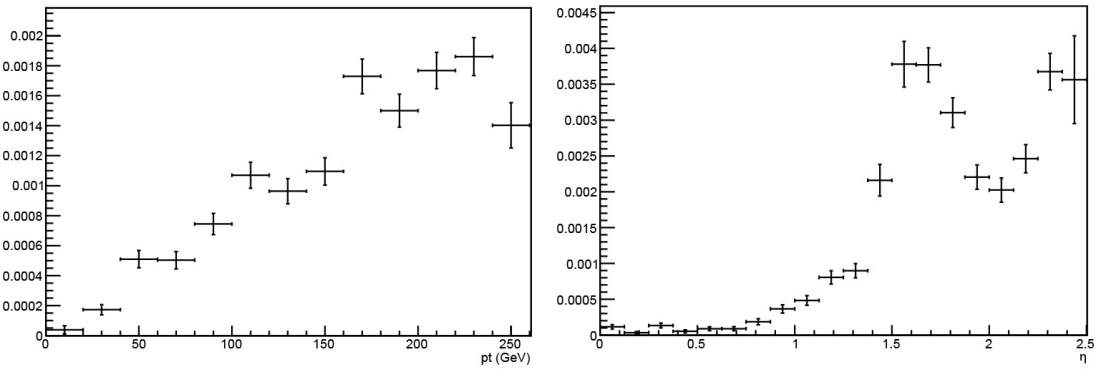


Figure 6.21: Electron charge misidentification probability obtained from Monte Carlo as a function of p_T (left) and η (right).

Chapter 7

Systematic Uncertainties

As this is a counting experiment the dominant sources of systematic uncertainties are those associated with our understanding of the signal model, efficiencies, and the background prediction. Most of them are common among many other CMS searches and we briefly describe them in this chapter. Particular attention will be devoted to the systematic uncertainty on the fake background prediction since this is specific to this analysis and is also the dominant systematic uncertainty.

7.1 General systematic uncertainties

The systematic uncertainties are listed and described below:

- *Integrated luminosity*: The systematic uncertainty on the integrated luminosity is 2.2%.

- *Parton distribution functions*: The ALPGEN signal Monte Carlo parton distribution function uncertainty is estimated using the method in Ref. [21]. The resulting uncertainty on the signal yield is 6%.

- Q^2 *scale:* We vary the event Q^2 scale in the ALPGEN signal Monte Carlo generator from $4Q^2$ to $Q^2/4$. The resulting uncertainty on the signal yield is 1%.

- *Muon trigger and selection*: Based upon the results of our tag and probe study and the trigger studies in Ref. [45] we assign a 2% uncertainty on the trigger effeciency and 2% on the muon selection effeciency.

- *Electron trigger and selection*: We assign an uncertainty on the electron selection efficiency as a function of Majorana mass, as listed in Table 5.17. Additional uncertainties of 4% due to the electron trigger efficiency and 10% due to the electron selection efficiency, as studied in Ref. [14], are also accounted for.

- *Jet energy scale*: We scale the jet energy by the official uncertainty [17] and observe the resulting effect on the signal efficiency. The resulting systematic uncertainty is between $3.3 - 14.2\%$ and depends on the overall importance of the jet p_T cut. In the lower Majorana neutrino mass range it has a larger effect (up to 14.2% at $m_N = 50$ GeV), while at 210 GeV is decreases to 3.3%.

- *Jet energy resolution*: We scale the jet energy uncertainty by an additional 10% and again observe the additional effect this has on the signal efficiency. We see an additional systematic of $0.2\% - 1\%$, depending on Majorana neutrino mass.

- *Pileup model*: As recommended by Ref. [3] we smear the number of interactions by ± 0.6 for the purpose of calculating the uncertainty due to pile up modeling in our simulations. This results in a systematic uncertainty of 1%.

- *Data-driven background estimate*: The uncertainty on the fake background estimate is 35% and is discussed in Sec. 7.2, while for the charge misidentification is 25% as discussed in Sec. 6.3.

- *MC background estimate*: The uncertainty on the MC background estimates is given by the uncertainty on the cross-section propagated to the final event yield. NLO and NNLO theoretical cross-sections are used and the relative uncertainty is then propagated to obtain the systematic uncertainty.

The uncertainties in the muon transverse momentum energy scale were found to be between $1.3\% - 6\%$ depending on muon pseudorapidity [44]. The effect of this error on our overall signal effeciency is negligible. A summary of the systematic uncertainties is given in Table 7.1.

Table 7.1: Summary of systematic uncertainties.

	Majorana Neutrino Mass (GeV)		
Source	70	130	190
Mass-dependent uncertainties:			
Jet Energy Scale	12%	7%	4%
Jet Energy Resolution	0.68%	0.23%	0.23%
Pile Up Model	1.7%	0.48%	0.16%
Mass-independent uncertainties:			
Integrated Luminosity	2.2%	2.2%	2.2%
Parton distribution Functions	5%	5%	5%
Event Q^2 scale	1%	1%	1%
Muon Trigger	2%	2%	2%
Muon Selection	2%	2%	2%
Background Prediction	35%	35%	35%

7.2 Systematic uncertainty on fake muon background

As already mentioned in Sec. 6.2, we have performed detailed studies in order to assign the proper systematic uncertainty on the fake background prediction, since this a big contributor to the final sensitivity and exclusion power.

We have followed three different paths to have the best possible understanding of the background prediction:

1. MC closure test,

2. data closure test with $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV,

3. variation of the free parameter in the fake rates calculation (away jet p_T) and in the loose object definition ($RelIso$).

We have decided to use all the three methods to have an accurate measurement of possible systematics, cross-check discrepancies, and better understand their sources.

As already seen in Table 6.6 the MC closure tests give, in general, good agreements, with the only exception being the $t\bar{t}$ in which case the discrepancy is about 50%. The discrepancy has already been explained and it is due to the different kinematics of the QCD events used to calculate the FR compared to $t\bar{t}$ events. We want to point out here that the disagreement, even if large, affects only a small part of our total background. We have estimated it from MC to be less than 10% of the total fake muon background.

Secondly we have shown in Table 6.8 that the predicted and observed number of events in data agree rather well, with a discrepancy of about 10%.This result along with the MC closure tests would already suggest a systematic error of the order of 30% or less.

Eventually, as a final check, we have tested the stability of the prediction against variation of the definition of the loose object and the away jet p_T cut. Thus, we have computed different fake rate matrices starting with an away jet as low as $p_T > 20$ GeV (the cut used in the loose event selection) and increasing it up to to 60 GeV in steps of 10 GeV. The 20 GeV cut of course represents an extreme case, since in our final selection we have at least two jets, so a more energetic event, but we wanted to test the method to the limit. The other change we made was in the definition of the

loose object. We started with $RelIso < 0.4$ (0.3 is too close to the nominal selection at 0.1 and shows divergences) and relaxed this parameter up to 1.0 in steps of 0.1. The

$p_T \setminus RelIso$	0.4	0.5	0.6	0.7	0.8	0.9	1.0
20 GeV	−28%				−25%		−33%
30 GeV					−9.6%		
40 GeV	−8.0%	−4.6%	−3.3%	0.96%	0.00%	−2.8%	−3.0%
50 GeV					2.6%		
60 GeV	−2.8%				3.1%		−0.14%

Table 7.2: Percentage variation of the final prediction versus loose object definition ($RelIso$) and away jet p_T cut. The bin 40 GeV and $RelIso = 0.8$ represents the chosen values.

results of this study are shown in Table 7.2, not all the cells in the table are filled since they are not relevant for our purposes.

It is interesting to notice an effect that could be surprising from our experience with the $t\bar{t}$ closure test: the increase of the final prediction with higher p_T cut. This is exactly the opposite effect that we saw in the MC closure test with $t\bar{t}$, where the prediction actually decreases when increasing the away jet momentum requirement. To understand it we have to look back at Eq. 6.5 and in particular at the double-to-single correction: for events with large QCD (an therefore double fake) background, where the correction is relatively large, the reduction of the FR actually increases the total prediction.

It is easier to illustrate the behavior of the prediction using a similar technique and assumptions used to derive Eq. 6.7. In this case we chose our variable x to be the FR and the quantities $N_{n\bar{n}}$ and $N_{\bar{n}\bar{n}}$ are fixed. Assuming an average FR over p_T and η we can rewrite Eq. 6.5 as:

$$N_{pp} = \frac{x}{1-x} \times \left(N_{n\bar{n}} - \frac{x}{1-x} N_{\bar{n}\bar{n}} \right) \tag{7.1}$$

It is now straightforward to plot this function vs. FR using different values for $N_{n\bar{n}}$ and $N_{\bar{n}\bar{n}}$. For events with overwhelming single fake we simply use $N_{n\bar{n}} = 10$ and $N_{\bar{n}\bar{n}} = 1$ (see Fig.7.1). Then we increase $N_{\bar{n}\bar{n}}$ to 5 and 10 for events with more and more double fake, this is done in Fig. 7.2 and 7.3 respectively. Even if it is difficult to make an exact correspondence between these simplified plots and the actual fake rate prediction, they are helpful to understand the general trend of the total prediction in different conditions.

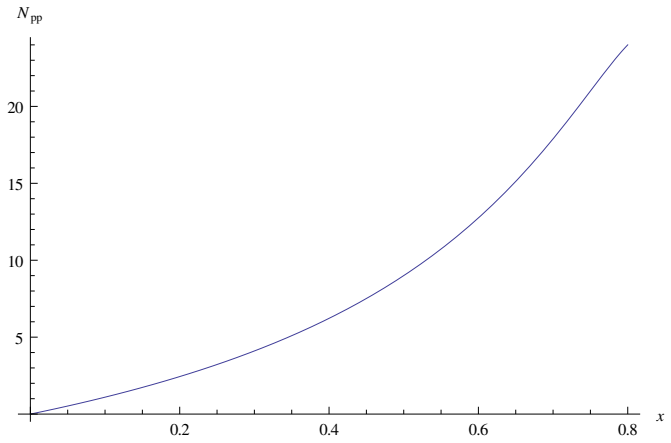


Figure 7.1: Simple plot of Eq. 7.1 with $N_{n\bar{n}} = 10$ and $N_{\bar{n}\bar{n}} = 1$. The x-axis represents roughly the FR, as x increases the total prediction follows it.

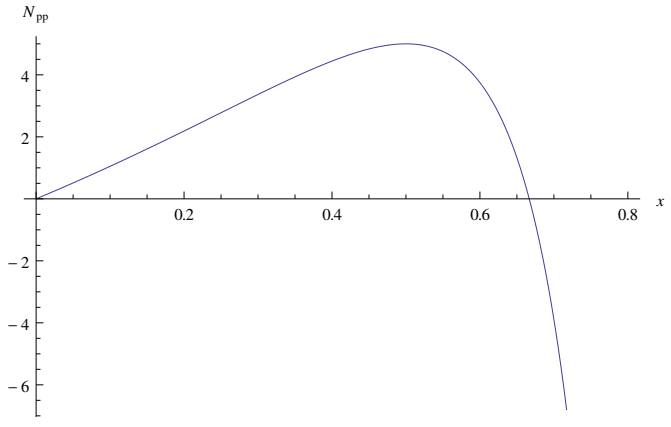


Figure 7.2: Simple plot of Eq. 7.1 with $N_{n\bar{n}} = 10$ and $N_{\bar{n}\bar{n}} = 5$. With a higher number of events with two loose objects the total prediction reaches a turn point and decreases even if the FR increases.

From Fig. 7.1, 7.2, and 7.3 we can see that the total prediction is not a mono-

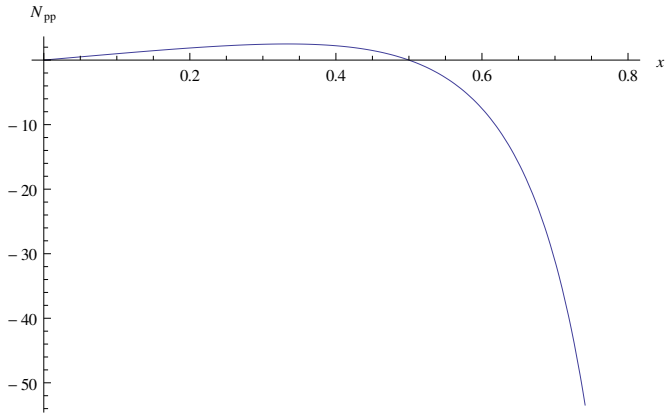


Figure 7.3: Simple plot of Eq. 7.1 with $N_{n\bar{n}} = 10$ and $N_{\bar{n}\bar{n}} = 10$. When the number of events with two loose objects matches the number of events with one loose and one tight the decrease of the total prediction happens sooner and is sharper.

tonic function of the FR for all possible values of $N_{n\bar{n}}$ and $N_{\bar{n}\bar{n}}$, we have confirmed that the double-to-single correction plays an important role when $N_{\bar{n}\bar{n}}$ grows and it is actually more important than the FR itself, therefore, the behavior seen in Table. 7.2 is expected.

This variational technique verifies our first estimate and confirms the important role played by the away jet in the event selection for the fake rates. It is worth mentioning also that the bins in Table 7.2 are not uncorrelated, in other words the loose object definition influences the effect that the away jet cut has on the final prediction and vice versa. The correlation is not 100% but it cannot be ignored.

The final systematic uncertainty on the fake background prediction has been chosen to be 35%. This value covers all the tests we have performed except the $t\bar{t}$ that, anyway, counts only for a small fraction of the events.

Chapter 8

Results and Conclusions

In this chapter we summarize the results obtained in the muon and electron channel for both tight and loose selections.

8.1 Muon channel

8.1.1 Loose selection

After all the selection cuts listed in Section 5.2 we observe 130 events in data for the loose selection. The final background estimates are given in Table 8.1 along with statistical and systematic uncertainties. The total background estimate is 162.7 ± 6.1 (stat.) ± 51.9 (syst.) events, with the dominant contribution arising from the fake muon background (148.2 ± 6.1 (stat.) ± 51.9 (syst.) events). The data yield is in good agreement with this estimate within errors, therefore we conclude we do not observe any excess and we can proceed setting limits.

We have also checked different distributions as we did for the data control region with $MET > 50$ GeV. Figure 8.1 shows the distributions of invariant mass of

Source	Events
Monte Carlo:	
WZ	7.36 ± 0.45 (stat.) ± 0.29 (syst.)
ZZ	1.77 ± 0.13 (stat.) ± 0.07 (syst.)
$W\gamma$	2.12 ± 0.50 (stat.) ± 0.21 (syst.)
$t\bar{t}W$	1.12 ± 0.06 (stat.) ± 0.11 (syst.)
W^+W^+	1.35 ± 0.09 (stat.) ± 0.13 (syst.)
W^-W^-	0.55 ± 0.03 (stat.) ± 0.06 (syst.)
dp $W^\pm W^\pm$	0.24 ± 0.03 (stat.) ± 0.02 (syst.)
Total Monte Carlo	14.5 ± 0.69 (stat.) ± 0.41 (syst.)
Data-driven background estimate:	
Fake muon background	148.22 ± 6.10 (stat.) ± 51.88 (syst.)
Total background	162.72 ± 6.14 (stat.) ± 51.88 (syst.)
Observed in data (4.98 fb^{-1})	130

Table 8.1: Observed event yields and estimated backgrounds with statistical and systematic uncertainties for the muon channel analysis (4.98 fb^{-1}) for the loose selection.

the second highest p_T muon and the two leading jets in the event for data, standard model backgrounds, and three choices for the Majorana neutrino signal: $m_N = 80$ GeV, $|V_{\mu N}|^2 = 0.025$, $m_N = 130$ GeV, $|V_{\mu N}|^2 = 0.025$, and $m_N = 210$ GeV, $|V_{\mu N}|^2 = 0.25$. We show also the leading muon p_T in Fig. 8.2, the invariant mass of the four selected particles in Fig. 8.1, and the invariant mass of the two selected jets in Fig. 8.4. All the distributions, even if not used to compute the final limit, are in good agreement with expectations.

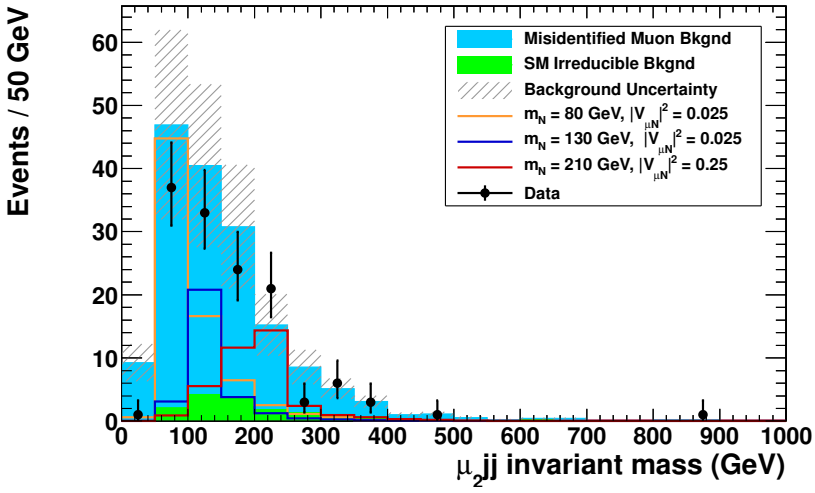


Figure 8.1: Invariant mass of the second leading p_T muon and the two leading jets for events passing the loose selection. The plot shows the data, standard model backgrounds, and three choices for the Majorana neutrino signal: $m_N = 80$ GeV, $|V_{\mu N}|^2 = 0.025$, $m_N = 130$ GeV, $|V_{\mu N}|^2 = 0.025$, and $m_N = 210$ GeV, $|V_{\mu N}|^2 = 0.25$.The black dots are data with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentified (fake) muon background prediction.

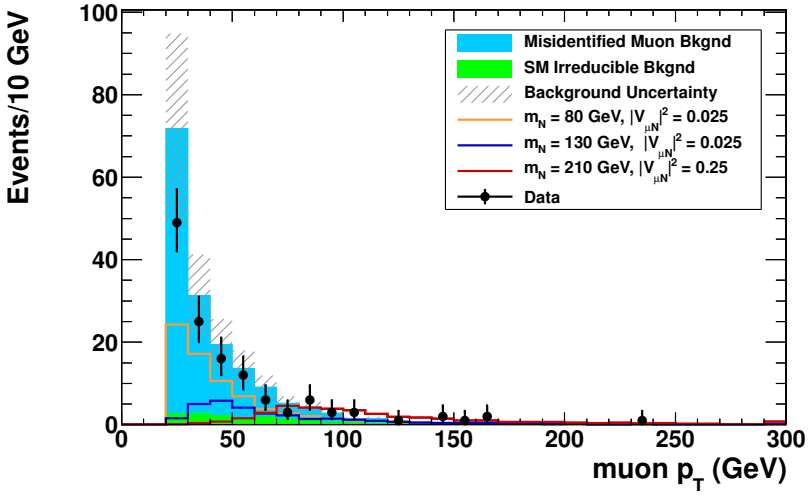


Figure 8.2: Leading μ momentum events passing the loose selection. The plot shows the data, standard model backgrounds, and three choices for the Majorana neutrino signal: $m_N = 80$ GeV, $|V_{\mu N}|^2 = 0.025$, $m_N = 130$ GeV, $|V_{\mu N}|^2 = 0.025$, and $m_N = 210$ GeV, $|V_{\mu N}|^2 = 0.25$.The black dots are data with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentified (fake) muon background prediction.

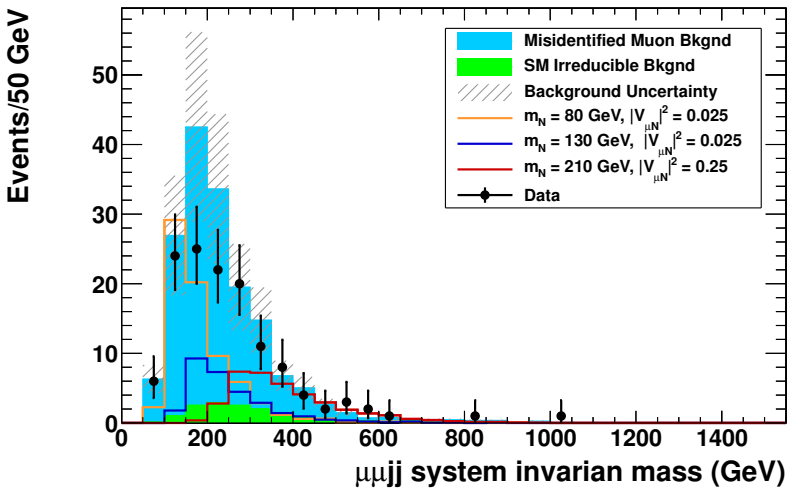


Figure 8.3: Invariant mass of the four selected particles in the events passing the loose selection. The plot shows the data, standard model backgrounds, and three choices for the Majorana neutrino signal: $m_N = 80$ GeV, $|V_{\mu N}|^2 = 0.025$, $m_N = 130$ GeV, $|V_{\mu N}|^2 = 0.025$, and $m_N = 210$ GeV, $|V_{\mu N}|^2 = 0.25$.The black dots are data with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentified (fake) muon background prediction.

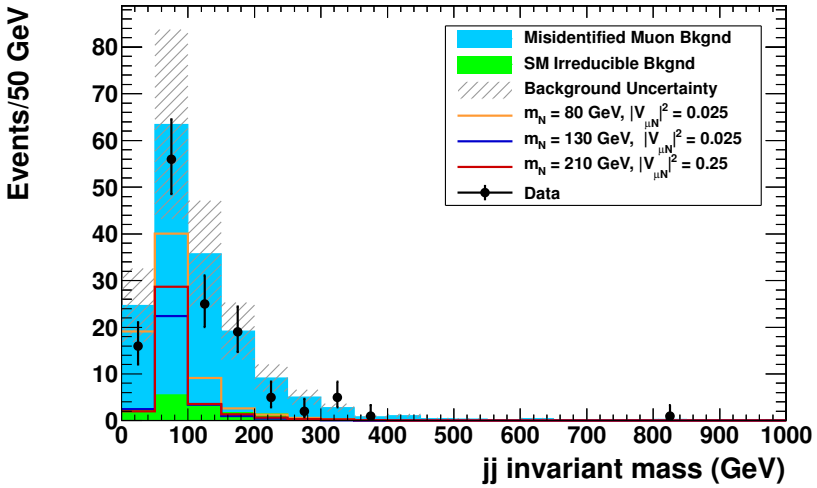


Figure 8.4: Invariant mass of the two selected jets in the events passing the loose selection. The plot shows the data, standard model backgrounds, and three choices for the Majorana neutrino signal: $m_N = 80$ GeV, $|V_{\mu N}|^2 = 0.025$, $m_N = 130$ GeV, $|V_{\mu N}|^2 = 0.025$, and $m_N = 210$ GeV, $|V_{\mu N}|^2 = 0.25$.The black dots are data with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentified (fake) muon background prediction.

8.1.2 Tight selection

To increase the sensitivity for high neutrino mass signals, we used a tight selection with the jet $p_T > 30$ GeV. For this selection we observe 65 events in data. The final background estimates are given in Table 8.2. The total background estimate is 70.4 ± 4.2 (stat.) ± 22.1 (syst.) events, again the dominant contribution is due to the fake muon background (63.1 ± 4.2 (stat.) ± 22.1 (syst.) events). The data yield is, again, in good agreement with the estimate within errors.

Source	Events
Monte Carlo:	
WZ	3.22 ± 0.30 (stat.) ± 0.13 (syst.)
ZZ	1.02 ± 0.10 (stat.) ± 0.04 (syst.)
$W\gamma$	0.75 ± 0.27 (stat.) ± 0.07 (syst.)
$t\bar{t}W$	1.06 ± 0.05 (stat.) ± 0.53 (syst.)
W^+W^+	0.76 ± 0.06 (stat.) ± 0.38 (syst.)
W^-W^-	0.45 ± 0.03 (stat.) ± 0.23 (syst.)
dp $W^\pm W^\pm$	0.07 ± 0.02 (stat.) ± 0.04 (syst.)
Total Monte Carlo	7.33 ± 0.42 (stat.) ± 0.71 (syst.)
Data-driven background estimate:	
Fake muon background	63.09 ± 4.20 (stat.) ± 22.08 (syst.)
Total background	70.42 ± 4.22 (stat.) ± 22.09 (syst.)
Observed in data (4.98 fb^{-1})	65

Table 8.2: Observed event yields and estimated backgrounds with statistical and systematic uncertainties for the muon channel analysis (4.98 fb^{-1}) for the tight selection.

As for the loose selection we did no observe any excess. Before proceeding to the limit calculation we checked the distributions of the neutrino mass (Fig. 8.5), the leading muon p_T (Fig. 8.6), the invariant mass of the four particles in the selected events (Fig. 8.7), and the invariant mass of the two jets (Fig. 8.8).

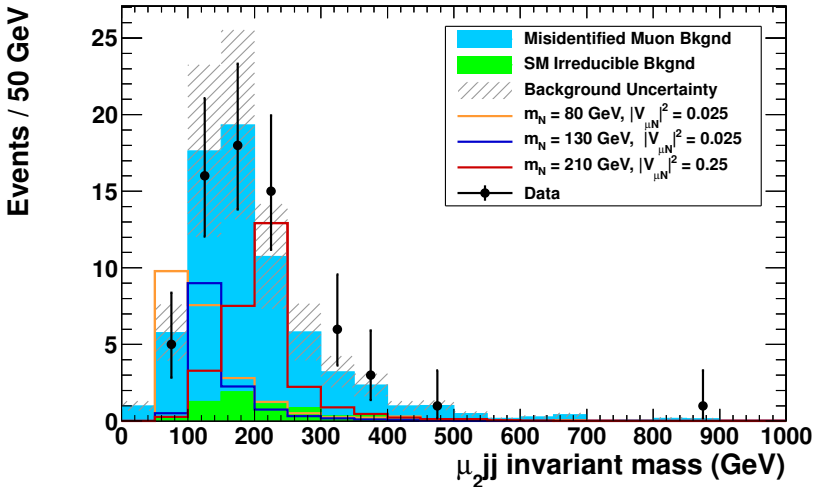


Figure 8.5: Invariant mass of the second leading p_T muon and the two leading jets for events passing the tight selection. The plot shows the data, standard model backgrounds, and three choices for the Majorana neutrino signal: $m_N = 80$ GeV, $|V_{\mu N}|^2 = 0.025$, $m_N = 130$ GeV, $|V_{\mu N}|^2 = 0.025$, and $m_N = 210$ GeV, $|V_{\mu N}|^2 = 0.25$.The black dots are data with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentified (fake) muon background prediction.

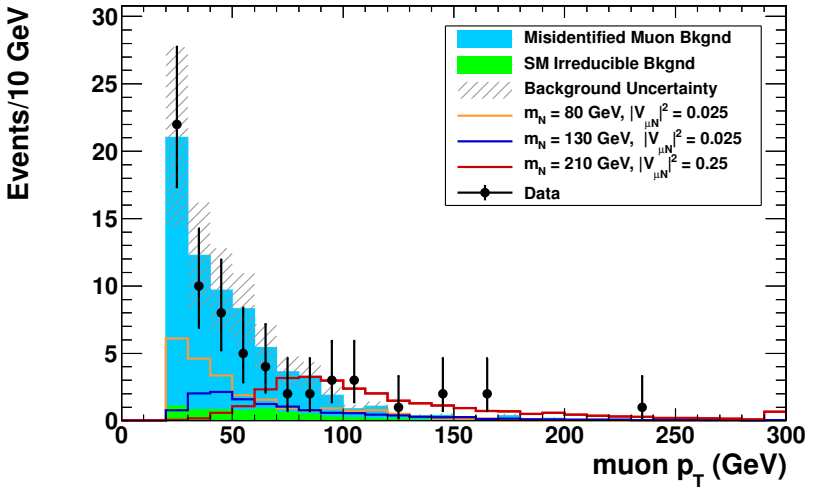


Figure 8.6: Leading μ momentum events passing the tight selection. The plot shows the data, standard model backgrounds, and three choices for the Majorana neutrino signal: $m_N = 80$ GeV, $|V_{\mu N}|^2 = 0.025$, $m_N = 130$ GeV, $|V_{\mu N}|^2 = 0.025$, and $m_N = 210$ GeV, $|V_{\mu N}|^2 = 0.25$.The black dots are data with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentified (fake) muon background prediction.

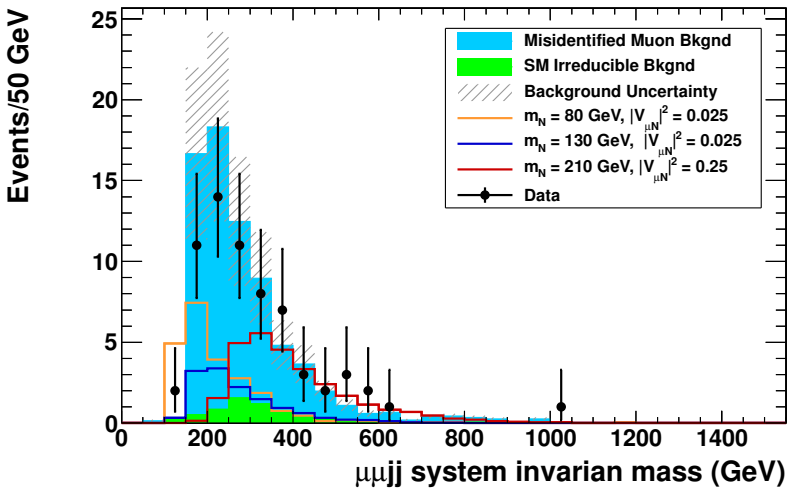


Figure 8.7: Invariant mass of the four selected particles in the events passing the tight selection. The plot shows the data, standard model backgrounds, and three choices for the Majorana neutrino signal: $m_N = 80$ GeV, $|V_{\mu N}|^2 = 0.025$, $m_N = 130$ GeV, $|V_{\mu N}|^2 = 0.025$, and $m_N = 210$ GeV, $|V_{\mu N}|^2 = 0.25$.The black dots are data with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentified (fake) muon background prediction.

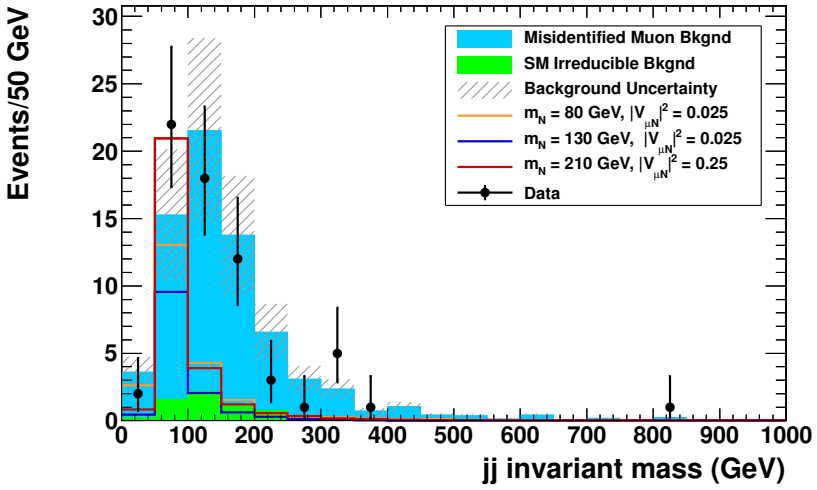


Figure 8.8: Invariant mass of the two selected jets in the events passing the tight selection. The plot shows the data, standard model backgrounds, and three choices for the Majorana neutrino signal: $m_N = 80$ GeV, $|V_{\mu N}|^2 = 0.025$, $m_N = 130$ GeV, $|V_{\mu N}|^2 = 0.025$, and $m_N = 210$ GeV, $|V_{\mu N}|^2 = 0.25$.The black dots are data with statistical error only, the green histogram is the sum of all the SM same-sign background sources and the blue histogram is the misidentified (fake) muon background prediction.

8.2 Electron channel

8.2.1 Loose selection

As for the muons we have defined two different selections. The results for the loose selection are summarized in Table 8.3. We did not observe any excess, and as before we also checked the invariant mass distribution of the second leading p_T electron and the two leading jets in the event (Fig. 8.9).

Source	Events
Monte Carlo:	
WZ	10.31 ± 0.46 (stat.) ± 0.41 (syst.)
ZZ	3.36 ± 0.15 (stat.) ± 0.13 (syst.)
$W\gamma$	6.77 ± 0.85 (stat.) ± 0.68 (syst.)
$t\bar{t}W$	0.70 ± 0.04 (stat.) ± 0.35 (syst.)
W^+W^+	0.88 ± 0.07 (stat.) ± 0.44 (syst.)
W^-W^-	0.34 ± 0.02 (stat.) ± 0.17 (syst.)
dp $W^\pm W^\pm$	0.17 ± 0.03 (stat.) ± 0.09 (syst.)
Total Monte Carlo	22.53 ± 0.98 (stat.) ± 1.00 (syst.)
Charge misidentification:	103.36 ± 8.97 (stat.) ± 25.85 (syst.)
Data driven background estimate:	
Fake electron background	610.50 ± 52.20 (stat.) ± 213.68 (syst.)
Total Background	736.39 ± 52.2 (stat.) ± 213.68 (syst.)
Observed in data (4.98 fb^{-1})	616

Table 8.3: Observed event yields and estimated backgrounds with statistical and systematic uncertainties for the electron channel analysis for the loose selection.

8.2.2 Tight selection

To increase the sensitivity to possible signals at high mass we have used a tight selection also for the electron channel, with the same 30 GeV cut on the jets p_T. The results are listed in Table 8.4 and we can see the distribution of the invariant mass of a possible Majorana neutrino candidate in Fig. 8.10. As before we do not see any excess

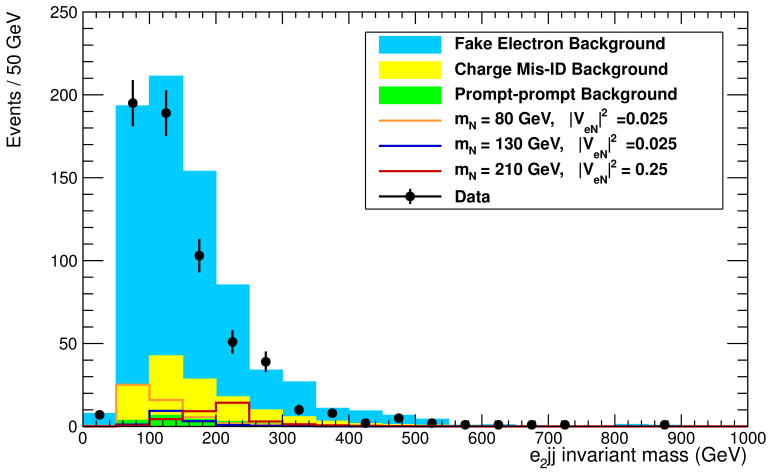


Figure 8.9: Invariant mass of the second leading p_T electron and the two leading jets for events passing the loose selection. The plot shows the data, standard model backgrounds, and three choices for the Majorana neutrino signal: $m_N = 80$ GeV, $|V_{eN}|^2 = 0.025$, $m_N = 130$ GeV, $|V_{eN}|^2 = 0.025$, and $m_N = 210$ GeV, $|V_{eN}|^2 = 0.25$.

Source	Events
Monte Carlo:	
WZ	4.89 ± 0.32 (stat.) ± 0.20 (syst.)
ZZ	2.12 ± 0.12 (stat.) ± 0.08 (syst.)
$W\gamma$	1.72 ± 0.43 (stat.) ± 0.17 (syst.)
$t\bar{t}W$	0.62 ± 0.04 (stat.) ± 0.31 (syst.)
W^+W^+	0.73 ± 0.07 (stat.) ± 0.37 (syst.)
W^-W^-	0.27 ± 0.02 (stat.) ± 0.13 (syst.)
dp $W^\pm W^\pm$	0.19 ± 0.03 (stat.) ± 0.10 (syst.)
Total Monte Carlo	10.55 ± 0.56 (stat.) ± 0.58 (syst.)
Charge misidentification:	31.90 ± 2.70 (stat.) ± 7.98 (syst.)
Data driven background estimate:	
Fake electron background	176.80 ± 4.70 (stat.) ± 61.88 (syst.)
Total Background	219.25 ± 5.45 (stat.) ± 62.39 (syst.)
Observed in data (4.98 fb^{-1})	201

Table 8.4: Observed event yields and estimated backgrounds with statistical and systematic uncertainties for the electron channel analysis for the tight selection.

so we can set limits.

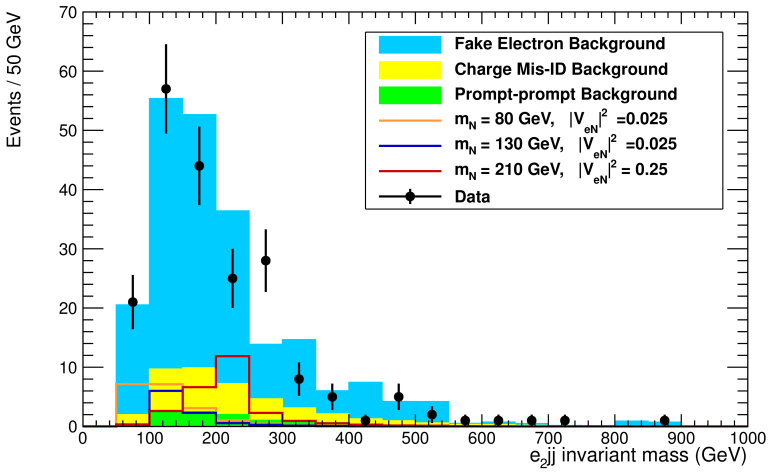


Figure 8.10: Invariant mass of the second leading p_T electron and the two leading jets for events passing the tight selection. The plot shows the data, standard model backgrounds, and two choices for the Majorana neutrino signal: $m_N = 80$ GeV, $|V_{eN}|^2 = 0.025$, $m_N = 130$ GeV, $|V_{eN}|^2 = 0.025$, and $m_N = 210$ GeV, $|V_{eN}|^2 = 0.25$.

8.3 Interpretation of results

We do not see any significant excess in the data beyond our predicted background. Therefore, we set limits on the square of the Majorana neutrino mixing element based on the number of observed events, the predicted number of background events, and the predicted number of events from Majorana neutrino production. The 95% confidence level limit on the cross section for heavy Majorana production is obtained from the standard CMS RooStatsCl95 package [83, 1] using the CLs method. From this the 95% C.L. limit on the mixing element squared can be calculated using the "bare" cross section (defined as the cross section for a mixing of 1.0) obtained from the Monte Carlo simulation.

For Majorana neutrino masses above 90 GeV we find that the expected limits are better for the tight event selection than for the loose selection. This is because the higher jet p_T cut of 30 GeV suppresses the background more effectively, while retaining

good signal efficiency at masses above 90 GeV. For lower masses the signal efficiency is significantly reduced, so the expected limit is not as good. Therefore, since we are significantly less sensitive than the LEP limits at masses below 90 GeV, we use the tight event selection to set our final limits. The resulting limits on $|V_{\mu N}|^2$ and $|V_{eN}|^2$, as a function of m_N, are shown in Figure 8.11 and Figure 8.12 respectively.

We also calculate the 95% C.L. upper limit on the cross section times acceptance times efficiency ($\sigma A\epsilon$). The observed limit for the muon channel is $(\sigma A\epsilon)_{95} = 5.39$ fb and the expected limit is $(\sigma A\epsilon)_{95} = 5.26$ fb, for the electron channel the observed limit is $(\sigma A\epsilon)_{95} = 14.56$ fb and the expected limit is $(\sigma A\epsilon)_{95} = 14.01$ fb.

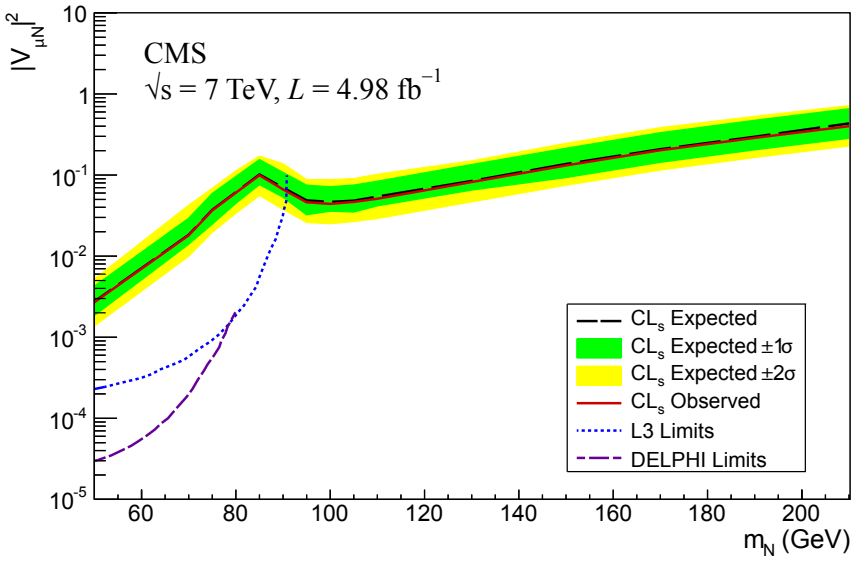


Figure 8.11: Exclusion region in the Majorana neutrino mixing element squared vs. mass plane obtained using the CLs method. The black line is the expected limit (with one and two sigma bands in green and yellow respectively). The red line is the observed exclusion contour. Also shown are the limits from DELPHI [7] and L3 [8].

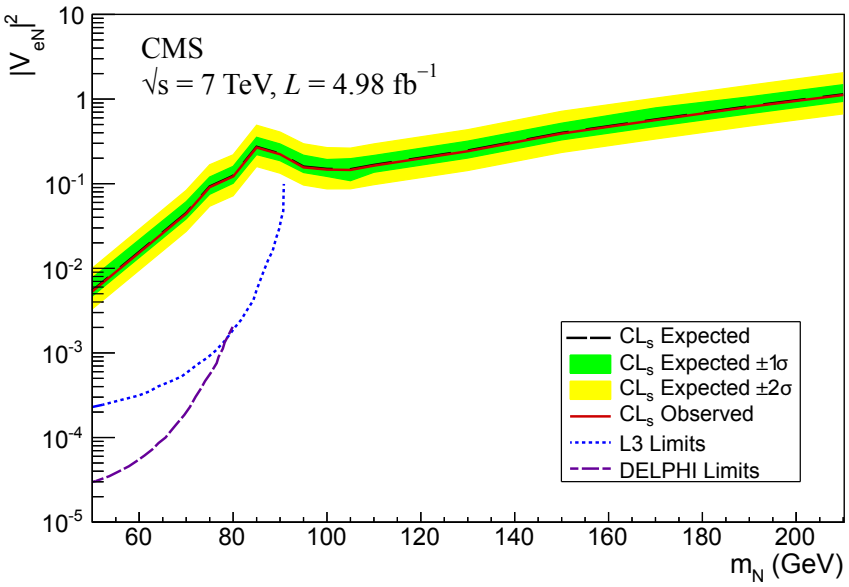


Figure 8.12: Exclusion region in the Majorana neutrino mixing element squared vs. mass plane obtained using the CLs method. The black line is the expected limit (with one and two sigma bands in green and yellow respectively). The red line is the observed exclusion contour. Also shown are the limits from DELPHI [7] and L3 [8].

Bibliography

[1] *CLs method Guide.*
https://twiki.cern.ch/twiki/bin/viewauth/CMS/RooStatsCl95.

[2] *CMSSW Fast Simulation Offline Guide.*
https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideFastSimulation.

[3] *CMSSW pileup studies Guide.*
https://twiki.cern.ch/twiki/bin/view/CMS/PileupSystematicErrors.

[4] *ROOT Guide.*
http://root.cern.ch.

[5] G. Aad and et al. Inclusive search for same-sign dilepton signatures in pp collisions at $\sqrt{s} = 7$ TeV with the atlas detector. *Journal of High Energy Physics*, 2011:1–48, 2011.

[6] G. Aad and et al. Search for heavy neutrinos and right-handed w bosons in events with two leptons and jets in pp collisions at $\sqrt{s} = 7$ TeV with the atlas detector. *The European Physical Journal C*, 72:1–22, 2012.

[7] P. Abreu et al. Search for neutral heavy leptons produced in Z decays. *Z.Phys.*, C74:57–71, 1997.

[8] O. Adriani et al. Search for isosinglet neutral heavy leptons in Z^0 decays. *Phys.Lett.*, B295:371–382, 1992.

[9] Jr. Almeida, F.M.L., Yara Do Amaral Coutinho, Jose Antonio Martins Simoes, and M.A.B. do Vale. On a signature for heavy Majorana neutrinos in hadronic collisions. *Phys.Rev.*, D62:075004, 2000. Latex2e(epsfig), 12 pages, 8 figures, to appear Physical Review D.

[10] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nucl. Phys. B*, (126):298, 1977.

[11] Johan Alwall, Pavel Demin, Simon de Visscher, Rikkert Frederix, Michel Herquet, Fabio Maltoni, Tilman Plehn, David L Rainwater, and Tim Stelzer. Madgraph/madevent v4: The new web generation. *J. High Energy Phys.*, 09:028. 38 p, Jun 2007.

[12] Anupama Atre, Tao Han, Silvia Pascoli, and Bin Zhang. The Search for Heavy Majorana Neutrinos. *JHEP*, 0905:030, 2009.

[13] D. Barge, C. Campagnari, P. Kalavase, D. Kovalskyi, V. Krutelyov, J. Ribnik W. Andrews, D. Evans, F. Golf, J. Mulmenstadt, S. Padhi, Y. Tu, F. Wurthwein, A. Yagil, J. Yoo, L. Bauerdick, I. Bloch, K. Burkett, I. Fisk, Y. Gao, O. Gutsche, and B. Hooberman. Inclusive search for new physics with same-sign dileptons using early lhc data. *CMS Analysis Note*, AN-2010-247, 2010.

[14] D. Barge, C. Campagnari, P. Kalavase, D. Kovalskyi, V. Krutelyov, J. Ribnik, W. Andrews, G. Cerati, D. Evans, F. Golf, I. MacNeill, S. Padhi, Y. Tu, F. Wurthwein, A. Yagil, J Yoo, L. Bauerdick, I. Bloch, K. Burkett, I. Fisk, Y. Gao, O. Gutsche, B. Hooberman, S. Jindariani, and J. Linacre. Search for new physics with same-sign dileptons using the 2011 dataset of cms. *CMS Analysis Note*, AN-2011-258, 2011.

[15] G. Bauer, J. Bendavid, E. Butz, et al. Lepton effiencies for the inclusive w cross section measurement with 36.1 pb^{-1}. *CMS Analysis Note*, AN-2011-097, 2011.

[16] P. Beneš, Amand Faessler, S. Kovalenko, and F. Šimkovic. Sterile neutrinos in neutrinoless double beta decay. *Phys. Rev. D*, 71:077901, Apr 2005.

[17] J. Berger, D. del Re, R. Eusebi, T. Hauth, I. Iashvili, S. Jain, M. Kaya, O. Kaya, K. Mishra, F. Pandolfi, D. Piparo, G. Quast, P. Schieferdecker, S. Sengupta, and M. Voutilainen. Absolute jet energy correction and its uncertainty with $36\,\mathrm{pb}^{-1}$. *CMS Analysis Note*, AN-2011/105, 2011.

[18] J. Beringer and et al. Review of particle physics. *Phys. Rev. D*, 86:010001, Jul 2012.

[19] Samoil Bilenky. *Introduction to the Physics of massive and Mixed Neutrinos.* Springer, 2010.

[20] R.K. Bock. *Data Analysis Techniques For High-Energy Physics Experiments.* Cambridge University Press, 1990.

[21] D Bourilkov, R C Group, and M R Whalley. LHAPDF: PDF use from the Tevatron to the LHC. 2006.

[22] S. Braibant, G. Giacomelli, and M. Spurio. *Particles and fundamental Interactions.* Springer, 2009.

[23] Nicola Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, Jun 1963.

[24] J. M. Campbell, J. W. Huston, and W. J. Stirling. Hard interactions of quarks and gluons: a primer for lhc physics. *Rep. Prog. Phys*, (70):89, Nov 2006. arXiv: hep-ph/0611148.

[25] S. Catani, F. Krauss, R. Kuhn, and B.R. Webber. Qcd matrix elements + parton showers. *JHEP*, (0111):063, 2001. arXiv: hep-ph/0109231.

[26] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012.

[27] GEANT4 Collaboration. Geant4: A simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.

[28] The CMS collaboration. *Technical Proposal*. Number CERN/LHCC 94-34. CERN, Geneva, 1994.

[29] The CMS collaboration. *The Electromagnetic Calorimeter: Technical Design Report*. Number CERN/LHCC 97-33. CERN, Geneva, 1997.

[30] The CMS collaboration. *The Hadronic Calorimeter: Technical Design Report*. Number CERN/LHCC 97-31. CERN, Geneva, 1997.

[31] The CMS collaboration. *The Magnet Project: Technical Design Report*. Number CERN/LHCC 97-10. CERN, Geneva, 1997.

[32] The CMS collaboration. *The Muon Project: Technical Design Report*. Number CERN/LHCC 97-32. CERN, Geneva, 1997.

[33] The CMS collaboration. *The Tracker Project: Technical Design Report*. Number CERN/LHCC 98-6. CERN, Geneva, 1998.

[34] The CMS collaboration. *The triggere and data acquisition project, Volume 1, The L-1 Trigger: Technical Design Report*. Number CERN/LHCC 2000-038, CMS TDR 6.1. CERN, Geneva, 2000.

[35] The CMS collaboration. *The triggere and data acquisition project, Volume 1, The High Level Trigger: Technical Design Report*. Number CERN/LHCC 2002-036, CMS TDR 6.2. CERN, Geneva, 2002.

[36] The CMS Collaboration. Reconstruction of electrons with the gaussian-sum filter in the cms tracker at the lhc. *CMS note*, (CMS NOTE-05-001), 2005.

[37] The CMS Collaboration. Algorithm for b-jet identification in cms. *CMS Physics Analysis Summary*, (BTV-09-001), 2009.

[38] The CMS Collaboration. Determination of the relative jet energy scale at cms from dijet balance. *CMS Physics Analysis Summary*, (JME-08-003), 2009.

[39] The CMS Collaboration. Jet energy calibration with photon+jet events. *CMS Physics Analysis Summary*, (JME-09-004), 2009.

[40] The CMS Collaboration. Met reconstruction, performance, and validation. *CMS Analysis Note*, (AN-08-089), 2009.

[41] The CMS Collaboration. Particle-flow based reconstruction of jets, taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$. *CMS Analysis Note*, (CMS AN-09-039), 2009.

[42] The CMS Collaboration. Commissioning of b-jet identification with pp collisions at $\sqrt{s} = 7$ TeV. *CMS Physics Analysis Summary*, (BTV-10-001), 2010.

[43] The CMS Collaboration. Determination of the jet energy scale in cms with pp collisions at $\sqrt{s} = 7$ TeV. *CMS Physics Analysis Summary*, (JME-10-010), 2010.

[44] The CMS Collaboration. Performace of CMS muon reconstruction in pp collisions at $\sqrt{s} = 7tev$. *CMS Physics Analysis Summary*, CMS-PAS-MOU-10-004, 2010.

[45] The CMS Collaboration. Search for higgs boson decay to two w bosons in the fully leptonic final state at $\sqrt{7}$ tev with 2011 data with the CMS detector. *CMS Analysis Note*, AN-11-155, 2011.

[46] A. Datta, M. Guchait, and A. Pilaftsis. Probing lepton number violation via majorana neutrinos at hadron supercolliders. *Phys.Rev.*, D50:3195–3203, 1994.

[47] F. del Aguila, J.A. Aguilar-Saavedra, J. de Blas, and M. Perez-Victoria. Electroweak constraints on see-saw messengers and their implications for LHC. 2008.

[48] F. del Aguila, J.A. Aguilar-Saavedra, and R. Pittau. Heavy neutrino signals at large hadron colliders. *JHEP*, 0710:047, 2007.

[49] Francisco del Aguila, Shaouly Bar-Shalom, Amarjit Soni, and Jose Wudka. Heavy majorana neutrinos in the effective lagrangian description: Application to hadron colliders. *Physics Letters B*, 670(4-5):399 – 402, 2009.

[50] Duane A. Dicus, Debra Dzialo Karatas, and Probir Roy. Lepton nonconservation at supercollider energies. *Phys. Rev. D*, 44:2033–2037, Oct 1991.

[51] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[52] C. Albajar et al. Events with large missing transverse energy at the cern collider: Iii. mass limits on supersymmetric particles. *Phys. Lett. B*, (198):261–270, 1987.

[53] C. Arnison et al. Hadronic jet production at the cern proton - anti-proton collider. *Phys. Lett. B*, (132):214–222, 1983.

[54] C. Bromberg et al. Observation of the production of jets of particles at high transverse momentum and comparison with inclusive single particle reactions. *Phys. Rev. Lett.*, (38):1447–1450, 1977.

[55] D. Boussard et al. *The Large Handron Collider conceptual design.* Number CERN/AC/95-05. CERN, Geneva, 1995.

[56] Dokshitzer et al. Better jet clustering algorithms. *JHEP*, (08):001, 1997.

[57] G. Abbiendi et al. Muon reconstruction in the cms detector. *CMS Analysis Note*, (097), 2008.

[58] S. Catani et al. Longitudinally invariant k_t clustering algorithms for hadron hadron collisions. *Nucl. Phys. B*, (406):187–224, 1993.

[59] Salam et al. A practical seedless infrared-safe cone jet algorithm. *JHEP*, (05):086, 2007.

[60] A. Ferrari, J. Collot, M-L. Andrieux, B. Belhorma, P. de Saintignon, J-Y. Hostachy, Ph. Martin, and M. Wielers. Sensitivity study for new gauge bosons and right-handed majorana neutrinos in pp collisions at $\sqrt{s} = 14$ TeV. *Phys. Rev. D*, 62:013001, May 2000.

[61] R. Fruhwirth. Application of kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth. A*, (262):444–450, 1987.

[62] F. Gasperini, M. Benettoni, E. Conti, and G. Zumerle. *Changes in the design of the CMS barrel drift chambers.* Number CMS NOTE 1999-007. CERN, Geneva, 1999.

[63] M. Gell-Mann, P. Ramond, and R. Slansky. In P.V. Nieuwenhuizen and D.Z. Freedman, editors, *Supergravity: proceedings of the Supergravity Workshop at Stony Brook*, page 315. North-Holland, 1979.

[64] D. Griffiths. *Introduction to elementary Particles.* Wiley, 2^{nd} edition, 2008.

[65] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.

[66] Tao Han and Bin Zhang. Signatures for Majorana neutrinos at hadron colliders. *Phys.Rev.Lett.*, 97:171804, 2006.

[67] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[68] Stefan Hoeche, Frank Krauss, Nils Lavesson, Leif Lonnblad, Michelangelo Mangano, Andreas Schaelicke, and Steffen Schumann. Matching parton showers and matrix elements. *HERA and the LHC workshop*, 2006. arXiv: hep-ph/0602031.

[69] K. Huitu, J. Maalampi, A. Pietil, and M. Raidal. Doubly charged higgs at lhc. *Nuclear Physics B*, 487(1-2):27 – 42, 1997.

[70] C. Itzykson and J.B. Zuber. *Quantum Field Theory.* McGraw-Hill, 2006.

[71] N. Jarosik and et al. Seven-Year Wilson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results. *Astrophysical Journal Supplement*, 192:14, 2011.

[72] Wai-Yee Keung and Goran Senjanović. Majorana neutrinos and the production of the right-handed charged gauge boson. *Phys. Rev. Lett.*, 50:1427–1430, May 1983.

[73] T. W. B. Kibble. Symmetry breaking in non-abelian gauge theories. *Phys. Rev.*, 155:1554–1561, Mar 1967.

[74] Makoto Kobayashi and Toshihide Maskawa. *cp*-violation in the renormalizable theory of weak interaction. *Progress of Theoretical Physics*, 49(2):652–657, 1973.

[75] H.L. Lai, J. Huston, S. Kuhlmann, J. Morfin, F. Olness, J.F. Owens, J. Pumplin, and W.K. Tung. Global qcd analysis of parton structure of the nucleon: Cteq5 parton distributions. *The European Physical Journal C - Particles and Fields*, 12:375–392, 2000.

[76] L.D. Landau and E.M. Lifshitz. *Quantum Electrodynamics*, volume IV. Elsevier, 2008.

[77] Mangano M., L., Moretti M., Piccinini F., and et al. Alpgen, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.

[78] M. Maggiore. *A Modern Introduction to Quantum Field Theory.* Oxford university press, 2005.

[79] A.D. Martin, R.G. Roberts, W.J. Stirling, and R.S. Thorne. Physical gluons and high E_T jets. *Phys. Lett. B*, (604):61–68, 2004.

[80] B.R. Martin and G. Shaw. *Particle Physics.* Wiley, 3^{rd} edition, 2008.

[81] Peter Minkowski. $\mu \to e\gamma$ at a rate of one out of 10^9 muon decays? *Physics Letters B*, 67(4):421 – 428, 1977.

[82] Rabindra N. Mohapatra and Goran Senjanović. Neutrino mass and spontaneous parity nonconservation. *Phys. Rev. Lett.*, 44:912–915, Apr 1980.

[83] Lorenzo Moneta, Kevin Belasco, Kyle S. Cranmer, S. Kreiss, Alfio Lazzaro, et al. The RooStats Project. *PoS*, ACAT2010:057, 2010.

[84] O. Panella, M. Cannoni, C. Carimalo, and Y.N. Srivastava. Signals of heavy Majorana neutrinos at hadron colliders. *Phys.Rev.*, D65:035005, 2002.

[85] Bruno. Pontecorvo. Mesonium and antimesonium. *Sov.Phys.JETP*, 6:429, 1957.

[86] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, P. Nadolsky, and W.K. Tung. New generation of parton distributions with uncertainties from global qcd analysis. *JHEP*, (0207):012, 2002.

[87] Andrei D. Sakharov. Violation of CP in variance, C asymmetry, and baryon asymmetry of the universe. *Soviet Physics Uspekhi*, 34:392, 1991.

[88] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. Pythia 6.4 physics and manual. *JHEP*, 05:026, 2006.

[89] T Sjstrand. Monte carlo generators. oai:cds.cern.ch:999717. (CERN-LCGAPP-2006-06):23 p, Nov 2006. arXiv: hep-ph/0611247.

[90] Jennifer A. Thomas and Patricia L. Vahle, editors. *Neutrino Oscillations*. Word Scientific, 2008.

[91] V. Trimble. Existence and nature of dark matter in the universe. *Annual Review of Astronomy and Astrophysics*, 25:425–472, 1987.

[92] M.R. Wahlley, D. Bourilkov, and R.C. Group. The les houches accord PDFs (LHAPDF) and LHAGLUE. 2005.

[93] S. Weinberg. *The Quantum Theory of Fields*. Cambridge university press, 2005.

[94] T. Yanagida. In O. Sawada and A. Sugamoto, editors, *Proceedings of the Workshop on the Unified Theory and the Baryon Number in the Universe*, page 95. National Labratory for High Energy Physics (KEK), 1979.

[95] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96:191–195, Oct 1954.

[96] Kai Zuber. *Neutrino physics*. IoP, 2004.